Analysis für Informatik

Prof. Michael Struwe

ETH Zürich

Vorwort

Wofür benötigen Informatikingenieure Kenntnisse in Analysis? Sollten nicht Logik, diskrete Mathematik und Kombinatorik ausreichen, um sämtliche Konzepte zu liefern, die relevant sind für den Umgang mit Maschinen, die konstruktionsbedingt nur endlich viele Zustände annehmen können? – Das Beispiel der komplexen Zahlen zeigt jedoch, dass bereits elementare Rechenoperationen wie die Bildung einer Quadratwurzel mehr als nur Programmierkenntnisse erfordern. Das Konzept der Konvergenz ist zentral für Anwendungen des Computers in numerischen Simulationen. Häufig führen auch elementare Fragestellungen der diskreten Mathematik auf schwierige Probleme der Analysis. So gelang Edmund Landau der Nachweis der Abschätzung $|A(R) - \frac{4\pi}{3}R^3| \leq CR^{3/2}$ für die Abweichung der Anzahl $A(R)$ von Punkten mit ganzzahligen Koordinaten innerhalb einer Kugel vom Radius R vom erwarteten Wert nur mit raffinierten Methoden der analytischen Zahlentheorie.

Zudem ist die moderne Informatik keine isolierte Disziplin; Teilgebiete wie die Computer Graphik erfordern ein Zusammenwirken von Informatikern mit Materialwissenschaftlern, Physikern und Mathematikern, wobei der Analysis eine wichtige Rolle zukommt. Verfahren wie der "Dielectric shader" entwerfen realistische Darstellungen von virtuellen Objekten mit brechenden Oberflächen und variabler optischer Dichte (z.B. ein halb gefülltes Glas Wasser), indem sie Absorbtion, Reflektion und Brechung der Lichtwellen aus den zugrundeliegenden physikalischen Gesetzen (Fresnel-Gleichungen, Snellsches Gesetz, Beersches Gesetz) herleiten. Die Lösung der von David Immel et al. sowie von James Kajiya im Jahre 1986 aufgestellten "rendering equation", einer Integralgleichung, ist ein anderer Ansatz zum Erzeugen realitätsnaher Bilder mittels Geometrischer Optik.

Natürlich können wir in dieser Vorlesung nicht im Detail auf derartige Anwendungen eingehen. Vielmehr werden Grundbegriffe und Konzepte bereitgestellt, die Voraussetzung sind für eine spätere Vertiefung dieser und weiterer Themen.

Das vorliegendende Skript entstand parallel zu meiner gleichnamigen Vorlesung im akademischen Jahr 2008/09. Das Skript wurde im Jahr 2009/10 überarbeitet, wobei auch das Erscheinungsbild verbessert wurde und zusätzliche Graphiken eingefügt wurden.

Ich danke Frau Manuela Dübendorfer für ihre Hilfe beim Erfassen meiner Vorlesungsunterlagen in LaTeX. Ebenso danke ich Frau Melanie Rupflin für eine Vielzahl von Anregungen und Hilfe beim Korrekturlesen des Skripts. Auch den

Studierenden dieser Jahrgänge danke ich für anregende Kommentare und Korrekturhinweise, vor allem Herrn Simon Eugster, der diese Hinweise im akademischen Jahr 2008/09 gesammelt und weitergereicht hat, und Herrn Jorim Jaggi, der mir im Sommer 2010 bei der Verbesserung der Gestaltung des Skripts geholfen und die Graphiken erstellt hat.

Zürich, 5.11.2010 *Michael Struwe*

Inhaltsverzeichnis

1 Logik und Grundlagen **3**

1.1 Logik 3

1.2 Mengenlehre 5

1.3 Funktionen 7

2 Zahlen und Vektoren **11**

2.1 Elementare Zahlen 11

2.2 Die reellen Zahlen 12

2.3 Supremum und Infimum 16

2.4 Der euklidische Raum 19

2.5 Komplexe Zahlen 22

3 Folgen und Reihen **27**

3.1 Beispiele 27

3.2 Grenzwert einer Folge 27

3.3 Konvergenzkriterien 30

3.4 Teilfolgen, Häufungspunkte 34

3.5 Cauchy-Kriterium 37

3.6 Folgen in $\mathbb{R}^d$ oder $\mathbb{C}$ 38

3.7 Reihen 40

3.8 Absolute Konvergenz 45

3.9 Die Exponentialreihe und die Funktion e^x 48

4 Stetigkeit **51**

4.1 Grenzwerte von Funktionen 51

4.2 Stetige Funktionen 56
4.3 Ein wenig Topologie 58
4.4 Äquivalente Normen 63
4.5 Topologisches Kriterium für Stetigkeit 66
4.6 Zwischenwertsatz und Folgerungen 68
4.7 Supremumsnorm 73
4.8 Punktweise und gleichmässige Konvergenz 75

5 Differentialrechnung auf $\mathbb{R}$ 79
5.1 Differential und Differentiationsregeln 79
5.2 Der Mittelwertsatz und Folgerungen 83
5.3 Die trigonometrischen Funktionen 89
5.4 Funktionen der Klasse C^1 91
5.5 Taylor-Formel 96
5.6 Gewöhnliche Differentialgleichungen 103
5.7 Inhomogene Differentialgleichungen 111

6 Integration 117
6.1 Stammfunktionen 117
6.2 Das Riemannsche Integral 127
6.3 Integrationsregeln, Hauptsatz 135
6.4 Uneigentliches Riemann-Integral 140
6.5 Differentialgleichungen 142

7 Differentialrechnung im $\mathbb{R}^n$ 153
7.1 Partielle Ableitungen und Differential 153
7.2 Differentiationsregeln 158
7.3 Differentialformen und Vektorfelder 164
7.4 Wegintegrale 166
7.5 Höhere Ableitungen 171
7.6 Vektorwertige Funktionen 177
7.7 Der Umkehrsatz 179
7.8 Implizite Funktionen 184
7.9 Extrema mit Nebenbedingungen 189

8 Integration im $\mathbb{R}^n$ **195**

8.1 Riemannsches Integral über einem Quader 195

8.2 Der Satz von Fubini 198

8.3 Jordan-Bereiche 201

8.4 Der Satz von Green 205

8.5 Substitutionsregel 210

8.6 Oberflächenmass und Fluss-Integral 216

8.7 Der Satz von Stokes im $\mathbb{R}^3$ 220

8.8 Der Satz von Gauss 224

Kapitel 1

Logik und Grundlagen

1.1 Logik

Beispiele für mathematische **Aussagen**:

i) "$4 > 2$" (wahr)

ii) "$\forall n \in \mathbb{N} : n > 4 \to n > 2$" (wahr)

iii) "$5 < 3$" (falsch)

In der Mathematik stützen wir uns auf gewisse Grundannahmen *"Axiome"* , die wir als gegeben ansehen. Eine dieser Annahmen ist der folgende Satz über die möglichen Wahrheitswerte von Aussagen.

Satz vom ausgeschlossenen Dritten (Tertium non datur): Eine zulässige mathematische Aussage ist entweder wahr oder falsch, jedoch nie beides zugleich.

Bemerkung 1.1.1. i) Dieses Axiom ist eine *mathematische Abstraktion*, wir bewegen uns in einer künstlichen Welt. In der wirklichen Welt gibt es Graustufen, zum Beispiel hängt der Wahrheitswert der Aussage "Das Wetter ist schön" vom subjektiven Befinden ab.

ii) Nicht alle Aussagen sind zulässig. Die rückbezügliche Aussage "Diese Aussage ist falsch." ist weder falsch (dann wäre sie wahr) noch wahr (dann wäre sie falsch). Analog: "Ich lüge jetzt." Aber: "Ich lüge immer" könnte falsch sein, falls ich je mal die Wahrheit gesagt habe.

Die Axiome der Logik sind insofern unvollständig. Wir werden dies aber niemals als Einschränkung empfinden.

Mit Aussagen kann man "rechnen". Es seien A, B mathematische Aussagen. Die Negation $(\neg A)$, "und" $(A \wedge B)$, "oder" $(A \vee B)$, die Implikation $(A \to B)$ und die Äquivalenz $(A \leftrightarrow B)$ sind definiert durch die **Wahrheitstafel**.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \to B$	$A \leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Beispiel 1.1.1. i) "$(n > 4) \quad \to \quad (n > 2)$". *Beachte:* Weder die Annahme (Voraussetzung) "$n > 4$" noch die Folgeaussage "$n > 2$" ist für alle $n \in \mathbb{N}$ erfüllt, die Implikation gilt jedoch immer.

ii) In der Politik macht man sich dies gern zunutze: Die Aussage "Wenn das Volk damals anders entschieden hätte, dann ..." ist bei beliebiger Fortsetzung wahr (Conjunctivus irrealis).

Die Implikation $A \to B$ ist die für den Aufbau der Mathematik wichtigste Verknüpfung.

Eine **wahre** Implikation $A \to B$ bezeichnen wir auch als "Folgerung" und schreiben $A \Rightarrow B$. (*"A ist hinreichend für B, "wenn A, dann B"*)

Bemerkung 1.1.2. Die Implikation ist transitiv:

$$(A \to B) \wedge (B \to C) \;\Rightarrow\; (A \to C)$$

Wir können daher über eine Kette von Folgerungen

$$A \Rightarrow B \Rightarrow \cdots \Rightarrow S$$

einen mathematischen "Satz" S aus einer "Annahme" A herleiten. (**Prinzip des mathematischen Beweises**).

Äquivalenz: Anstelle von $(A \to B) \wedge (B \to A)$ schreiben wir $A \leftrightarrow B$.
Anstelle von $(A \Rightarrow B) \wedge (B \Rightarrow A)$ schreiben wir $A \Leftrightarrow B$; in diesem Fall ist also die Aussage A wahr genau dann, wenn B wahr ist.

Kontraposition (Umkehrschluss): Falls $A \Rightarrow B$, so kann A nicht wahr sein, wenn B falsch ist. (*"B ist notwendig für A."*) Die Aussage "$A \Rightarrow B$" ist somit gleichbedeutend mit "$(\neg B) \Rightarrow (\neg A)$":

$$(A \to B) \;\Leftrightarrow\; (\neg B \to \neg A).$$

Prinzip des indirekten Beweises: Zum Beweis der Aussage $A \Rightarrow B$ genügt es, die Aussage $(\neg B) \Rightarrow (\neg A)$ zu zeigen, oder die Annahme $A \wedge (\neg B)$ zum Widerspruch zu führen.

Beispiel 1.1.2. Es seien A die üblichen Axiome über $\mathbb{N}$, B die Aussage:

"Es gibt keine grösste natürliche Zahl."

Wir zeigen: $A \Rightarrow B$.

Beweis (indirekt). Nimm an, es gibt ein maximales $n_0 \in \mathbb{N}$; das heisst, $n_0 \geq l$ für jedes $l \in \mathbb{N}$. Nach einem der Axiome für $\mathbb{N}$ hat n_0 jedoch einen Nachfolger $n_0 + 1 \in \mathbb{N}$, und $n_0 + 1 > n_0$. Widerspruch! □

Auf den Eigenschaften der natürlichen Zahlen beruht ein weiteres Beweisprinzip, das **Prinzip der vollständigen Induktion**: Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Weiter gelte $A(1)$, und für jedes $n \in \mathbb{N}$ gelte $A(n) \Rightarrow A(n+1)$. Dann gilt $A(n)$ für jedes $n \in \mathbb{N}$, denn mit der Kette

$$A(1) \Rightarrow A(2) \Rightarrow \cdots \Rightarrow A(n-1) \Rightarrow A(n)$$

erhalten wir $A(n)$ in endlich vielen Schritten aus $A(1)$.

Beispiel 1.1.3. Für jedes $n \in \mathbb{N}$ gilt

$$1 + 3 + 5 + \cdots + (2n-1) = \sum_{k=1}^{n} (2k-1) = n^2 \,.$$

Beweis (vollständige Induktion). Der Beweis besteht aus zwei Teilen:

Induktions-Verankerung $(n = 1)$**:** $1 = 1^2$.

Induktions-Schluss $(n \to n+1)$**:** Nach Induktionsannahme gilt

$$\underbrace{1 + \cdots + (2n-1)}_{=n^2} + \underbrace{(2(n+1)-1)}_{=2n+1} = n^2 + 2n + 1 = (n+1)^2.$$

□

1.2 Mengenlehre

Nach Georg Cantor ist eine **Menge** die "ungeordnete Zusammenfassung verschiedener Objekte (sogenannter 'Elemente') zu einem Ganzen."

Beispiel 1.2.1. o) Für $a \neq b$ gilt $\{a, b\} = \{b, a\} = \{a, b, a\}$,

i) $\mathbb{N} = \{1, 2, 3, \ldots\}$,

ii) $\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\} = \mathbb{N} \cup \{0\}$,

iii) $\{n \in \mathbb{N}; n \text{ teilt } 15\} = \{1, 3, 5, 15\}$,

iv) $\emptyset = \{\}$: die **leere Menge**.

Wie bei Aussagen müssen wir jedoch rückbezügliche Definitionen vermeiden:

Beispiel 1.2.2. (Bertrand Russel) Die "Menge M aller Mengen, die sich selbst nicht als Element enthalten" gibt es nicht.

(Sonst müsste eine der Aussagen $M \in M$ oder $M \notin M$ gelten. Jedoch führt die Annahme $M \in M$ nach Definition von M zum Widerspruch $M \notin M$, während die Annahme $M \notin M$ zum Widerspruch $M \in M$ führt).

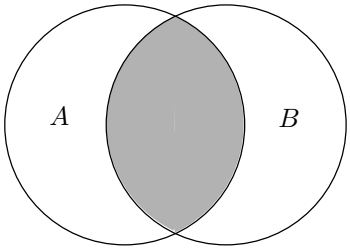
Das Russelsche Beispiel lässt sich leicht in die Alltagssprache übersetzen: Definiert man den Dorfbarbier als den Mann, der alle Männer rasiert, die sich nicht selbst rasieren, so kommt man auf analoge Weise zu einem Widerspruch.

Mengenoperationen. Die folgenden Verknüpfungen sind für beliebige Mengen A, B erklärt:

$$A \cup B := \{x; x \in A \vee x \in B\}, \textbf{ Vereinigungsmenge},$$

$$A \cap B := \{x; x \in A \wedge x \in B\}, \textbf{ Durchschnitt},$$

$$A \backslash B := \{x \in A;\ x \notin B\}, \textbf{ Differenzmenge}$$



Zudem sind für Mengen A, B, X die folgenden Relationen erklärt:

$A \subset X$:	Teilmenge
$X \backslash A =: A^c$	Komplement von A in einer (festen) Grundmenge X.
$A = B$:	falls A und B dieselben Elemente enthalten.

Beispiel 1.2.3. Für $A, B \subset X$ gilt

$$(A \cap B)^c = A^c \cup B^c.$$

Beweis. Für $x \in X$ gilt

$$x \in (A \cap B)^c \Leftrightarrow x \notin A \cap B \Leftrightarrow x \notin A \vee x \notin B$$
$$\Leftrightarrow x \in A^c \vee x \in B^c \Leftrightarrow x \in A^c \cup B^c\,.$$

Vgl. Übung 1.4. □

Wir können **Quantoren** benutzen, um Aussagen über Elemente einer Menge zu machen:

$\forall$: der Allquantor ("für alle"),

$\exists$: der Existenzquantor ("es gibt").

Beispiel 1.2.4. i) $\forall n \in \mathbb{N} : n > 0$ (wahr).

ii) $\exists n_0 \in \mathbb{N} \ \forall k \in \mathbb{N} : \ k \leq n_0$. (Dies ist die (falsche) Aussage: "Es gibt eine grösste natürliche Zahl $n_0 \in \mathbb{N}$", siehe Beispiel 1.1.2.)

iii) $\forall n_0 \in \mathbb{N} \ \exists k \in \mathbb{N} : \ k > n_0$. (Diese (wahre) Aussage ist die Verneinung von ii).)

Im Beispiel 1.2.4 erkennen wir folgende Regeln für die **Verneinung** von Aussagen mit Quantoren:

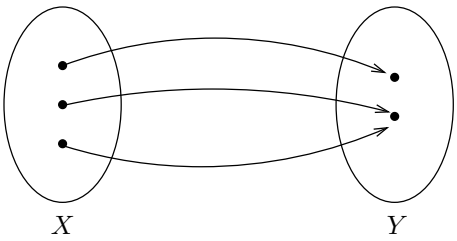
$$\neg \big(\forall n \in \mathbb{N} : A(n)\big) \Leftrightarrow \exists n \in \mathbb{N} : \ \neg A(n),$$
$$\neg \big(\exists n \in \mathbb{N} : A(n)\big) \Leftrightarrow \forall n \in \mathbb{N} : \ \neg A(n).$$

1.3 Funktionen

In der Schule haben wir **Funktionen** oder **Abbildungen** in der Form von Zuordnungsvorschriften $y = f(x)$ für reelle Zahlen kennengelernt, z.B.

$$y = f(x) = x - x^3, \ -1 \leq x \leq 1.$$

Allgemein betrachten wir im folgenden Abbildungen $f : X \to Y$ zwischen beliebigen Mengen X und Y, welche jedem $x \in X$ genau ein "Bild" $y = f(x) \in Y$ zuordnen. Jedes $z \in X$ mit $y = f(z)$ heisst dann ein "Urbild" von y. Die Begriffe "Funktion" und "Abbildung" verwenden wir synonym.



Somit ist eine Funktion erklärt durch Angabe

- des **Definitionsbereiches** (hier X)
- des **Bild- oder Wertebereiches** (hier Y)
- der **Abbildungsvorschrift** ($x \mapsto f(x)$)

Beispiel 1.3.1. i) $f : [-1, 1] \to \mathbb{R}$, $x \mapsto x - x^3$

ii) $g : \mathbb{R} \to [-1, 1]$, $x \mapsto \sin(x)$

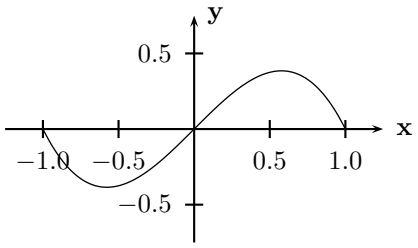
iii) $h : \mathbb{R} \to [0, \infty[$, $x \mapsto x^2$

iv) $id_X : X \to X$, $x \mapsto x = id_X(x)$: Identität.

Wir können Funktionen $f : D(f) \subseteq \mathbb{R} \to \mathbb{R}$ durch ihren **Graphen** darstellen

$$\mathcal{G}(f) = \big\{\big(x, f(x)\big);\ x \in \mathcal{D}(f)\big\} \subset \mathbb{R} \times \mathbb{R}\,.$$

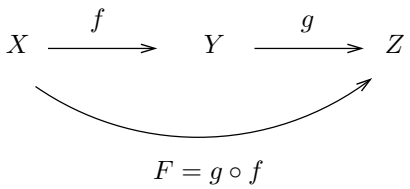
Beispiel 1.3.2. $f : [-1.1] \to \mathbb{R}$, $f(x) = x - x^3$.



Dies geht auch allgemein (jedoch abstrakt).

Komposition. Abbildungen $f : X \to Y$, $g : Y \to Z$ kann man hintereinander ausführen. Dies ergibt eine neue Abbildung

$$F := g \circ f : X \to Z,\ \ x \mapsto g\big(f(x)\big)\,,$$



Diese Komposition ist **assoziativ**: Für $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ gilt

$$F_1 := h \circ (g \circ f) = (h \circ g) \circ f =: F_2 : X \to W.$$

Die Definitionsbereiche von F_1 und F_2 sind nämlich offenbar dieselben ($= X$), ebenso die Wertebereiche ($= W$), und für jedes $x \in X$ gilt

$$\begin{aligned} F_1(x) &= h\big((g \circ f)(x)\big) = h\big(g(f(x))\big) \\ &= (h \circ g)\,\big(f(x)\big) = F_2(x)\,. \end{aligned}$$

Z.B. ergibt für f, g, h aus Beispiel 1.3.1 und $x = 1$ die Rechnung

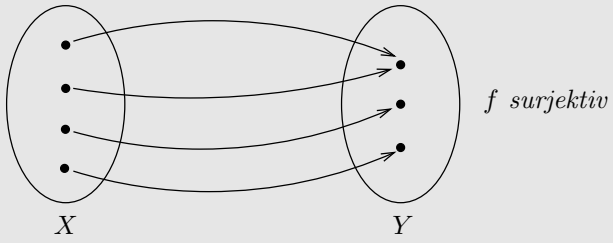
$$(h \circ g \circ f)(1) = (sin(x - x^3))^2\big|_{x=1} = sin^2(0) = 0$$

.

Definition 1.3.1. *Sei $f : X \to Y$ eine Abbildung.*

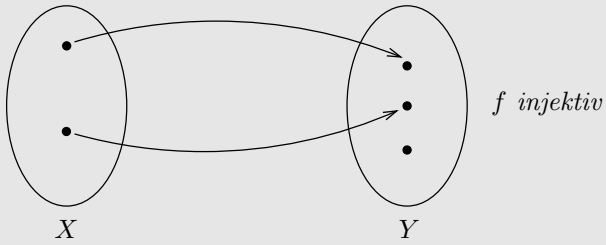
i) f heisst **surjektiv**, *falls jedes $y \in Y$* **mindestens** *ein Urbild hat; d.h., falls*

$$\forall y \in Y \; \exists x \in X : \; f(x) = y \, .$$

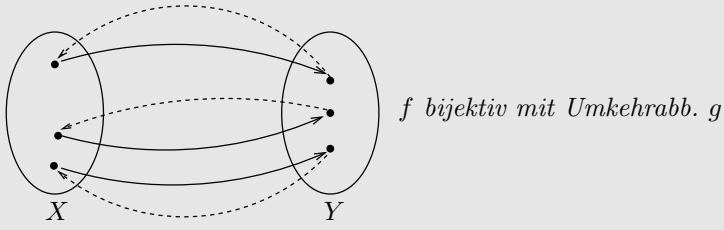


ii) f heisst **injektiv**, *falls jedes $y \in Y$* **höchstens** *ein Urbild hat, d.h. falls*

$$\forall x_1, x_2 \in X : \; f(x_1) = f(x_2) \Rightarrow x_1 = x_2 \, .$$



iii) f heisst **bijektiv**, *falls jedes $y \in Y$* **genau** *ein Urbild hat, d.h. falls f sowohl injektiv als auch surjektiv ist.*



Falls f bijektiv (und nur in diesem Fall), können wir eine Abbildung $g : Y \to X$ einführen, welche jedem $y \in Y$ das eindeutig bestimmte Urbild $x \in X$ unter f zuordnet, mit

$$g \circ f = id_X, \quad f \circ g = id_Y \, .$$

Dieses g heisst die **Umkehrabbildung** von f, $g = f^{-1}$.

Andererseits kann man bei jeder Abbildung $f : X \to Y$ zu jeder Teilmenge $B \subset Y$ deren **Urbild** $f^{-1}(B) \subset X$ betrachten mit

$$f^{-1}(B) := \{x \in X; \; f(x) \in B\}.$$

Beispiel 1.3.3. Sei $f : [-1, 1] \to \mathbb{R}$, $x \mapsto x - x^3$, und sei $B = \{0\}$. Dann gilt

$$f^{-1}(B) = \{x \in [-1, 1]; f(x) = 0\} = \{-1, 0, 1\}.$$

Falls f bijektiv mit Umkehrabbildung $g = f^{-1} : Y \to X$, so gilt offenbar für jedes $y \in Y$

$$f^{-1}(\{y\}) = \{f^{-1}(y)\},$$

wobei f^{-1} im 1. Ausdruck die Urbildfunktion, im 2. Ausdruck die Umkehrabbildung bezeichnet.
Allgemein ist f bijektiv genau dann, wenn für jedes $y \in Y$ das Urbild $f^{-1}(\{y\})$ genau ein Element enthält.

Kapitel 2

Zahlen und Vektoren

2.1 Elementare Zahlen

Mit den **natürlichen Zahlen**

$$\mathbb{N} = \{1, 2, 3, \dots\},\ \mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$$

kann man Objekte abzählen. Zahlen in $\mathbb{N}$ kann man addieren und multiplizieren. In den **ganzen Zahlen**

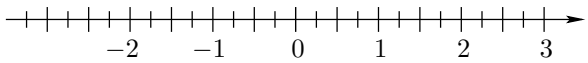
$$\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$$

ist zusätzlich die Subtraktion möglich. In den **rationalen Zahlen**

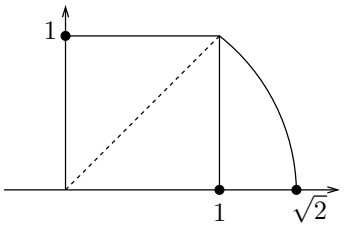
$$\mathbb{Q} = \left\{ \frac{p}{q}\ ; \quad p, q \in \mathbb{Z},\ q > 0 \right\}$$

kann man zudem (ausser durch 0) dividieren: $\mathbb{Q}$ ist ein **Zahlkörper**.

Offenbar kann man diese **elementaren Zahlen** $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$ der Grösse nach auf dem **Zahlenstrahl** anordnen.



Irrationale Zahlen. Zwischen je zwei rationalen Zahlen $r_1 < r_2$ liegt eine weitere, z.B. die Zahl $\frac{r_1+r_2}{2} \in \mathbb{Q}$, welche den halben Abstand zu r_1 hat wie r_2; die rationalen Zahlen liegen somit dicht auf der Zahlengeraden. Jedoch erkannten bereits die Pythagoräer, dass die Länge der Diagonalen im Einheitsquadrat durch kein $r \in \mathbb{Q}$ dargestellt wird.



Satz 2.1.1. *Es gibt keine Zahl $r \in \mathbb{Q}$ mit $r^2 = 2$.*

Beweis (indirekt).. Nimm an, es gibt $r = \frac{p}{q} \in \mathbb{Q}$ mit $r^2 = 2$. Nach Kürzen gemeinsamer Teiler dürfen wir annehmen, dass p, q teilerfremd sind (keine gemeinsamen Teiler haben), und $p, q > 0$.

Aus der Gleichung $r^2 = \frac{p^2}{q^2} = 2$ folgt nach Multiplikation mit q^2 zunächst

$$p^2 = 2 \cdot q^2 \,.$$

Da die Zahl 2 prim ist, enthält p den Teiler 2; es gilt also $p = 2s$ für ein $s \in \mathbb{N}$ und somit

$$2 \cdot q^2 = p^2 = 2^2 \cdot s^2.$$

Nach Kürzen des Faktors 2 erhalten wir

$$q^2 = 2 \cdot s^2,$$

und wie oben folgt $q = 2t$ für ein $t \in \mathbb{N}$. Die Zahl 2 teilt also sowohl p als auch q im Widerspruch zu unserer Annahme, dass p und q teilerfremd sind. □

$\mathbb{Q}$ weist also "Lücken" auf. Wir können jedoch $\mathbb{Q}$ erweitern zum Körper $\mathbb{R}$ der **reellen Zahlen**, der die Zahlengerade "lückenlos" überdeckt. Dies gelingt z.B. mit dem Begriff des "Dedekindischen Schnittes" oder über "Fundamentalfolgen". Die Zahlengerade ist ein geometrisches Modell für $\mathbb{R}$. Wir überspringen hier jedoch die entsprechende Konstruktion und nehmen $\mathbb{R}$ als gegeben an.

2.2 Die reellen Zahlen

Wichtig für das folgende sind die für das Rechnen mit reellen Zahlen geltenden Regeln, die **Axiome** für $\mathbb{R}$, die wir im folgenden aufführen.

Es gibt eine Operation, genannt **Addition:** $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x + y$, auf $\mathbb{R}$ mit den Eigenschaften:

A.i) Assoziativität: $\forall x, y, z \in \mathbb{R} : x + (y + z) = (x + y) + z$,

A.ii) Neutrales Element: $\exists\, 0 \in \mathbb{R} \quad \forall x \in \mathbb{R}$: $x + 0 = x$,

A.iii) Inverses Element: $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R} :\ x + y = 0$,

A.iv) Kommutativität: $\forall x, y \in \mathbb{R} :\ x + y = y + x$.

D.h. $\mathbb{R}$ bildet eine Abelsche (kommutative) **Gruppe** bezüglich der Addition.

Bemerkung 2.2.1. Das zu $x \in \mathbb{R}$ inverse Element $y = -x$ ist eindeutig bestimmt.

Beweis. Falls y und z zu x invers, so folgt

$$z \overset{(A.ii))}{=} z + \underbrace{(x+y)}_{=0} \overset{(A.i),iv))}{=} \underbrace{(x+z)}_{=x+z=0} + y \overset{(A.iv))}{=} y + 0 \overset{(A.ii))}{=} y\,.$$

□

Es gibt eine weitere Operation, genannt **Multiplikation:** $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(x, y) \mapsto x \cdot y = xy$, auf $\mathbb{R}$ mit den Eigenschaften:

M.i) Assoziativität: $\forall x, y, z \in \mathbb{R} : x \cdot (y \cdot z) = (x \cdot y) \cdot z$,

M.ii) Neutrales Element: $\exists 1 \in \mathbb{R}\backslash\{0\}\ \forall x \in \mathbb{R}$: $x \cdot 1 = x$,

M.iii) Inverses Element: $\forall x \in \mathbb{R}\backslash\{0\}\ \exists y \in \mathbb{R} :\ x \cdot y = 1$,

M.iv) Kommutativität: $\forall x, y \in \mathbb{R} :\ x \cdot y = y \cdot x$.

Die Multiplikation ist verträglich mit der Addition wegen dem **Distributivitäts-Gesetz**

D) $\forall x, y, z \in \mathbb{R} :\ x \cdot (y + z) = x \cdot y + x \cdot z$.

Bemerkung 2.2.2. i) $\forall x \in \mathbb{R}$: $x \cdot 0 = 0$.

ii) $\forall x, y \in \mathbb{R} :\ x \cdot y = 0 \Rightarrow x = 0$ oder $y = 0$.

Beweis. i) $x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0$. Addiere $-(x \cdot 0)$!

ii) Falls $x \cdot y = 0$, wobei $x \neq 0$ mit multiplikativ Inversem x^{-1}, so folgt

$$y = \underbrace{(x^{-1} \cdot x)}_{=1} \cdot y = x^{-1} \underbrace{(x \cdot y)}_{=0} = 0\,.$$

□

Wegen Bemerkung 2.2.2 bildet auch $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ bezüglich der Multiplikation eine abelsche Gruppe.

Zudem gibt es auf $\mathbb{R}$ eine **Ordnung** $\leq$ mit den folgenden Eigenschaften:

O.i) Reflexivität: $\forall x \in X$, $x \leq x$,

O.ii) Transitivität: $\forall x, y, z \in \mathbb{R}$: $x \leq y \wedge y \leq z \Rightarrow x \leq z$,

O.iii) Identitivität: $\forall x, y \in \mathbb{R}$: $x \leq y \wedge y \leq x \Rightarrow x = y$,

O.iv) Die Ordnung ist total: $\forall x, y \in \mathbb{R}$: $x \leq y$ oder $y \leq x$.

Die Ordnung ist **konsistent** mit Addition und Multiplikation:

K.i) $\forall x, y, z \in \mathbb{R} :\ x \leq y \Rightarrow x + z \leq y + z$

K.ii) $\forall x, y, z \in \mathbb{R} :\ x \leq y,\ 0 \leq z \Rightarrow x \cdot z \leq y \cdot z$.

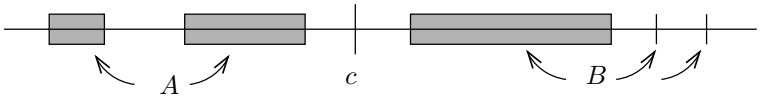
Die reellen Zahlen bilden somit einen linear geordneten Zahlkörper mit den Operationen Addition und Multiplikation. Diese Eigenschaft und die entsprechenden Axiome A.i) - iv), M.i) - iv), D, O.i) - iv), K.i) - ii) gelten bereits in $\mathbb{Q}$. Die entscheidende weitere Eigenschaft von $\mathbb{R}$ ist das **Vollständigkeitsaxiom**:

V) $\mathbb{R}$ ist **ordnungsvollständig**: Zu je zwei nicht leeren Mengen $A, B \subset \mathbb{R}$ mit

$$a \leq b \text{ für alle } a \in A, b \in B$$

gibt es ein $c \in \mathbb{R}$, sodass gilt

$$a \leq c \leq b, \quad \forall a \in A,\ b \in B\,.$$



Einige elementare Folgerungen aus den Axiomen:

Folgerung 2.2.1. i) $\forall x \in \mathbb{R} : (-1) \cdot x = -x$.

Beweis. Es gilt

$$x + (-1) \cdot x \overset{(M.ii)}{=} 1 \cdot x + (-1) \cdot x \overset{D}{=} \big(1 + (-1)\big) \cdot x = 0 \cdot x = 0.$$

Da das additiv Inverse zu x nach Bemerkung 2.2.1 eindeutig bestimmt ist, folgt die Behauptung. □

ii) $(-1) \cdot (-1) = 1$.

Beweis. Spezialfall von i), da mit $(-1) + 1 = 0$ folgt $1 = -(-1)$. Setze nun $x = -1$ in i). □

iii) $\forall x \in \mathbb{R}$: $x^2 \geq 0$.

Beweis. Sei $x \in \mathbb{R}$ beliebig gewählt. Mit O.iv) gilt $x \geq 0$ oder $x \leq 0$.

a) $x \geq 0$. Mit K.ii) folgt $x^2 \geq 0 \cdot x = 0$.

b) $x \leq 0$. Mit K.i) folgt $-x \geq 0$, und mit i) und ii) sowie a) folgt

$$0 \leq (-x)^2 = \big((-1) \cdot x\big)^2 = (-1)^2 \cdot x^2 = x^2\,.$$

□

iv) $0 < 1 < 2 < \ldots$.

Beweis. $1 \overset{ii)}{=} (-1)^2 \overset{iii)}{\geq} 0$, und $1 \neq 0$ nach M.ii). Also ist $0 < 1$ und mit K.i) folgt die Behauptung. □

v) $\forall x > 0$: $x^{-1} > 0$.

Beweis. Annahme $x^{-1} \leq 0$. Nach Multiplikation mit $x > 0$ folgt

$$1 = x^{-1} \cdot x \leq 0 \cdot x = 0$$

im Widerspruch zu iv). □

vi) $\forall x, y \geq 0$: $x \leq y \Leftrightarrow x^2 \leq y^2$.

Beweis. Ohne Beschränkung der Allgemeinheit gelte $x + y > 0$. (Sonst $x = y = 0 = x^2 = y^2$.) Beachte:

$$y^2 - x^2 = \underbrace{(y + x)}_{>0}(y - x).$$

"$\Rightarrow$": Sei $y \geq x$, also $y - x \geq 0$. Mit K.ii) folgt $y^2 \geq x^2$.

"$\Leftarrow$": Nach v) gilt $(y + x)^{-1} > 0$. K.ii) liefert Behauptung. □

vii) Es gibt $c \in \mathbb{R}$ mit $c^2 = 2$.

Beweis. Setze

$$A = \{a \in [1, 2];\ a^2 < 2\}, \quad B = \{b \in [1, 2];\ b^2 \geq 2\}\ .$$

Dann gilt offenbar $1 \in A$, $2 \in B$; also $A \neq \emptyset \neq B$. Weiter folgt mit vi)

$$a < b \quad \forall a \in A, b \in B.$$

Das Vollständigkeitsaxiom V liefert somit eine Zahl $c \in \mathbb{R}$ mit

$$a \leq c \leq b \quad \forall a \in A, b \in B;$$

insbesondere folgt sofort $1 \leq c \leq 2$.

Wir zeigen, dass $c^2 = 2$. Andernfalls gilt nach dem Ordnungsaxiom O.iv) entweder a) $c^2 < 2$ oder b) $c^2 > 2$. Im Falle a) gibt es $0 < \epsilon \leq 1/5$ mit $c^2 = 2 - 5\epsilon$. Für $a := c + \epsilon$ erhalten wir

$$a^2 = c^2 + 2\epsilon c + \epsilon^2 < c^2 + 5\epsilon = 2\ ;$$

also $a \in A$ im Widerspruch zur Trennungseigenschaft von c. Analog erhalten wir im Falle b) eine Zahl $0 < \epsilon \leq 2/5$ mit $c^2 = 2 + 5\epsilon$. Für $b := c - \epsilon$ folgt dann

$$b^2 = c^2 - 2\epsilon c + \epsilon^2 > c^2 - 4\epsilon > 2;$$

also $b \in B$. Erneut ergibt sich ein Widerspruch zur Trennungseigenschaft von c, und es bleibt nur die Möglichkeit $c^2 = 2$. □

Bemerkung 2.2.3. Es gilt $A, B \subset \mathbb{Q}$; die Mengen A und B werden aber durch kein $c \in \mathbb{Q}$ getrennt. Wie wir oben gesehen haben, ist das die Mengen A, B trennende $c \in \mathbb{R}$ nämlich eindeutig bestimmt, und es erfüllt $c^2 = 2$, gehört nach Satz 2.1.1 also nicht zu $\mathbb{Q}$. Der Körper $\mathbb{Q}$ ist daher nicht ordnungsvollständig.

Definition 2.2.1. *Der* **Absolutbetrag** *einer Zahl $x \in \mathbb{R}$ ist die Zahl*

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

Offenbar gilt $|x| \geq 0$ für alle x. Weiter hat der Absolutbetrag die Eigenschaften

viii) $x \leq |x|, \quad \forall x \in X$

ix) $|xy| = |x|\,|y|, \quad \forall x, y \in \mathbb{R}.$

Satz 2.2.1. *(Dreiecks-Ungleichung). Es gilt*

$$|x + y| \leq |x| + |y|, \quad \forall x, y \in \mathbb{R}\,.$$

Beweis. Mit vi) folgt die Behauptung aus

$$\begin{aligned} |x+y|^2 &\overset{ii)}{=} (x+y)^2 = x^2 + 2xy + y^2 \\ &\overset{viii),ix)}{\leq} |x|^2 + 2|x|\,|y| + |y|^2 = (|x| + |y|)^2\,. \end{aligned}$$

□

Satz 2.2.2. *(Young) Für $x, y \in \mathbb{R}$, $\epsilon > 0$ gilt*

$$2|x \cdot y| \leq \epsilon x^2 + \frac{1}{\epsilon}\, y^2\,.$$

Beweis. Setze $\delta = \sqrt{\epsilon} > 0$. Ohne Beschränkung der Allgemeinheit gelte $x \cdot y \geq 0$. Die Behauptung folgt aus

$$0 \leq \left(\delta x - \frac{y}{\delta}\right)^2 = \delta^2 x^2 - 2x \cdot y + \frac{1}{\delta^2}\, y^2.$$

□

2.3 Supremum und Infimum

Definition 2.3.1. *Eine Menge $A \subset \mathbb{R}$ heisst* **nach oben beschränkt**, *falls gilt*

$$\exists b \in \mathbb{R}\ \forall a \in A:\ a \leq b.$$

Jedes derartige b heisst eine **obere Schranke** *für A. (Analog:* **nach unten beschränkt**, **untere Schranke**.*)*

Beispiel 2.3.1. Das Intervall

$$]-1, 1[= \{x \in \mathbb{R};\ -1 < x < 1\}$$

ist nach oben (z.B. durch $b = 1$) und unten (z.B. durch $a = -1$) beschränkt.

Sei nun $\emptyset \neq A \subset \mathbb{R}$ nach oben beschränkt,

$$B = \{b \in \mathbb{R};\ b \text{ ist obere Schranke für } A\}.$$

Dann gilt $B \neq \emptyset$, und

$$a \leq b \qquad \text{für alle } a \in A,\ b \in B.$$

Mit dem Vollständigkeitsaxiom folgt die Existenz einer Zahl $c \in \mathbb{R}$ mit

$$a \leq c \leq b \qquad \text{für alle } a \in A,\ \ b \in B.$$

Offenbar ist c obere Schranke für A; also $c \in B$. Da zugleich gilt $c \leq b$ für alle $b \in B$, ist c die **kleinste obere Schranke** für A. Hierdurch ist c eindeutig bestimmt.

Satz 2.3.1. i) *Jede nicht leere, nach oben beschränkte Menge $A \subset \mathbb{R}$ besitzt eine kleinste obere Schranke $c =: \sup A$, das* **Supremum** *von A.*

ii) *Analog besitzt jede nicht leere, nach unten beschränkte Menge $A \subset \mathbb{R}$ eine grösste untere Schranke $c' = \inf A$, das* **Infimum** *von A.*

Beispiel 2.3.2. i) Sei $A =]-1, 1[\subset \mathbb{R}$. Dann gilt

$$\sup A = 1, \qquad \inf A = -1.$$

ii) Sei $A = [-1, 1] = \{x \in \mathbb{R};\ \ -1 \leq x \leq 1\}$. Dann gilt

$$\begin{aligned} \sup A = 1 = \max A: &\quad \text{das } \textbf{Maximum} \text{ von } A, \\ \inf A = -1 = \min A: &\quad \text{das } \textbf{Minimum} \text{ von } A. \end{aligned}$$

iii) Besteht die Menge A aus nur endlich vielen Elementen $a_1 < a_2 < \cdots < a_k$, so gilt offenbar

$$\sup A = a_k = \max A.$$

Die Beispiele zeigen, dass $\sup A$, $\inf A$ im allgemeinen nicht zur Menge A gehören. Gehört $\sup A$ jedoch zu A, so sagen wir, **"das Supremum wird in A angenommen"**, und wir schreiben $\sup A = \max A$.

In diesem Fall gehört $c = \max A$ sowohl zu A als auch zu B, der Menge der oberen Schranken. In der die Zahl c charakterisierenden Beziehung

$$a \leq c \leq b \qquad \text{für alle } a \in A,\ b \in B$$

ist also auf beiden Seiten Gleichheit nicht ausgeschlossen.

Falls $\inf A \in A$ sagen wir analog **"das Infimum wird in A angenommen"** und schreiben $\inf A = \min A$.

Beispiel 2.3.3. i) Sei $A \subset \mathbb{R}$ die Menge

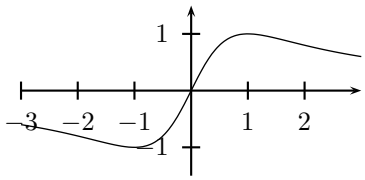
$$A = \Big\{\frac{2x}{1+x^2};\ \ x \in \mathbb{R}\Big\}.$$

Behauptung: $\sup A = 1$.

Beweis.

$$1 - \frac{2x}{1+x^2} = \frac{1+x^2-2x}{1+x^2} = \frac{(1-x)^2}{1+x^2} \geq 0$$

und Gleichheit gilt, falls $x = 1$. □



Der Beweis zeigt, dass $\sup A$ für $x = 1$ sogar angenommen wird, und zwar in der Maximalstelle $x = 1$ der Funktion

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

ii) Sei $A = \{\arctan x;\ x \in \mathbb{R}\}$. Dann gilt

$$\sup A = \pi/2, \qquad \inf A = -\pi/2,$$

und diese Werte werden nicht angenommen.

Als weitere Folgerung aus dem Axiom V ergibt sich, dass jede Zahl in $\mathbb{R}$ eine endliche Grösse besitzt.

Satz 2.3.2. (Archimedisches Prinzip) *Zu jeder Zahl $0 < b \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $b < n$.*

Beweis (indirekt). Andernfalls gibt es $b \in \mathbb{R}$ mit

$$n \leq b, \quad \forall n \in \mathbb{N}.$$

Dann ist b eine obere Schranke für $\mathbb{N}$, und es existiert $c = \sup \mathbb{N} \in \mathbb{R}$.

Mit $n \in \mathbb{N}$ ist jedoch auch $n + 1 \in \mathbb{N}$, also

$$n + 1 \leq c, \quad \forall n \in \mathbb{N}.$$

Somit folgt

$$n \leq c - 1, \quad \forall n \in \mathbb{N}$$

im Widerspruch zur Minimalität von c. □

Vereinbarung: Für nach oben unbeschränkte Mengen $A \neq \emptyset$ setzen wir

$$\sup A = \infty,$$

analog für nach unten unbeschränkte Mengen $A \neq \emptyset$

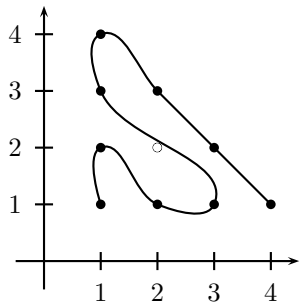
$$\inf A = -\infty.$$

Wegen Satz 2.3.2 definieren die Symbole $\pm\infty$ keine reellen Zahlen. Formal definieren wir noch

$$\infty + \infty = \infty,\ \infty + x = \infty,\ \forall x \in \mathbb{R};$$

jedoch ist der Ausdruck $\infty - \infty$ nicht sinnvoll erklärt.

Kardinalität: Gibt es mehr rationale oder mehr irrationale Zahlen? Wir sagen, zwei Mengen X und Y sind gleichmächtig, falls es eine bijektive Abbildung $f\colon X \to Y$ gibt. Die rationalen Zahlen kann man mit dem ersten Cantorschen Diagonalverfahren abzählen, wie in der Abbildung unten dargestellt; $\mathbb{Q}$ und $\mathbb{N}$ sind demnach gleichmächtig.



Kann man auch die reellen Zahlen abzählen? Dann könnte man auch alle Zahlen der Art $a = 0.a_1a_2a_3\ldots$ mit $a_i \in \{0,1\}$ abzählen. (Solche unendlichen Dezimalzahlen definieren nach Axiom V genau ein $a \in \mathbb{R}$.) Sei

$$\begin{aligned} a^{(1)} &= 0.a_1^{(1)}a_2^{(1)}a_3^{(1)}\ldots \\ a^{(2)} &= 0.a_1^{(2)}a_2^{(2)}a_3^{(2)}\ldots \\ a^{(3)} &= 0.a_1^{(3)}a_2^{(3)}a_3^{(3)}\ldots \\ &\vdots \end{aligned}$$

solch eine Abzählung. Setze

$$b = 0.b_1b_2\ldots$$

mit

$$b_i = a_i^{(i)} + 1 \mod 2.$$

Dann gilt offenbar $b_i \neq a_i^{(i)}$ und damit $b \neq a^{(i)}$ für jedes $i \in \mathbb{N}$; d.h. b kommt in der Abzählung nicht vor. (Dies ist das 2. Cantorsches Diagonalverfahren.) $\mathbb{R}$ ist somit "mächtiger" als die Menge der natürlichen Zahlen $\mathbb{N}$.

Die Kontinuumshypothese: Die Frage, ob jede Teilmenge von $\mathbb{R}$ entweder abzählbar ist oder gleichmächtig ist wie $\mathbb{R}$, hat die Mathematik lange beschäftigt. Gödel (1937) und Cohen (1964) konnten schliesslich zeigen, dass diese Frage nicht aus den Axiomen entscheidbar ist. (Vergleiche Davis-Hersch: *Erfahrung Mathematik*, S.336.)

2.4 Der euklidische Raum

Die **euklidische Ebene** $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x,y);\ x,y \in \mathbb{R}\}$ ist unsere Zeichenebene.

Beispiel 2.4.1. Den Graphen einer Funktion $f\colon \mathbb{R} \to \mathbb{R}$ können wir bequem in der euklidischen Ebene darstellen.

Der **3-dimensionale euklidische Raum** $\mathbb{R}^3 = \mathbb{R}\times\mathbb{R}\times\mathbb{R} = \{(x, y, z);\ x, y, , z \in \mathbb{R}\}$ ist unser Anschauungsraum.

Beispiel 2.4.2. 1. Die Bewegung eines Massepunktes kann man durch dessen Orts- und Geschwindigkeitsvektor beschreiben.

2. Eine Schar von N Massepunkten (Atome in einem Gas, Planeten im Sonnensystem) können wir gleichzeitig mit ihrem jeweiligen Ort $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)})$, $1 \le i \le N$, erfassen, indem wir diese Koordinaten in einen langen Vektor $x = (x_1, \dots, x_{3N})$ eintragen. Wir können dann wie gewohnt komponentenweise damit rechnen.

Für beliebiges $n \in \mathbb{N}$ erhalten wir so den **n-dimensionalen euklidischen Raum**

$$\mathbb{R}^n = \{x = (x_1, \dots, x_n);\ x_k \in \mathbb{R},\ 1 \le k \le n\}$$

mit komponentenweiser **Addition**

$$x + y = (x_1 + y_1, \dots, x_n + y_n),\ \forall x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n) \in \mathbb{R}^n$$

und **Skalarmultiplikation**

$$\lambda x = (\lambda x_1, \dots, \lambda x_n),\ \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n,\ \lambda \in \mathbb{R}.$$

Offensichtlich "erbt" der Raum $\mathbb{R}^n$ bezüglich der Addition die Struktur einer abelschen Gruppe. Das neutrale Element ist

$$0 = (0, \dots, 0) :\ \text{der } \textbf{Nullvektor}.$$

Bezüglich der Skalarmultiplikation gelten die Regeln

S.i) Distributivgesetz: $(\alpha + \beta)x = \alpha x + \beta x$,

S.ii) Distributivgesetz: $\alpha(x + y) = \alpha x + \alpha y$,

S.iii) Assoziativität: $\alpha(\beta x) = (\alpha\beta)x$,

S.iv) Einselement: $1 \cdot x = x$

für alle $x, y \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$.

Dies ist die Struktur eines $\mathbb{R}$**-Vektorraums**; vgl. Lineare Algebra.

Bezüglich der **Standardbasis**

$$e_i = (0, \dots, 0, \underbrace{1}_{i-\text{te Stelle}}, 0, \dots, 0) \in \mathbb{R}^n,\ 1 \le i \le n,$$

lässt sich jeder Vektor $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ in eindeutiger Weise als **Linearkombination**

$$x = (x_1, \dots, x_n) = x_1 e_1 + \cdots + x_n e_n = \sum_{i=1}^{n} x_i e_i$$

darstellen.

Skalarprodukt: Für $x = (x_i)_{1\le i\le n},\ \ y = (y_i)_{1\le i\le n} \in \mathbb{R}^n$ setze

$$x \cdot y = x_1y_1 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i \in \mathbb{R}.$$

Das so definierte Skalarprodukt hat die Eigenschaften

SP.i) Symmetrie: $x \cdot y = y \cdot x$,

SP.ii) (Bi-)Linearität: $x \cdot (y + z) = x \cdot y + x \cdot z$,

SP.iii) (Bi-)Linearität: $x \cdot (\alpha y) = \alpha(x \cdot y)$

für alle $x, y \in \mathbb{R}^n,\ \ \alpha \in \mathbb{R}$.

Beispiel 2.4.3. i) Für $x = (2, 0, 3),\ y = (-3, 1, 2)$ gilt

$$x \cdot y = -2 \cdot 3 + 0 \cdot 1 + 3 \cdot 2 = 0;$$

d.h. x und y stehen **senkrecht** aufeinander.

ii) Dies gilt auch für verschiedene Standardbasisvektoren

$$e_i \cdot e_j = 0 \qquad (i \neq j).$$

Euklidische Norm: Mit Hilfe des Skalarprodukts können wir die Länge von Vektoren messen, indem wir setzen

$$\|x\| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2} \qquad \text{(positive Wurzel)}.$$

Beispiel 2.4.4. i) Es gilt $\|e_i\| = 1$, $1 \le i \le n$. Die Standardbasisvektoren sind also paarweise **orthogonal** und auf Länge 1 **normiert**; sie sind **orthonormal**.

ii) Nach Pythagoras ist der Abstand des Punktes (x_1, x_2) vom Nullpunkt

$$l = \sqrt{x_1^2 + x_2^2} = \|(x_1, x_2)\|\,.$$

iii) Insbesondere hat die Diagonale im Einheitsquadrat die Länge $l_2 = \sqrt{2}$, im Einheitswürfel im $\mathbb{R}^3$ die Länge $l_3 = \sqrt{3}$, im Einheitshyperwürfel im $\mathbb{R}^n$ die Länge $l_n = \sqrt{n}$.

Satz 2.4.1. (Cauchy-Schwarz) *Für alle $x, y \in \mathbb{R}^n$ gilt*

$$|x \cdot y| \le \|x\| \cdot \|y\|\,.$$

Beweis. OBdA $x \neq 0 \neq y$. Mit Satz 2.2.2 (Young) können wir bei Wahl von $\epsilon = \frac{\|y\|}{\|x\|} > 0$ abschätzen

$$\begin{aligned} 2\ |x \cdot y| &= 2\ |x_1 y_1 + \dots x_n y_n| \leq 2\ |x_1 y_1| + \dots + 2\ |x_n y_n| \\ &\leq \epsilon x_1^2 + \frac{1}{\epsilon} y_1^2 + \dots + \epsilon x_n^2 + \frac{1}{\epsilon} y_n^2 = \epsilon \|x\|^2 + \frac{1}{\epsilon} \|y\|^2 = 2\ \|x\| \|y\| . \end{aligned}$$

□

Wir können Satz 2.4.1 auch geometrisch herleiten: OBdA sei $x \neq 0$. Zerlege y orthogonal

$$y = \underbrace{\frac{x}{\|x\|}\big(\frac{x}{\|x\|} \cdot y\big)}_{=y^{||}} + \underbrace{\big(y - \frac{x}{\|x\|}\big(\frac{x}{\|x\|} \cdot y\big)\big)}_{=y^{\perp}} .$$

Offenbar gilt $x \cdot y^{\perp} = 0$, also auch $y^{||} \cdot y^{\perp} = 0$. Mit Pythagoras folgt nun die gewünschte Ungleichung $\frac{|x \cdot y|}{\|x\|} = \big\|y^{||}\big\| \leq \|y\|$ sofort.

Satz 2.4.2. *Die euklidische Norm hat die Eigenschaften*

i) Definitheit: $\forall x \in \mathbb{R}^n : \quad \|x\| \geq 0, \quad \|x\| = 0 \Rightarrow x = 0,$

ii) Positive Homogenität: $\forall x \in \mathbb{R}^n,\ \alpha \in \mathbb{R} :\ \|\alpha x\| = |\alpha| \cdot \|x\|,$

iii) Dreiecks-Ungleichung: $\forall x, y \in \mathbb{R}^n :\ \|x + y\| \leq \|x\| + \|y\|.$

Beweis. i) und ii) folgen direkt aus der Definition.
iii) Wie im Beweis von Satz 2.2.1 schätzen wir mit Satz 2.4.1 ab

$$\begin{aligned} \|x + y\|^2 &= (x + y) \cdot (x + y) = x \cdot x + 2\ x \cdot y + y \cdot y \\ &\leq \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Beispiel 2.4.5. Für $x = (1, 1) = e_1 + e_2 \in \mathbb{R}^2$ gilt

$$\sqrt{2} = \|x\| \leq \|e_1\| + \|e_2\| = 2.$$

2.5 Komplexe Zahlen

In $\mathbb{R}^2$ können wir zusätzlich zur Addition eine weitere Verknüpfung einführen, die **komplexe Multiplikation**

$$\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \ni (a, b), (c, d) \mapsto (ac - bd, ad + bc) \in \mathbb{R}^2.$$

Diese Operation ist assoziativ mit neutralem Element $(1, 0)$. Weiter gilt für $(a, b) \neq (0, 0)$ die Gleichung

$$(a, b) \cdot \Big(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\Big) = (1, 0); \tag{2.5.1}$$

d.h.

$$(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}) \in \mathbb{R}^2$$

ist zu (a, b) invers.

Schliesslich ist die komplexe Multiplikation kommutativ, und es gilt das Distributivgesetz

$$((a_1, b_1) + (a_2, b_2)) \cdot (c, d) = (a_1, b_1) \cdot (c, d) + (a_2, b_2) \cdot (c, d).$$

D.h. $\mathbb{R}^2$ bildet bzgl. Addition und komplexer Multiplikation einen Zahlkörper, den **Körper der komplexen Zahlen** $\mathbb{C}$.

Bemerkung 2.5.1. i) Wir können $\mathbb{R}$ in $\mathbb{C}$ „einbetten“ mittels

$$\mathbb{R} \ni x \mapsto (x, 0) \in \mathbb{C}.$$

Diese Einbettung ist verträglich mit den Körperoperationen, da gilt

$$\begin{aligned} x + y &\mapsto (x + y, 0) = (x, 0) + (y, 0), \\ xy &\mapsto (xy, 0) = (x, 0) \cdot (y, 0). \end{aligned}$$

Zudem ist sie verträglich mit der Skalarmulitplikation in $\mathbb{R}^2$, denn

$$\alpha(x, y) = (\alpha x, \alpha y) = (\alpha, 0) \cdot (x, y).$$

ii) Somit können wir den Standardbasisvektor $e_1 = (1, 0) \in \mathbb{R}^2$ „identifizieren“ mit $1 \in \mathbb{R}$. Für $e_2 = (0, 1) \in \mathbb{R}^2$ führen wir das Symbol i ein,

$$i = (0, 1): \quad \text{„imaginäre Einheit“},$$

mit

$$i^2 = (-1, 0) = -1.$$

Somit hat jedes $z = (x, y) \in \mathbb{C}$ die eindeutige Darstellung

$$z = xe_1 + ye_2 = x + iy$$

mit **Realteil** $x = Re(z)$ und **Imaginärteil** $y = Im(z)$.

Konjugation. Zu $z = x + iy \in \mathbb{C}$ sei

$$\overline{z} = x - iy \in \mathbb{C}$$

die zu z konjugierte Zahl. Die Konjugation hat die Eigenschaften:

i) Für alle $z = x + iy = (x, y) \in \mathbb{C} = \mathbb{R}^2$ gilt

$$\begin{aligned} z \cdot \overline{z} &= (x + iy) \cdot (x - iy) = x^2 - i^2 y^2 = x^2 + y^2 \\ &= x^2 + y^2 = \|z\|^2 . \end{aligned} \tag{2.5.2}$$

ii) Für alle $z_{1,2} \in \mathbb{C}$ gilt

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \qquad \overline{z_1 z_2} = \overline{z_1} \cdot \overline{z_2}. \tag{2.5.3}$$

Beweis.

$$(x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

□

Folgerung 2.5.1. i) Mit (2.5.2) folgt

$$z^{-1} = \frac{\overline{z}}{\|z\|^2}, \qquad \forall z \in \mathbb{C}\backslash\{0\};$$

in Übereinstimmung mit (2.5.1).

Beispiel:

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

ii) Mit (2.5.2) erhalten wir

$$\|zw\|^2 = (zw) \cdot \overline{(zw)} = zw\overline{zw} = \|z\|^2 \|w\|^2;$$

d.h. wie in $\mathbb{R}$ gilt

$$\|zw\| = \|z\| \|w\|, \quad \forall z, w \in \mathbb{C}.$$

Zur Abkürzung schrieben wir im folgenden daher $|z| = \|z\|$ für den **Absolutbetrag** der Zahl $z \in \mathbb{C}$.

Polarform: Führen wir (r, ϕ) ein als Polarkoordinaten in der Ebene, so gilt für $z = x + iy \in \mathbb{C}$ offenbar

$$r = |z|, \quad x = r\cos\phi, \quad y = r\sin\phi,$$

d.h.

$$z = r \underbrace{(\cos\phi + i\sin\phi)}_{=:e^{i\phi}\ \text{(Euler)}} = re^{i\phi}.$$

Die Additionstheoreme

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

für cos und sin ergeben die Beziehung

$$\begin{aligned}e^{i\phi}e^{i\psi} &= (\cos\phi + i\sin\phi)(\cos\psi + i\sin\psi) \\ &= (\cos\phi\cos\psi - \sin\phi\sin\psi) + i(\sin\phi\cos\psi + \cos\phi\sin\psi) \\ &= \cos(\phi+\psi) + i\sin(\phi+\psi) = e^{i(\phi+\psi)}.\end{aligned}$$

Somit folgt für

$$z = re^{i\phi}, \quad w = se^{i\psi} \in \mathbb{C}$$

die einfache Darstellung

$$zw = rse^{i(\phi+\psi)}.$$

Beispiel 2.5.1. i) $(1+i) = \sqrt{2}e^{i\pi/4}$, also $(1+i)^2 = 2e^{i\pi/2} = 2i$.

Zur Probe können wir dies Ergebnis auch direkt berechnen: $(1+i)\cdot(1+i) = 0 + i\cdot 2 = 2i$.

ii) Welchen Wert hat die Zahl

$$z = \frac{(1-i)^4}{(\sqrt{3}+i)^3}\,?$$

Setze

$$z_1 = 1 - i = \sqrt{2}e^{-i\pi/4},$$
$$z_2 = \sqrt{3} + i = 2e^{i\phi_2},$$

wobei

$$\phi_2 = \arctan\frac{1}{\sqrt{3}} = \frac{\pi}{6}.$$

Es folgt

$$z_1^4 = \sqrt{2}^4 e^{-i\pi} = -4,$$
$$z_2^3 = 2^3 e^{i3\phi_3} = 8e^{i\pi/2} = 8i;$$

d.h.

$$z = \frac{z_1^4}{z_2^3} = \frac{-4}{8i} = \frac{4i}{8} = \frac{i}{2}.$$

In $\mathbb{C}$ kann man die Gleichung $z^2 = c$ für jede Zahl $c = se^{i\psi}$ lösen; d.h. man kann aus jeder Zahl c Quadratwurzeln ziehen. Der Ansatz $z = re^{i\phi}$ führt auf

$$z^2 = r^2 e^{i2\phi} = se^{i\psi};$$

d.h.

$$r = \sqrt{s}, \quad \phi = \psi/2 \mod \pi$$

oder

$$z = \pm\sqrt{s}e^{i\psi/2}.$$

Allgemein gilt für jede Zahl

$$c = se^{i\psi} \in \mathbb{C},\ q \in \mathbb{N},$$

dass

$$z = \sqrt[q]{s}e^{i\phi}, \quad \text{wobei} \quad \phi = \frac{\psi}{q} \mod \frac{2\pi}{q}$$

die q verschiedenen Lösungen der Gleichung

$$z^q = c$$

beschreibt.

Beispiel 2.5.2. Für $c = 1$, $q \in \mathbb{N}$ erhält man so die q-ten **Einheitswurzeln** $z = e^{i\frac{2\pi k}{q}}$, $k = 0, \ldots, q-1$.

Wir erkennen hier bereits, dass es nicht ohne weiteres sinnvoll ist, in $\mathbb{C}$ irrationale oder imaginäre Potenzen zu bilden, da das **Argument** ϕ einer Zahl $z = re^{i\phi}$ nur modulo 2π bestimmt ist. So wäre z. B. mit

$$1 = e^{2\pi k i},\ k \in \mathbb{Z},$$

die Zahl 1^i als die Menge

$$1^i = \{e^{-2\pi k};\quad k \in \mathbb{Z}\}$$

zu deuten, was wenig sinnvoll scheint.

Auch gibt es keine mit den Körperoperationen verträgliche **Ordnung** auf $\mathbb{C}$; sonst wäre gemäss Folgerung 2.2.1 iii)

$$i^2 \geq 0$$

und mit $1 = 1^2 > 0$ folgt

$$0 = 1 + i^2 > 0.$$

Hingegen ist $\mathbb{C}$ im Unterschied zu $\mathbb{R}$ **algebraisch vollständig:** Nicht nur die Gleichung $z^2 + 1 = 0$ hat in $\mathbb{C}$ die Lösungen $z = \pm i$, sondern es gilt der

Fundamentalsatz der Algebra: Jedes Polynom

$$p(z) = z^n + a_{n-1}z^{n-1} + \cdots + a_0$$

vom Grad $n \geq 1$ hat in $\mathbb{C}$ eine Nullstelle.

Den Beweis müssen wir jedoch auf später verschieben.

Kapitel 3

Folgen und Reihen

3.1 Beispiele

Die folgenden Beispiele sind aus der Mittelschule bekannt:

i) Die **Fibonacci Zahlen**

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13, \dots$$

entstehen aus einem einfachen Populationsmodell gemäss dem Gesetz

$$a_0 = 1,\ a_1 = 1,\ a_{n+1} = a_n + a_{n-1}, \quad \forall n \in \mathbb{N}.$$

ii) Die **Zinsfaktoren** bei $\frac{1}{n}$-tel jährlicher Verzinsung,

$$a_n = (1 + \frac{1}{n})^n, \quad n \in \mathbb{N},$$

streben für $n \to \infty$ gegen die Eulersche Zahl

$$e = 2.718 \dots,$$

den "Limes der kontinuierlichen Verzinsung".

iii) Die **geometrische Reihe**

$$S_n = 1 + q + q^2 + \cdots + q^n = \sum_{k=0}^{n} q^k, \quad n \in \mathbb{N},$$

hat für $-1 < q < 1$ den "Grenzwert" $S = \frac{1}{1-q}$; vergleiche Beispiel 3.7.1.i).

3.2 Grenzwert einer Folge

Sei $(a_n)_{n\in\mathbb{N}} = (a_1, a_2, a_3, \dots)$ eine Folge in $\mathbb{R},\ a \in \mathbb{R}$.

Definition 3.2.1. i) *Die Folge* $(a_n)_{n\in\mathbb{N}}$ **konvergiert** *gegen* a *für* $n \to \infty$*, falls gilt*

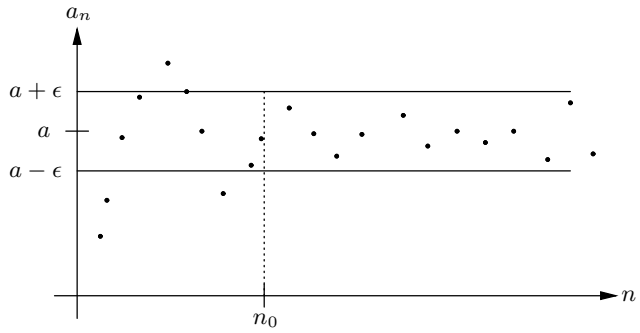
$$\forall \epsilon > 0 \ \exists n_0 = n_0(\epsilon) \in \mathbb{N} \ \forall n \geq n_0 : \ |a_n - a| < \epsilon;$$

*d.h. falls zu jeder (noch so kleinen) „***Fehlerschranke***“* $\epsilon > 0$ *ab einem genügend grossen Index* $n_0 = n_0(\epsilon)$ *alle Folgenglieder sich um weniger als* ϵ *von* a *unterscheiden. Wir schreiben dann*

$$a = \lim_{n\to\infty} a_n \quad \textit{oder} \quad a_n \to a \ (n \to \infty)$$

und nennen a *den* **Grenzwert** *oder* **Limes** *der Folge* $(a_n)_{n\in\mathbb{N}}$*.*

ii) *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *heisst* **konvergent***, falls sie einen Limes besitzt; andernfalls heisst die Folge* **divergent***.*



Beispiel 3.2.1. i) Für $a_n = \frac{1}{n}$, $n \in \mathbb{N}$, gilt $a_n \to 0 \ \ (n \to \infty)$.

Beweis. Nach Satz 2.3.2 gibt es zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $n_0 > \frac{1}{\epsilon}$, d.h. $\frac{1}{n_0} < \epsilon$. Es folgt

$$-\epsilon < 0 < \frac{1}{n} \leq \frac{1}{n_0} < \epsilon, \quad \forall n \geq n_0,$$

wie gewünscht. □

ii) Sei $q \in \mathbb{R}$ mit $0 < q < 1$. Dann gilt

$$q^n \to 0 \ \ (n \to \infty).$$

Beweis. Schreibe $\frac{1}{q} = 1 + \delta$ mit $\delta > 0$. Da gemäss der Bernoullischen Ungleichung $(1 + x)^n \geq 1 + nx$ für alle $x \geq -1$, $n \in \mathbb{N}$ (vgl. Übung 1.1), folgt

$$\frac{1}{q^n} = \left(\frac{1}{q}\right)^n = (1 + \delta)^n \geq 1 + n\delta \geq n\delta, \ \ \forall n \in \mathbb{N};$$

also

$$0 < q^n \leq \frac{1}{n\delta}, \quad \forall n \in \mathbb{N}.$$

Zu $\epsilon > 0$ wähle $n_0 = n_0(\epsilon)$ mit $\frac{1}{\epsilon\delta} < n_0$. Es folgt

$$0 < q^n \leq \frac{1}{n\delta} \leq \frac{1}{n_0\delta} < \epsilon, \quad \forall n \geq n_0.$$

□

iii) Es gilt $\sqrt[n]{n} \to 1 \ \ (n \to \infty)$.

Beweis. Für $0 < a, b \in \mathbb{R}$ gilt

$$a^n - b^n = (a-b)(\underbrace{a^{n-1} + ba^{n-2} + \cdots + b^{n-2}a + b^{n-1}}_{>0}); \tag{3.2.1}$$

somit folgt

$$a \geq b \Leftrightarrow a^n \geq b^n, \quad \forall a, b > 0, \ n \in \mathbb{N}. \tag{3.2.2}$$

Sei nun $\epsilon > 0$ beliebig vorgegeben. Schätze ab

$$(1+\epsilon)^n = \underbrace{1 + n\epsilon}_{>0} + \binom{n}{2}\epsilon^2 + \underbrace{\binom{n}{3}\epsilon^3 + \cdots + \epsilon^n}_{\geq 0} > \binom{n}{2}\epsilon^2 = \frac{n(n-1)}{2}\epsilon^2 \geq n,$$

falls n so gross gewählt, dass

$$\frac{n-1}{2}\epsilon^2 \geq 1.$$

Setze

$$n_0 \geq \frac{2}{\epsilon^2} + 1.$$

Dann gilt für $n \geq n_0$ stets

$$\frac{n-1}{2}\epsilon^2 \geq \frac{n_0 - 1}{2}\epsilon^2 \geq 1,$$

also auch

$$(1+\epsilon)^n \geq n \geq 1,$$

und mit (3.2.2) folgt

$$1 \leq \sqrt[n]{n} < 1 + \epsilon, \quad \forall n \geq n_0;$$

d.h.

$$\left| \sqrt[n]{n} - 1 \right| < \epsilon, \quad \forall n \geq n_0.$$

□

Nicht jede Folge $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ ist konvergent.

Beispiel 3.2.2. i) Sei $a_n = (-1)^n$, $n \in \mathbb{N}$. Offenbar gilt für jedes $a \in \mathbb{R}$, $n \in \mathbb{N}$:

$$|a_n - a| + |a_{n+1} - a| \geq |(a_n - a) - (a_{n+1} - a)| = 2,$$

und kein $a \in \mathbb{R}$ kann Grenzwert von (a_n) sein.

ii) Sei $a_n = n$, $n \in \mathbb{N}$. Zu jedem $a \in \mathbb{R}$ gibt es n_0 mit $a < n_0$; also

$$|a_n - a| = n - a \geq n_0 - a > 0, \quad \forall n \geq n_0,$$

und kein $a \in \mathbb{R}$ kann Grenzwert von (a_n) sein.

iii) Ebenso gilt für die Fibonacci Zahlen $(F_n)_{n\in\mathbb{N}_0}$ induktiv $F_n \geq n$ für alle $n \in \mathbb{N}$, und kein $a \in \mathbb{R}$ kann Grenzwert von (F_n) sein.

Beweis (Induktion). Es gilt $F_0 = 1,\ F_1 = 1$ und daher auch $F_n \geq 1$ für alle $n \in \mathbb{N}_0$. Falls wir annehmen $F_n \geq n$ für ein $n \geq 1$, so folgt auch $F_{n+1} = F_n + F_{n-1} \geq n + 1$. □

iv) Seien $p \in \mathbb{N}$, $q \in \mathbb{R}$ mit $0 < q < 1$ fest. Dann gilt

$$\lim_{n\to\infty} n^p q^n = 0;$$

d.h. die Exponentialfunktion wächst schneller als jede Potenz.

Beweis. Setze

$$s = q^{1/p} = \sqrt[p]{q} < 1, \quad s > 0,$$

so dass

$$a_n = n^p q^n = (ns^n)^p = (s\sqrt[n]{n})^{np}, \quad n \in \mathbb{N}.$$

Wähle $\epsilon > 0$ mit

$$s = \frac{1}{(1+\epsilon)^2},$$

dazu $n_0 = n_0(\epsilon) \in \mathbb{N}$ gemäss Beispiel 3.2.1.iii), so dass

$$\sqrt[n]{n} < 1 + \epsilon, \quad \forall n \geq n_0.$$

Damit erhalten wir

$$0 < a_n = \left(\frac{\sqrt[n]{n}}{(1+\epsilon)^2}\right)^{np} < \left(\frac{1}{1+\epsilon}\right)^{pn} = r^n$$

mit

$$r = \left(\frac{1}{1+\epsilon}\right)^p < 1.$$

Mit Beispiel 3.2.1.ii) folgt $0 < a_n < r^n \to 0 \ \ (n \to \infty,\ n \geq n_0)$. □

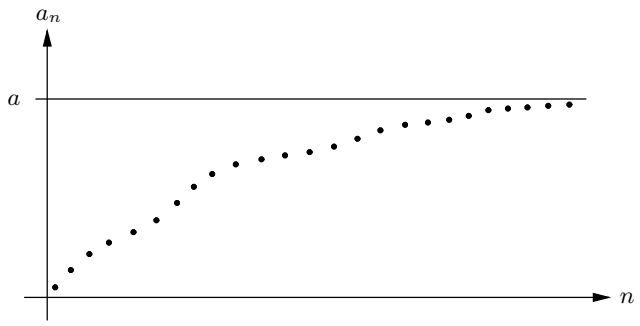
3.3 Konvergenzkriterien

Kann man es einer Folge $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ ansehen, ob sie konvergiert, ohne den Limes zu kennen? Es gibt einige nützliche Kriterien, die dies erleichtern.

Satz 3.3.1. (Monotone Konvergenz) *Sei $(a_n) \subset \mathbb{R}$ nach oben beschränkt und monoton wachsend; d.h. mit einer Zahl $b \in \mathbb{R}$ gelte:*

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq b, \quad \forall n \in \mathbb{N}.$$

Dann ist (a_n) konvergent, und $\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n$.



Beweis. Setze $A = \{a_n;\ n \in \mathbb{N}\}$. Nach Annahme ist $A \neq \emptyset$ nach oben beschränkt; also existiert

$$a = \sup A = \sup_{n\in\mathbb{N}} a_n$$

gemäss Satz 2.3.1.

Behauptung Es gilt $a = \lim_{n\to\infty} a_n$.

Beweis. Sei $\epsilon > 0$ beliebig vorgegeben. Da $a \in \mathbb{R}$ die **kleinste** obere Schranke für A ist, gibt es $n_0 = n_0(\epsilon) \in \mathbb{N}$ mit $a_{n_0} > a - \epsilon$. Monotonie ergibt

$$a - \epsilon < a_{n_0} \leq a_n \leq \sup_{l\in\mathbb{N}} a_l = a < a + \epsilon, \quad \forall n \geq n_0,$$

wie gewünscht. □

Beispiel 3.3.1. Jeder unendliche Dezimalbruch

$$x = x_0.x_1 \ldots x_k \ldots$$

definiert eine monoton gegen die Zahl $x \in \mathbb{R}$ konvergente Folge.

Satz 3.3.2. *Seien die Folgen $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ konvergent mit $\lim_{n\to\infty} a_n = a$, $\lim_{n\to\infty} b_n = b$. Dann konvergieren die Folgen $(a_n + b_n)_{n\in\mathbb{N}}$, $(a_n \cdot b_n)_{n\in\mathbb{N}}$, und*

i) $\lim_{n\to\infty} (a_n + b_n) = a + b = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$,

ii) $\lim_{n\to\infty} (a_n \cdot b_n) = a \cdot b = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$.

iii) *Falls zusätzlich $b \neq 0 \neq b_n$ für alle n, so gilt auch $\lim_{n\to\infty} (a_n/b_n) = a/b$.*

iv) *Falls $a_n \leq b_n$ für $n \in \mathbb{N}$, so auch $a \leq b$.*

Beweis. Zu $\epsilon > 0$ sei im folgenden $n_0 = n_0(\epsilon) \in \mathbb{N}$ stets so gewählt, dass

$$|a_n - a| < \epsilon, \quad |b_n - b| < \epsilon, \quad \forall n \geq n_0.$$

i), ii) OBdA sei $\epsilon < 1$. Es folgt

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < 2\epsilon, \ \forall n \geq n_0,$$

und wegen $|b_n| \leq |b_n - b| + |b| \leq |b| + 1$ analog auch

$$\begin{aligned} |a_n b_n - ab| &= |(a_n - a)b_n + a(b_n - b)| \\ &\leq |b_n| \cdot |a_n - a| + |a| \cdot |b_n - b| \leq \big(|a| + |b| + 1\big)\epsilon, \quad \forall n \geq n_0. \end{aligned}$$

Da $\epsilon > 0$ beliebig, folgt die Behauptung.

iii) Wegen ii) genügt es, den Fall $a = a_n = 1$ für alle $n \in \mathbb{N}$ zu betrachten. OBdA gelte auch $0 < \epsilon < \frac{|b|}{2}$, also

$$|b_n| = |b_n - b + b| \geq |b| - |b_n - b| \geq |b| - \epsilon > |b|/2, \quad \forall n \geq n_0.$$

Es folgt

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \left|\frac{b_n - b}{b_n b}\right| \leq \frac{2}{|b|^2}\,|b_n - b| \leq \frac{2}{|b|^2} \cdot \epsilon, \quad \forall n \geq n_0.$$

Da $\epsilon > 0$ beliebig, folgt die Behauptung.

iv) (indirekt). Falls wir widerspruchsweise annehmen, dass $a > b$, folgt bei Wahl von $a - b =: 2\epsilon > 0$ die Ungleichung

$$b_n < b + \epsilon = a - \epsilon < a_n, \ \forall n \geq n_0,$$

im Widerspruch zur Annahme, dass $a_n \leq b_n$ für alle $n \in \mathbb{N}$. □

Bemerkung 3.3.1. Wie die Beispiele 3.2.1.i)-iii) zeigen, folgt aus $a_n < b_n$, $n \in \mathbb{N}$, im Allgemeinen nicht die strikte Ungleichung $a < b$.

Beispiel 3.3.2. i) Durch Kombination der Aussagen Satz 3.3.2.i)-iii) sieht man sofort ein, dass

$$a_n = \frac{3n^4 - 7n^3 + 5}{2n^4 + 6n^2 + 3} = \frac{3 - 7/n + 5/n^4}{2 + 6/n^2 + 3/n^4} \to \frac{3}{2} \ \ (n \to \infty).$$

ii) Eulersche Zahl: Betrachte die Folgen

$$a_n = (1 + \frac{1}{n})^n < b_n = (1 + \frac{1}{n})^{n+1}, \qquad n \in \mathbb{N}.$$

Behauptung Es gilt

$$2 = a_1 \leq a_2 \leq \cdots \leq a_{n-1} \leq a_n < b_n \leq b_{n-1} \leq \cdots \leq b_1 = 4, \quad \forall n \in \mathbb{N}.$$

Beweis. Wir schätzen ab

$$\begin{aligned} \frac{a_n}{a_{n-1}} &= \left(\frac{1 + \frac{1}{n}}{1 + \frac{1}{n-1}}\right)^n \big(1 + \frac{1}{n-1}\big) = \left(\frac{\big(\frac{n+1}{n}\big)}{\big(\frac{n}{n-1}\big)}\right)^n \cdot \frac{n}{n-1} \\ &= \Big(\frac{n^2 - 1}{n^2}\Big)^n \cdot \frac{n}{n-1} = \big(1 - \frac{1}{n^2}\big)^n \cdot \frac{1}{1 - \frac{1}{n}} \geq \big(1 - \frac{1}{n}\big)\frac{1}{1 - \frac{1}{n}} = 1, \end{aligned}$$

wobei wir im letzten Schritt die Bernoullische Ungleichung verwenden. Analog erhalten wir unter Verwendung von $\frac{n}{n^2-1} \geq \frac{n}{n^2} = \frac{1}{n}$ die Abschätzung

$$\begin{aligned}\frac{b_{n-1}}{b_n} &= \left(\frac{1+\frac{1}{n-1}}{1+\frac{1}{n}}\right)^n \cdot \frac{1}{1+\frac{1}{n}} = \Big(\frac{n^2}{n^2-1}\Big)^n \cdot \frac{1}{1+\frac{1}{n}} \\ &= \Big(1+\frac{1}{n^2-1}\Big)^n \cdot \frac{1}{1+\frac{1}{n}} \geq \big(1+\frac{n}{n^2-1}\big) \cdot \frac{1}{1+\frac{1}{n}} \geq 1.\end{aligned}$$

Die Behauptung folgt. □

Gemäss Satz 3.3.1 und Satz 3.3.2 existieren

$$a = \lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n \leq b = \lim_{n\to\infty} b_n = \inf_{n\in\mathbb{N}} b_n.$$

Weiter gilt für beliebiges $n \in \mathbb{N}$

$$0 \leq b - a \leq b_n - a_n = \Big(\big(1+\frac{1}{n}\big) - 1\Big) \cdot a_n \leq \frac{4}{n};$$

also $a = b =: e$, die Eulersche Zahl.

iii) Sei $c > 1$. Setze $a_1 = c$ und

$$a_{n+1} = \frac{1}{2}\big(a_n + \frac{c}{a_n}\big) = a_n + \frac{c-a_n^2}{2a_n}, \quad n \in \mathbb{N}.$$

Behauptung $\lim_{n\to\infty} a_n = a$ existiert, und $a^2 = c$.

Beweis. Es gilt $a_1^2 = c^2 > c$, und weiter

$$a_{n+1}^2 = \Big(a_n + \frac{c-a_n^2}{2a_n}\Big)^2 = a_n^2 + (c-a_n^2) + \underbrace{\Big(\frac{c-a_n^2}{2a_n}\Big)^2}_{\geq 0} \geq c, \quad \forall n \in \mathbb{N};$$

insbesondere folgt somit auch $a_{n+1} \leq a_n$, $\forall n \in \mathbb{N}$.

Behauptung $a_n \geq 1, \quad \forall n \in \mathbb{N}$.

Beweis (Induktion). $n = 1$: Nach Voraussetzung gilt $a_1 = c \geq 1$.

$n \mapsto n+1$: Mit der Induktionsannahme $a_n \geq 1 > 0$ erhalten wir zunächst

$$a_{n+1} = \frac{1}{2}\big(a_n + \frac{c}{a_n}\big) \geq \frac{a_n}{2} \geq \frac{1}{2} > 0.$$

Da andererseits $a_{n+1}^2 \geq c \geq 1$, folgt $a_{n+1} \geq 1$. □

Somit erhalten wir

$$1 \leq a_{n+1} \leq a_n \leq \cdots \leq a_1 = c, \quad \forall n \in \mathbb{N},$$

und gemäss Satz 3.3.1 existiert $a = \lim_{n\to\infty} a_n$. Mit Satz 3.3.2 folgt schliesslich

$$a = \lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{1}{2}\big(a_n + \frac{c}{a_n}\big) = \frac{1}{2}\big(a + \frac{c}{a}\big),$$

also erfüllt a die Gleichung $a^2 = c$, wie gewünscht. □

3.4 Teilfolgen, Häufungspunkte

Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.

Definition 3.4.1. *Sei $\Lambda \subset \mathbb{N}$ eine unendliche Teilmenge, $\mathbb{N} \ni n \mapsto l(n) \in \Lambda$ eine monotone Abzählung von Λ. Dann heisst die Folge $(a_l)_{l\in\Lambda} = (a_{l(n)})_{n\in\mathbb{N}}$ eine* **Teilfolge** *von $(a_n)_{n\in\mathbb{N}}$.*

Beispiel 3.4.1. i) Die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = (-1)^{n+1}$, $n \in \mathbb{N}$, hat die konstanten Teilfolgen $(a_{2n})_{n\in\mathbb{N}}$, $(a_{2n+1})_{n\in\mathbb{N}}$.

ii) Die Folge $(2^n)_{n\in\mathbb{N}}$ ist eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ mit $a_n = n$, $n \in \mathbb{N}$.

Definition 3.4.2. *$a \in \mathbb{R}$ heisst* **Häufungspunkt** *von $(a_n)_{n\in\mathbb{N}}$, falls $(a_n)_{n\in\mathbb{N}}$ eine gegen a konvergente Teilfolge besitzt; d.h. falls*

$$a = \lim_{l\to\infty,\ l\in\Lambda} a_l,$$

Bemerkung 3.4.1. Offenbar ist $a \in \mathbb{R}$ genau dann Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, falls gilt

$$\forall \epsilon > 0 \ \ \forall n_0 \in \mathbb{N} \ \ \exists l \geq n_0 : \ |a - a_l| < \epsilon.$$

Beweis. Falls $a_l \to a$ $(l \to \infty,\ l \in \Lambda)$, so existiert offenbar zu jedem $\epsilon > 0$ ein $l \geq n_0$ mit $|a - a_l| < \epsilon$. Umgekehrt wähle $l(1) = 1$ und für jedes $n > 1$ einen Index $l = l(n) > l(n-1)$ mit $|a - a_l| < 1/n$. Induktion liefert $l(n) \geq n$. □

Beispiel 3.4.2. Die Folge $a_n = (-1)^{n+1}$ hat die Häufungspunkte $+1$ und -1.

Limes superior, limes inferior. Sei $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ beschränkt, d.h.

$$\exists M \in \mathbb{R} \ \ \forall n \in \mathbb{N} : \ |a_n| < M.$$

Für $k \in \mathbb{N}$ existieren dann

$$c_k = \inf_{n\geq k} a_n \leq \sup_{n\geq k} a_n = b_k.$$

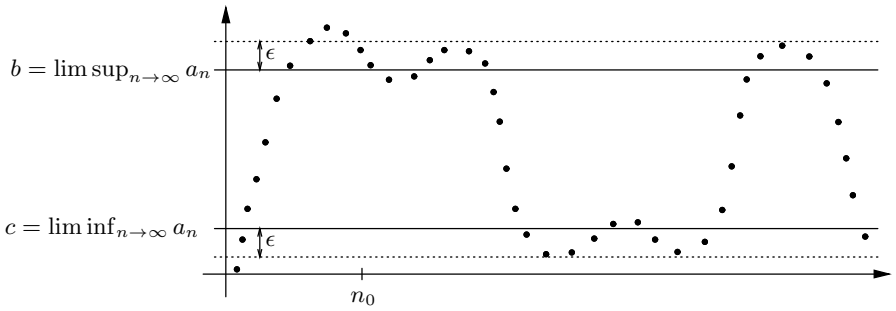
Offenbar gilt

$$-M \leq c_1 \leq \cdots \leq c_k \leq c_{k+1} \leq b_{k+1} \leq b_k \leq \cdots \leq b_1 \leq M, \ \ \forall k \in \mathbb{N}.$$

Mit Satz 3.3.1 folgt die Existenz von

$$b = \lim_{k\to\infty} b_k =: \limsup_{n\to\infty} a_n \qquad (\text{„Limes superior“}),$$
$$c = \lim_{k\to\infty} c_k =: \liminf_{n\to\infty} a_n \qquad (\text{„Limes inferior“}),$$

und $c \leq b$ wegen Satz 3.3.2.



Lemma 3.4.1. *b und c sind Häufungspunkte von $(a_n)_{n\in\mathbb{N}}$.*

Beweis. Seien $\epsilon > 0$ und $n_0 \in \mathbb{N}$ vorgegeben. Da $b = \lim\limits_{k\to\infty} b_k$, gibt es $k_0 = k_0(\epsilon)$ mit

$$|b_k - b| < \epsilon, \quad \forall k \geq k_0;$$

d.h.

$$b - \epsilon < b_k = \sup_{n\geq k} a_n < b + \epsilon, \quad \forall k \geq k_0.$$

OBdA $k_0 \geq n_0$. (Ersetze sonst k_0 durch n_0.) Fixiere $k = k_0$. Nach Definition des Supremums gibt es $l \geq k_0$ mit

$$a_l \geq \sup_{n\geq k_0} a_n - \epsilon = b_{k_0} - \epsilon > b - 2\epsilon;$$

Weiter gilt

$$a_l \leq \sup_{n\geq k_0} a_n < b + \epsilon < b + 2\epsilon;$$

also

$$|a_l - b| < 2\epsilon.$$

Die Behauptung folgt mit Bemerkung Bem:3.4.1.

Analog ist c Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$. □

Es folgt

Satz 3.4.1. (Bolzano Weierstrass) *Jede beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ besitzt eine konvergente Teilfolge, also auch einen Häufungspunkt.*

Bemerkung 3.4.2. i) Sei $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ beschränkt, und seien (b_k), b, (c_k) und c wie in Lemma 3.4.1. Zu $\epsilon > 0$ wähle $k_0 = k_0(\epsilon)$ mit

$$b_k = \sup_{n\geq k} a_n < b + \epsilon, \quad \forall k \geq k_0,$$
$$c_k = \inf_{n\geq k} a_n > c - \epsilon, \quad \forall k \geq k_0.$$

Für $k = k_0$ folgt dann

$$c - \epsilon < a_n < b + \epsilon, \quad \forall n \geq k_0;$$

d.h. für jedes $\epsilon > 0$ liegen alle bis auf endlich viele Glieder der Folge $(a_n)_{n\in\mathbb{N}}$ in $]c - \epsilon, b + \epsilon[$.

ii) Insbesondere ist b der grösste und ist c der kleinste Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$, und

iii) falls $b = c$, so ist $(a_n)_{n\in\mathbb{N}}$ konvergent mit

$$\lim_{n\to\infty} a_n = b = c.$$

iv) Umgekehrt konvergiert jede Teilfolge einer Folge $a_n \to a \ \ (n \to \infty)$ ebenfalls gegen a.

Beispiel 3.4.3. Sei $g_1 = 1$, $g_{n+1} = 1 + \frac{1}{g_n}$, $n \in \mathbb{N}$. Offenbar gilt $1 \le g_n \le 2$, $\forall n \in \mathbb{N}$; aber

$$(g_n)_{n\in\mathbb{N}} = (1, 2, \frac{3}{2}, \frac{5}{3}, \dots)$$

ist nicht monoton. Beachte jedoch, dass mit der Rekursionsformel

$$g_{n+2} = 1 + \frac{1}{g_{n+1}} = 1 + \frac{1}{1+\frac{1}{g_n}} = \frac{1+2g_n}{1+g_n} = 2 - \frac{1}{1+g_n}$$

folgt

$$g_{n+2} - g_n = \frac{1}{1+g_{n-2}} - \frac{1}{1+g_n} = \frac{g_n - g_{n-2}}{(1+g_n)(1+g_{n-2})}, \quad \forall n \ge 3.$$

Da $g_1 = 1 < g_3 = 3/2$, ist die Teilfolge $(g_{2n-1})_{n\in\mathbb{N}}$ somit monoton wachsend; analog ist wegen $g_2 = 2 > g_4 = 5/3$ die Teilfolge $(g_{2n})_{n\in\mathbb{N}}$ monoton fallend. Da (g_n) zudem beschränkt, existieren

$$a = \lim_{n\to\infty} g_{2n-1}, \quad b = \lim_{n\to\infty} g_{2n}.$$

Mit Satz 3.3.2 folgt $1 \le a, b \le 2$; weiter erhalten wir

$$a = \lim_{n\to\infty} g_{2n+1} = \lim_{n\to\infty} \Big(1 + \frac{1}{g_{2n}}\Big) = 1 + \frac{1}{b},$$
$$b = \lim_{n\to\infty} g_{2n} = \lim_{n\to\infty} \Big(1 + \frac{1}{g_{2n-1}}\Big) = 1 + \frac{1}{a}.$$

Multiplikation dieser Gleichungen mit b, bzw. mit a ergibt die Beziehung

$$ab = 1 + b = 1 + a,$$

also $a = b =: g$, und g löst die Gleichung

$$g = 1 + \frac{1}{g}$$

Setzen wir schliesslich noch $h = \frac{1}{g}$, so erhalten wir die Beziehung

$$\frac{1}{h} = g = 1 + \frac{1}{g} = 1 + h = \frac{1+h}{1}$$

und erkennen h als goldenen Schnitt.

Da jede Teilfolge $(g_l)_{l\in\Lambda}$ entweder unendlich viele Folgenglieder g_{2n} oder unendlich viele Folgenglieder g_{2n+1} enthält, folgt

$$g = \limsup_{n\to\infty} g_n = \liminf_{n\to\infty} g_n,$$

und $(g_n)_{n\in\mathbb{N}}$ ist konvergent mit $\lim_{n\to\infty} g_n = g$.

3.5 Cauchy-Kriterium

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$.

Definition 3.5.1. $(a_n)_{n\in\mathbb{N}}$ *heisst* **Cauchy-Folge**, *falls gilt*

$$\forall \epsilon > 0 \ \exists n_0 = n_0(\epsilon) \in \mathbb{N} \ \ \forall n, l \geq n_0 : \ |a_n - a_l| < \epsilon. \tag{3.5.1}$$

Satz 3.5.1. *(Cauchy-Kriterium) Für* $(a_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ *sind äquivalent:*

i) $(a_n)_{n\in\mathbb{N}}$ *ist konvergent,*

ii) $(a_n)_{n\in\mathbb{N}}$ *ist Cauchy-Folge.*

Beweis: *i)⇒ ii):* Sei $(a_n)_{n\in\mathbb{N}}$ konvergent, $a = \lim\limits_{n\to\infty} a_n$. Zu $\epsilon > 0$ wähle $n_0 \in \mathbb{N}$ mit

$$|a_n - a| < \epsilon, \qquad \forall n \geq n_0.$$

Es folgt

$$|a_n - a_l| \leq |a_n - a| + |a_l - a| < 2\epsilon, \qquad \forall n, l \geq n_0.$$

ii) ⇒ i): Sei $(a_n)_{n\in\mathbb{N}}$ Cauchy-Folge.

Behauptung 1 $(a_n)_{n\in\mathbb{N}}$ ist beschränkt.

Beweis. Zu $\epsilon = 1 > 0$ wähle n_0 mit

$$|a_n - a_l| < 1, \qquad \forall l, n \geq n_0.$$

Fixiere $n = n_0$. Dann folgt

$$|a_l| < |a_{n_0}| + 1, \qquad \forall l \geq n_0;$$

also

$$|a_l| < \max\big\{|a_1|, \ldots, |a_{n_0-1}|, |a_{n_0}| + 1\big\}.$$

□

Behauptung 2 $(a_n)_{n\in\mathbb{N}}$ ist konvergent.

Beweis. Nach Satz 3.4.1 gibt es $\Lambda \subset \mathbb{N}, \ a \in \mathbb{R}$ mit

$$a_l \to a \ \ (l \to \infty, \ l \in \Lambda).$$

Zu $\epsilon > 0$ wähle $n_0 = n_0(\epsilon)$ mit

$$\begin{aligned} |a_l - a| &< \epsilon, \quad \forall l \geq n_0, \ l \in \Lambda \\ |a_l - a_n| &< \epsilon, \quad \forall l, n \geq n_0. \end{aligned}$$

Für beliebiges $l \in \Lambda, \ l \geq n_0$ erhalten wir so die Abschätzung

$$|a_n - a| < |a_n - a_l| + |a_l - a| < 2\epsilon, \ \forall n \geq n_0;$$

d.h.

$$a_n \to a \quad (n \to \infty).$$

□

Beispiel 3.5.1. i) Die **harmonische Reihe**

$$a_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = \sum_{k=1}^{n} \frac{1}{k}, \quad n \in \mathbb{N},$$

ist divergent, da zum Beispiel gilt

$$a_{2n} - a_n = \frac{1}{n+1} + \cdots + \frac{1}{2n} \geq \frac{n}{2n} = \frac{1}{2}, \quad \forall n \in \mathbb{N}.$$

Also ist $(a_n)_{n\in\mathbb{N}}$ keine Cauchy-Folge und divergiert nach Satz 3.5.1.

ii) Die **alternierende harmonische Reihe**

$$a_n = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + (-1)^{n+1}\frac{1}{n} = \sum_{k=1}^{n} \frac{(-1)^{k+1}}{k}, \quad n \in \mathbb{N},$$

erfüllt wegen $\frac{1}{2k-1} - \frac{1}{2k} > 0$ für alle $k \in \mathbb{N}$ offenbar die Ungleichung

$$a_{2k-2} < a_{2k} < a_{2k-1} < a_{2k+1}, \ k \in \mathbb{N}, \tag{3.5.2}$$

und die Teilfolgen $(a_{2k})_{k\in\mathbb{N}}$, bzw. $(a_{2k+1})_{k\in\mathbb{N}}$ sind nach Satz 3.3.1 monoton konvergent. Da

$$|a_n - a_{n+1}| = \frac{1}{n+1}, \ n \in \mathbb{N}, \tag{3.5.3}$$

haben sie zudem denselben Limes a, und $(a_n)_{n\in\mathbb{N}}$ konvergiert nach Bemerkung 3.4.2.iii). Schliesslich erhalten wir mit (3.5.2) und (3.5.3) noch die Fehlerabschätzung

$$|a_n - a| \leq \frac{1}{n+1}, \ n \in \mathbb{N}\ .$$

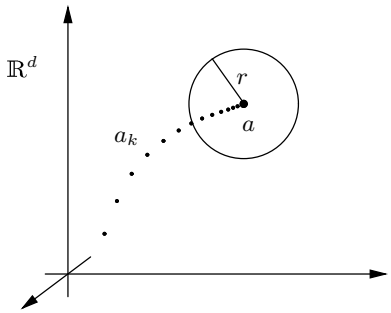
iii) Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge positiver Zahlen mit $a_{k-1} \geq a_k \to 0$ $(k \to \infty)$. Analog zu ii) besitzt dann die **Leibniz-Reihe**

$$S_n = a_1 - a_2 + a_3 - \cdots + (-1)^{n+1}a_n, \ n \in \mathbb{N},$$

einen Limes S, und es gilt die Fehlerabschätzung $|S_n - S| \leq a_{n+1}$, $n \in \mathbb{N}$.

3.6 Folgen in $\mathbb{R}^d$ oder $\mathbb{C}$

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}^d$ mit $a_n = (a_n^1, \ldots, a_n^d) \in \mathbb{R}^d$, $n \in \mathbb{N}$, und sei $a = (a^1, \ldots, a^d) \in \mathbb{R}^d$.



Definition 3.6.1. $a_n \to a \quad (n \to \infty)$, *falls* $\|a_n - a\| \to 0 \quad (n \to \infty)$.

Satz 3.6.1. *Es sind äquivalent*

i) $a_n \to a \ (n \to \infty)$,

ii) $\forall i \in \{1, \ldots, d\} : \ a_n^i \to a^i \ (n \to \infty)$.

Beweis. Für $x = (x^1, \ldots, x^d) \in \mathbb{R}^d$ gilt offenbar

$$\max_{1 \le i \le d} \left|x^i\right| \le \|x\| = \sqrt{\sum_{i=1}^{d} |x^i|^2} \le \sqrt{d} \max_{1 \le i \le d} \left|x^i\right|. \tag{3.6.1}$$

i) $\Rightarrow$ ii): Falls $a_n \to a \ (n \to \infty)$, folgt mit (3.6.1) für jedes $i \in \{1, \ldots, d\}$

$$\left|a_n^i - a^i\right| \le \|a_n - a\| \to 0 \quad (n \to \infty).$$

ii) $\Rightarrow$ i): Falls $a_n^i \to a^i \quad (n \to \infty)$ für $1 \le i \le d$, so

$$\max_{1 \le i \le d} \left|a_n^i - a^i\right| \le \sum_{i=1}^{d} \left|a_n^i - a^i\right| \to 0 \quad (n \to \infty)$$

gemäss Satz 3.3.2, also mit (3.6.1) auch $a_n \to a \ (n \to \infty)$. □

Mit Satz 3.5.1 und (3.6.1) folgt aus Satz 3.6.1:

Satz 3.6.2. *Es sind äquivalent:*

i) $(a_n)_{n \in \mathbb{N}}$ *konvergiert*

ii) $(a_n)_{n \in \mathbb{N}}$ *ist Cauchy-Folge.*

Definition 3.6.2. $(a_n)_{n \in \mathbb{N}}$ *ist* **beschränkt**, *falls gilt*

$$\exists C \in \mathbb{R} \ \ \forall n \in \mathbb{N} : \quad \|a_n\| \le C.$$

Mit Satz 3.4.1 und Satz 3.6.1 folgt

Satz 3.6.3. (Bolzano-Weierstrass) *Jede beschränkte Folge* $(a_n)_{n \in \mathbb{N}}$ *in* $\mathbb{R}^d$ *besitzt eine konvergente Teilfolge.*

Beweis. Für $1 \leq k \leq d$ schätze ab

$$\left|a_n^k\right| \leq \|a_n\| \leq C, \quad n \in \mathbb{N}.$$

Nach Satz 3.4.1 existieren Teilfolgen $N \supset \Lambda_1 \supset \cdots \supset \Lambda_d =: \Lambda$, $(a^1, \ldots, a^d) \in \mathbb{R}^d$ mit

$$a_n^k \to a^k \quad (n \to \infty,\ n \in \Lambda_k),\ 1 \leq k \leq d.$$

Mit Satz 3.6.1 folgt

$$a_n \to a := (a^1, \ldots, a^d) \quad (n \to \infty,\ n \in \Lambda).$$

□

3.7 Reihen

Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Betrachte die Folge $(S_n)_{n\in\mathbb{N}}$ der **Partialsummen**

$$S_n = a_1 + \cdots + a_n = \sum_{k=1}^{n} a_k, \quad n \in \mathbb{N}.$$

Definition 3.7.1. *Wir sagen,* **die Reihe $\sum_{k=1}^{\infty} a_k$ ist konvergent**, *falls*

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k =: \sum_{k=1}^{\infty} a_k$$

existiert.

Beispiel 3.7.1. i) Für $|q| < 1$ gilt

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

(Beweis: Induktion oder mit $(1-q)S_n = S_n - (S_{n+1} - 1) = 1 - q^{n+1}$.)

Also ist die **geometrische Reihe** $\sum_{k=0}^{\infty} q^k$ konvergent mit

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

ii) Die **harmonische Reihe** $\sum_{k=1}^{\infty} \frac{1}{k}$ ist nach Beispiel 3.5.1.i) divergent.

Konvergenzkriterien. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$ mit

$$S_n = \sum_{k=1}^{n} a_k, \quad n \in \mathbb{N}.$$

Der Satz 3.5.1 liefert sofort

Satz 3.7.1. *(Cauchy Kriterium) Die Reihe* $\sum_{k=1}^{\infty} a_k$ *ist konvergent genau dann, wenn gilt*

$$\left|\sum_{k=l}^{n} a_k\right| \to 0 \quad (n \geq l,\ l \to \infty).$$

Beweis. $|S_n - S_l| = \left|\sum_{k=l+1}^{n} a_k\right|$; benutze Satz 3.5.1. □

Bemerkung 3.7.1. i) Insbesondere ist die Bedingung $a_k \to 0 \ \ (k \to \infty)$ **notwendig** für Konvergenz von $\sum_{k=1}^{\infty} a_k$. (Wähle $n = l$ in Satz 3.7.1.)

ii) Die Bedingung $a_k \to 0 \ \ (k \to \infty)$ ist aber nicht hinreichend für Konvergenz, vergleiche die harmonische Reihe.

Im folgenden suchen wir möglichst leistungsfähige hinreichende Bedingungen für Konvergenz von $\sum_{k=1}^{\infty} a_k$. Alle aufgeführten Kriterien stützen sich auf den Vergleich mit der geometrischen Reihe.

Satz 3.7.2. (Quotientenkriterium) *Sei* $a_k \neq 0, \quad k \in \mathbb{N}$.

i) *Falls*

$$\limsup_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| < 1,$$

so ist $\sum_{k=1}^{\infty} a_k$ *konvergent;*

ii) *falls*

$$\liminf_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| > 1,$$

so ist $\sum_{k=1}^{\infty} a_k$ *divergent.*

Beweis. **i)** Setze

$$q_0 := \limsup_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{n\to\infty} \sup_{k\geq n} \left|\frac{a_{k+1}}{a_k}\right| < 1.$$

Wähle $q \in \mathbb{R}$ mit $q_0 < q < 1$. Dann gilt für genügend grosses $n_0 \in \mathbb{N}$

$$\sup_{k\geq n} \left|\frac{a_{k+1}}{a_k}\right| \leq q, \quad \forall n \geq n_0,$$

insbesondere bei Wahl von $n = n_0$ also

$$\left|\frac{a_{k+1}}{a_k}\right| \leq q, \quad \forall k \geq n_0.$$

Es folgt für $k \geq n_0$ die Abschätzung

$$|a_k| = \Big|\underbrace{\frac{a_k}{a_{k-1}} \cdot \frac{a_{k-1}}{a_{k-2}} \cdots \frac{a_{n_0+1}}{a_{n_0}}}_{(k-n_0) \text{ Faktoren}} a_{n_0}\Big| \leq q^{k-n_0} |a_{n_0}| = Cq^k,$$

wobei $C = q^{-n_0} |a_{n_0}|$. Für $n \geq l \geq n_0$ erhalten wir somit

$$\left|\sum_{k=l}^{n} a_k\right| \leq \sum_{k=l}^{n} |a_k| \leq C \sum_{k=l}^{n} q^k \leq Cq^l \frac{1}{1-q} \to 0 \quad (n \geq l,\ l \to \infty),$$

und $\sum\limits_{k=1}^{\infty} a_k$ ist konvergent nach Satz 3.7.1.

ii) Es gelte

$$\liminf_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{n\to\infty} \inf_{k\geq n} \left|\frac{a_{k+1}}{a_k}\right| > 1.$$

Dann existiert n_0 mit

$$\left|\frac{a_{k+1}}{a_k}\right| \geq \inf_{k\geq n_0} \left|\frac{a_{k+1}}{a_k}\right| \geq 1, \quad \forall k \geq n_0,$$

also

$$|a_k| = \left|\frac{a_k}{a_{k-1}}\right| \ldots \left|\frac{a_{n_0+1}}{a_{n_0}}\right| |a_{n_0}| \geq |a_{n_0}| > 0, \quad \forall k \geq n_0$$

und $\sum\limits_{k=1}^{\infty} a_k$ ist divergent nach Bemerkung 3.7.1.i).

□

Beispiel 3.7.2. i) Die **Exponentialreihe**

$$Exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

konvergiert für jedes $z \in \mathbb{C}$.

Beweis. OBdA sei $z \neq 0$, also auch $a_k := \frac{z^k}{k!} \neq 0$, $\forall k \in \mathbb{N}_0$. Wegen

$$\left|\frac{a_{k+1}}{a_k}\right| = \frac{|z|}{k+1} \to 0 \quad (k \to \infty)$$

erhalten wir mit Satz 3.7.2 Konvergenz. □

ii) Für welche $z \in \mathbb{C}$ konvergiert die Reihe

$$f(z) = \sum_{k=0}^{\infty} \frac{z^k k!}{k^k} \quad ?$$

OBdA sei $z \neq 0$. Setze $a_k = \frac{z^k k!}{k^k}$, $k \in \mathbb{N}$, mit

$$\frac{a_{k+1}}{a_k} = z\Big(\frac{k}{k+1}\Big)^k = z\frac{1}{(1+\frac{1}{k})^k} \overset{(k\to\infty)}{\to} \frac{z}{e}.$$

Mit Satz 3.7.2 folgt Konvergenz von $f(z)$ für $|z| < e$, Divergenz für $|z| > e$.

Offen bleibt das Verhalten von $f(z)$ für $|z| = e$, d.h. auf dem Rand des „Konvergenzkreises“.

Das Quotientenkriterium versagt, wenn unendlich viele a_k verschwinden oder falls die Folge der Quotienten $\frac{a_{k+1}}{a_k}$ stark oszilliert. Um auch solche Fälle behandeln zu können, benötigen wir ein leistungsfähigeres Kriterium. In der Tat ist das folgende Kriterium nahezu bestmöglich.

Satz 3.7.3. (Wurzelkriterium) *Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$.*

i) *Falls* $\limsup_{k\to\infty} \sqrt[k]{|a_k|} < 1$, *so konvergiert* $\sum_{k=1}^{\infty} a_k$;

ii) *falls* $\limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1$, *so divergiert* $\sum_{k=1}^{\infty} a_k$.

Beweis. **i)** Für $q \in \mathbb{R}$ mit

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|} < q < 1$$

und genügend grosses $n_0 \in \mathbb{N}$ gilt

$$\sqrt[k]{|a_k|} \leq q, \quad \forall k \geq n_0,$$

also

$$|a_k| \leq q^k, \quad \forall k \geq n_0.$$

Für $n \geq l \geq n_0$ folgt

$$\left|\sum_{k=l}^{n} a_k\right| \leq \sum_{k=l}^{\infty} q^k = \frac{q^l}{1-q} \to 0 \quad (l \to \infty),$$

und $\sum_{k=1}^{\infty} a_k$ ist konvergent nach Satz 3.7.1.

ii) Falls $\limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1$, so gibt es für alle $n_0 \in \mathbb{N}$ ein $k \geq n_0$ mit $\sqrt[k]{|a_k|} \geq 1$, also $\limsup_{k\to\infty} |a_k| \geq 1$, und $\sum_{k=1}^{\infty} a_k$ ist divergent nach Bemerkung 3.7.1.i).

□

Anwendung: Sei $(c_k)_{k\in\mathbb{N}_0}$ Folge in $\mathbb{R}$ oder $\mathbb{C}$. Betrachte die **Potenzreihe** in $z \in \mathbb{C}$:

$$p(z) := c_0 + c_1 z + c_2 z^2 + \cdots = \sum_{k=0}^{\infty} c_k z^k.$$

Setze $a_k = c_k z^k$, $k \in \mathbb{N}$, mit

$$\sqrt[k]{|a_k|} = |z| \cdot \sqrt[k]{|c_k|}, \quad k \in \mathbb{N}.$$

Mit Satz 3.7.3 erhalten wir sofort die folgende Charakterisierung des Konvergenzbereichs von p.

Satz 3.7.4. *Die Potenzreihe* $p(z) = \sum_{k=0}^{\infty} c_k z^k$ *ist konvergent für alle* $z \in \mathbb{C}$ *mit*

$$|z| < \rho := \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}} \in [0, \infty],$$

divergent für alle $|z| > \rho$.

Bemerkung 3.7.2. Insbesondere ist der Konvergenzbereich von p ein Kreis. Satz 3.7.4 liefert zudem eine **präzise** Charakterisierung des Konvergenzradius, was das Quotientenkriterium nicht zu leisten vermag.

Beispiel 3.7.3. Falls wir als Koeffizienten die Zahlen

$$c_k = \begin{cases} 1, & k \text{ ungerade} \\ \frac{1}{k}, & k \text{ gerade} \end{cases}$$

wählen, so liefert das Wurzelkriterium Konvergenz der Reihe $\sum_{k=0}^{\infty} c_k z^k$ für $|z| < 1$, während das Quotientenkriterium keine Aussage ermöglicht.

Beispiel 3.7.4. (Zeta-Funktion). Für $s > 0$ betrachte die Reihe

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}.$$

Für welche $s > 0$ ist $\zeta(s)$ konvergent?

i) Offenbar gilt für $s \leq 1$ stets

$$\sum_{k=1}^{n} \frac{1}{k^s} \geq \sum_{k=1}^{n} \frac{1}{k} \to \infty \quad (n \to \infty).$$

Also ist $\zeta(s)$ für solche s divergent.

ii) Sei $s > 1$. Setze $a_k = \frac{1}{k^s}$, $k \in \mathbb{N}$. Beachte

$$\frac{a_{k+1}}{a_k} = \Big(\frac{k}{k+1}\Big)^s \to 1 \quad (k \to \infty),$$
$$\sqrt[k]{k^s} = (\sqrt[k]{k})^s \to 1 \quad (k \to \infty);$$

Quotienten- und Wurzelkriterium versagen also. Zerlegen wir jedoch für beliebiges $L \in \mathbb{N}$ die Summe über $1 \leq k \leq 2^{L+1} - 1$ "dyadisch" in L Teilsummen

über $2^l \le k \le 2^{l+1}-1$, $0 \le l \le L$, so erhalten wir

$$\sum_{k=1}^{2^{L+1}-1} \frac{1}{k^s} = \sum_{l=0}^{L} \Big(\sum_{k=2^l}^{2^{l+1}-1} \frac{1}{k^s} \Big) = \sum_{l=0}^{L} b_l$$

mit

$$b_l = \sum_{k=2^l}^{2^{l+1}-1} \frac{1}{k^s} \le 2^l \cdot \frac{1}{2^{ls}} = 2^{l(1-s)} = q^l,$$

wobei $q = 2^{1-s} < 1$. Es folgt, $\zeta(s)$ konvergiert für $s > 1$. Zum Beispiel gilt $\zeta(2) = \frac{\pi^2}{6}$; aber niemand weiss, welchen Wert $\zeta(3)$ hat.

Die Zeta-Funktion steht im Mittelpunkt einer der berühmtesten Vermutungen der Zahlentheorie ("Riemannsche Vermutung", "Riemann hypothesis").

3.8 Absolute Konvergenz

Sei $(a_n)_{n\in\mathbb{N}}$ Folge in $\mathbb{R}$ oder $\mathbb{C}$.

Definition 3.8.1. *Die Reihe* $\sum_{k=1}^{\infty} a_k$ *konveriert* **absolut**, *falls* $\sum_{k=1}^{\infty} |a_k|$ *konvergiert.*

Bemerkung 3.8.1. i) Das Quotienten- und Wurzelkriterium sind Kriterien für absolute Konvergenz.

ii) Da die Partialsummen $S_n = \sum_{k=1}^{n} |a_k| \le S_{n+1}$ monoton wachsen, genügt nach Satz 3.3.1 die Beschränktheit dieser Folge für Konvergenz. **Notation:**

$$\sum_{k=1}^{\infty} |a_k| < \infty \Leftrightarrow \sum_{k=1}^{\infty} |a_k| \text{ konvergiert.}$$

Wir wollen Reihen $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ miteinander multiplizieren. Dabei kann man die Produkte $a_k b_l$ offenbar in sehr unterschiedlicher Weise summieren. Kommt es auf die Reihenfolge an?

Beispiel 3.8.1. $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ konvergiert nach Beispiel 3.5.1.ii), jedoch ist $\sum_{k=1}^{\infty} \frac{1}{2k}$ nach Beispiel 3.5.1.i) divergent. Wählt man zu $l \in \mathbb{N}$ den Index k_l so, dass $\sum_{k=1}^{k_l} \frac{1}{2k} > l + \sum_{k=1}^{l} \frac{1}{2k-1}$, und summiert man für jedes l die ersten k_l positiven Glieder der Folge $(\frac{(-1)^k}{k})_{k\in\mathbb{N}}$, wobei man nach jeweils k_j positiven Gliedern der Folge das j-te negative Folgenglied abzieht, so erhält man für die entsprechenden Partialsummen S_{k_l+l} der so umgeordneten alternierenden harmonischen Reihe

den Wert

$$S_{k_l+l} = \frac{1}{2} + \cdots + \frac{1}{2k_1} - 1 + \frac{1}{2(k_1+1)} + \cdots + \frac{1}{2k_2} - \frac{1}{3} + \cdots + \cdots - \ldots$$
$$+ \frac{1}{2k_{l-1}+1} + \cdots + \frac{1}{2k_l} - \frac{1}{2l-1} = \sum_{k=1}^{k_l} \frac{1}{2k} - \sum_{k=1}^{l} \frac{1}{2k-1} > l, \ \forall l \in \mathbb{N};$$

d.h. die so umgeordnete alternierende harmonische Reihe ist divergent!

Zunächst beantworten wir die obige Frage für **eine** Reihe.

Satz 3.8.1. *Sei $\sum_{k=1}^{\infty} a_k$ absolut konvergent, und sei $\varphi : \mathbb{N} \to \mathbb{N}$ bijektiv. Dann ist auch die "umgeordnete" Reihe $\sum_{k=1}^{\infty} a_{\varphi(k)}$ konvergent, und es gilt*

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{k=1}^{\infty} a_k.$$

Beweis. Zum Beweis der Konvergenz von $\sum_{k=1}^{\infty} a_{\varphi(k)}$ sei $\epsilon > 0$ vorgegeben. Da $\sum_{k=1}^{\infty} |a_k|$ konvergiert, gibt es $n_0 = n_0(\epsilon)$ mit

$$\sum_{k=n_0}^{\infty} |a_k| < \epsilon.$$

Setze $n_1 = \max\{\varphi^{-1}(1), \ldots, \varphi^{-1}(n_0)\}$. Da φ injektiv, folgt

$$\varphi(k) > n_0, \quad \forall k > n_1;$$

also für $n, l > n_1$:

$$\left| \sum_{k=n}^{l} a_{\varphi(k)} \right| \leq \sum_{k=n_0}^{\infty} |a_k| < \epsilon.$$

Also konvergiert $\sum_{k=1}^{\infty} a_{\varphi(k)}$. Weiter gilt

$$\left| \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_{\varphi(k)} \right| \leq \left| \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{n_1} a_{\varphi(k)} \right| + \left| \sum_{k=n_1+1}^{\infty} a_{\varphi(k)} \right| \leq 2 \sum_{k=n_0}^{\infty} |a_k| \leq 2\epsilon.$$

□

Beispiel 3.8.2. $\sum_{k=1}^{\infty} kq^k$ ist für $|q| < 1$ gemäss Satz 3.7.3 absolut konvergent. Mit Satz 3.8.1 dürfen wir die Terme der Reihe auch wie folgt zerlegen, um den

Wert dieser Reihe zu bestimmen.

$$\begin{aligned}\sum_{k=1}^{\infty} kq^k &= q + q^2 + q^3 + \dots \\ &\quad + q^2 + q^3 + \dots \\ &\quad + q^3 + \dots \\ &\quad \dots \\ &= \sum_{l=1}^{\infty}\Big(q^l \sum_{k=0}^{\infty} q^k\Big) = \frac{q}{(1-q)^2}.\end{aligned}$$

(Umordnung der Summation: Statt $\downarrow$ wird $\rightarrow$ summiert.)

Satz 3.8.2. *Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{R}$ oder $\mathbb{C}$, und seien die Reihen $\sum_{k=1}^{\infty} a_k$, $\sum_{k=1}^{\infty} b_k$ absolut konvergent. Dann konvergiert die Reihe der Produkte absolut mit*

$$\sum_{k,l=1}^{\infty} a_k b_l = \sum_{k=1}^{\infty} a_k \cdot \sum_{l=1}^{\infty} b_l,$$

unabhängig von der Summationsreihenfolge.

Beweis. Nach Satz 3.8.1 genügt es, die absolute Konvergenz der Reihe $\sum_{k,l=1}^{\infty} a_k b_l$ für **eine** Summationsreihenfolge zu zeigen. Mit Satz 3.3.2 erhalten wir dann auch

$$\sum_{k,l=1}^{\infty} a_k b_l = \lim_{n\to\infty} \sum_{k,l=1}^{n} a_k b_l = \lim_{n\to\infty}\Big(\sum_{k=1}^{n} a_k \cdot \sum_{l=1}^{n} b_l\Big) = \sum_{k=1}^{\infty} a_k \cdot \sum_{l=1}^{\infty} b_l,$$

und die Aussage des Satzes ist bewiesen. Somit folgt der Satz aus

Behauptung $\sum_{l=1}^{\infty}\Big(\sum_{k=1}^{\infty} |a_k b_l|\Big)$ konvergiert absolut.

Beweis. Setze $C_0 := \sum_{k=1}^{\infty} |a_k|$. (OBdA sei $C_0 > 0$). Zu $\epsilon > 0$ wähle $l_0 = l_0(\epsilon)$ mit

$$\sum_{l=l_0}^{\infty} |b_l| < \epsilon/C_0.$$

Für $n \geq m \geq l_0$ folgt mit Satz 3.3.2

$$\sum_{l=m}^{n}\Big(\sum_{k=1}^{\infty} |a_k b_l|\Big) = \sum_{l=m}^{n}\Big(|b_l| \overbrace{\sum_{k=1}^{\infty} |a_k|}^{\leq C_0}\Big) \leq C_0 \sum_{l=m}^{n} |b_l| \leq C_0 \sum_{l=l_0}^{\infty} |b_l| < \epsilon.$$

□

Beispiel 3.8.3. Sei

$$Exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

wie in Beispiel 3.7.2.i). Für $x, y \in \mathbb{C}$ gilt nach Satz 3.8.2

$$\begin{aligned} Exp(x) \cdot Exp(y) &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \sum_{l=0}^{\infty} \frac{y^l}{l!} \\ &= \sum_{k,l=0}^{\infty} \frac{x^k y^l}{k!\ l!} = \sum_{k=0}^{\infty} \Big(\sum_{l=0}^{\infty} \frac{x^k y^l}{k!\ l!} \Big) = (*). \end{aligned}$$

Substituiere bei festem k den Laufindex l durch die neue Summationsvariable $n := k + l$; d.h. ersetze l durch $n - k$. Wir erhalten

$$\begin{aligned} (*) &= \sum_{k=0}^{\infty} \Big(\sum_{n=k}^{\infty} \frac{x^k y^{n-k}}{k!\ (n-k)!} \Big) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \binom{n}{k} \frac{x^k y^{n-k}}{n!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{\Big(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} \Big)}_{=(x+y)^n} = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}. \end{aligned}$$

Es folgt

Korollar 3.8.1. (Additionstheorem für Exp) *Für $x, y \in \mathbb{C}$ gilt*

$$Exp(x) \cdot Exp(y) = Exp(x + y).$$

3.9 Die Exponentialreihe und die Funktion e^x

Die Exponentialreihe hat interessante Eigenschaften.

Satz 3.9.1. *Es gilt*

$$Exp(1) = \sum_{k=1}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \cdots = e = \lim_{n \to \infty} \Big(1 + \frac{1}{n}\Big)^n.$$

Beweis. Nach dem binomischen Lehrsatz gilt

$$\begin{aligned} \big(1 + \frac{1}{n}\big)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{n!}{k!\ (n-k)!\ n^k} \\ &= \sum_{k=0}^{n} \frac{1}{k!} \underbrace{\frac{n(n-1)\ldots(n-k+1)}{n^k}}_{=:a_k^{(n)}} < Exp(1), \end{aligned}$$

mit

$$0 < a_k^{(n)} < 1, \quad a_k^{(n)} \to 1 \quad (n \to \infty,\ k \text{ fest}).$$

Zu $\epsilon > 0$ wähle $n_0 = n_0(\epsilon)$ mit

$$\sum_{k=0}^{n_0} \frac{1}{k!} > Exp(1) - \epsilon,$$

dazu weiter $n_1 = n_1(\epsilon,\ n_0)$ mit

$$\sum_{k=0}^{n_0} 1 - a_k^{(n)} < \epsilon, \quad \forall n \geq n_1.$$

OBdA $n_1 \geq n_0$. Es folgt für $n \geq n_1 \geq n_0$:

$$0 < Exp(1) - \big(1 + \frac{1}{n}\big)^n < \sum_{k=0}^{n_0} \frac{1}{k!}(1 - a_k^{(n)}) + \epsilon < 2\epsilon.$$

□

Mit Korollar 3.8.1 folgt induktiv

$$Exp(n) = Exp(1) \cdot Exp(n-1) = \cdots = (Exp(1))^n = e^n, \quad \forall n \in \mathbb{N}.$$

Unter Beachtung von

$$Exp(n) \cdot Exp(-n) = Exp(0) = 1$$

erhalten wir dann weiter

$$Exp(-n) = \frac{1}{Exp(n)} = e^{-n}, \quad \forall n \in \mathbb{N}.$$

Analog gilt für $p, q \in \mathbb{N}$

$$Exp\Big(\frac{p}{q}\Big) = \Big(Exp\big(\frac{1}{q}\big)\Big)^p; \tag{3.9.1}$$

insbesondere folgt für $p = q \in \mathbb{N}$ zunächst

$$Exp\big(\frac{1}{q}\big) = \big(Exp(1)\big)^{\frac{1}{q}} = e^{\frac{1}{q}},$$

und mit (3.9.1) dann auch

$$Exp\Big(\frac{p}{q}\Big) = \big(e^{\frac{1}{q}}\big)^p = e^{\frac{p}{q}}, \quad \forall p, q \in \mathbb{N};$$

d.h. wir erhalten

Satz 3.9.2. *$Exp(x) = e^x$, $\forall x \in \mathbb{Q}$.*

Für rein imaginäre Argumente $z = iy$, $y \in \mathbb{R}$, können wir $Exp(iy)$ durch Umordnung gemäss Satz 3.8.1 in Real- und Imaginärteil zerlegen

$$\begin{aligned} Exp(iy) &= \sum_{k=0}^{\infty} \frac{(iy)^k}{k!} \\ &= \sum_{l=0}^{\infty} \frac{(-1)^l y^{2l}}{(2l)!} + i \sum_{l=0}^{\infty} \frac{(-1)^l y^{2l+1}}{(2l+1)!} \\ &=: Cos(y) + iSin(y). \end{aligned}$$

Mit Korollar 3.8.1 erhalten wir dann für $z = x + iy \in \mathbb{C}$

$$\begin{aligned} Exp(x+iy) &= Exp(x) \cdot Exp(iy) \\ &= Exp(x)\big(Cos(y) + iSin(y)\big). \end{aligned}$$

Bemerkung 3.9.1. Später werden wir Cos und Sin als die trigonometrischen Funktionen cos und sin wiedererkennen und die Identität herleiten

$$Exp(x+iy) = e^x\big(\cos(y) + i\sin(y)\big), \quad \forall z = x+iy \in \mathbb{C}.$$

Kapitel 4

Stetigkeit

4.1 Grenzwerte von Funktionen

Bisher konnten wir mit Satz 3.3.2 Grenzwerte von Ausdrücken der Form

$$y_k = \frac{a_k b_k + c_k}{d_k} \text{ mit } a_k \to a,\ b_k \to b,\ c_k \to c,\ d_k \to d \neq 0\ (k \to \infty)$$

behandeln. Die Punkte $x_k := (a_k, b_k, c_k, d_k)$ können wir auffassen als Elemente $x_k \in \mathbb{R}^4$ mit $x_k \to x_0 = (a, b, c, d)$ $(k \to \infty)$.

Allgemein untersuchen wir nun für eine Funktion $f\colon \Omega \to \mathbb{R}^n$ auf einer Menge $\Omega \subset \mathbb{R}^d$ die Konvergenz von Folgen $y_k = f(x_k)$, wobei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$. Der Limes x_0 der Folge $(x_k)_{k\in\mathbb{N}}$ muss dabei nicht notwendig selbst wieder in Ω liegen.

Beispiel 4.1.1. Sei $f(x) = \frac{x^2-1}{x-1}$, $x \neq 1$. Die Funktion f hat eine "Definitionslücke" bei $x = 1$. Wegen $x^2 - 1 = (x+1)(x-1)$ kann man jedoch für $x \neq 1$ den Faktor $(x-1)$ kürzen. Für eine Folge $1 \neq x_k \to 1$ für $k \to \infty$ erhalten wir so

$$f(x_k) = x_k + 1 \to 2 \quad (k \to \infty).$$

Wir könne also die Funktion f durch $f(1) = 2$ an der Stelle $x = 1$ "stetig ergänzen".

Definition 4.1.1. *Sei $\Omega \subset \mathbb{R}^d$. Der* **Abschluss** *von Ω ist die Menge*

$$\overline{\Omega} = \{x \in \mathbb{R}^d;\ \exists (x_k)_{k\in\mathbb{N}} \subset \Omega :\ x_k \to x\ (k \to \infty)\}.$$

Bemerkung 4.1.1. Offenbar gilt $\Omega \subset \overline{\Omega}$. (Zu $x \in \Omega$ betrachte die konstante Folge $x_k = x \to x$ $(k \to \infty)$.)

Beispiel 4.1.2. i) $\overline{\mathbb{R} \setminus \{1\}} = \mathbb{R}$.

ii) $\overline{]0, 1[} = [0, 1]$.

iii) Sei $B_r(x_0) = \{x \in \mathbb{R}^d;\ \|x - x_0\| < r\}$. Dann gilt

$$\overline{B_r(x_0)} = \{x \in \mathbb{R}^d;\ \|x - x_0\| \le r\}.$$

Beweis. OBdA sei $x_0 = 0$.

"$\supset$": Für $x \in \mathbb{R}^d$ mit $\|x\| \le r$ setze $x_k = \big(1 - \frac{1}{k}\big)x$, $k \in \mathbb{N}$, mit

$$\|x_k\| = \big(1 - \frac{1}{k}\big)\,\|x\| < r,\ x_k \to x\ (k \to \infty).$$

"$\subset$": Falls $\|x\| > r$, $x_k \to x$ $(k \to \infty)$, so existiert k_0 mit

$$\|x_k\| \ge \|x\| - \|x - x_k\| > r\ (k \ge k_0).$$

Es folgt $x_k \notin B_r(0)$ für $k \ge k_0$, und $x \notin \overline{B_r(0)}$. □

iv) $\overline{\mathbb{Q}} = \mathbb{R}$.

Beweis. Zu $x_0 \notin \mathbb{Q}$, $k \in \mathbb{N}$ sei x_k die an der k-ten Nachkommastelle abgebrochene Dezimaldarstellung von x_0. Offenbar gilt $x_k \in \mathbb{Q}$ für alle $k \in \mathbb{N}$ und $x_k \to x_0$ $(k \to \infty)$. □

v) $\overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}$.

Beweis. Sei $x_0 \in \mathbb{Q}$. Dann gilt $y_0 = x_0 + \sqrt{2} \notin \mathbb{Q}$. Für $k \in \mathbb{N}$ sei y_k wie in iv) die an der k-ten Nachkommastelle abgebrochene Dezimaldarstellung von y_0. Wir erhalten so die Folge $x_k = y_k - \sqrt{2} \notin \mathbb{Q}$, $k \in \mathbb{N}$, und $x_k \to x_0$ $(k \to \infty)$. □

Sei $f\colon \Omega \to \mathbb{R}^n$, $x_0 \in \overline{\Omega}$, $a \in \mathbb{R}^n$.

Definition 4.1.2. *f hat an der Stelle x_0 den* **Grenzwert** *a, falls für jede Folge $(x_k)_{k\in\mathbb{N}}$ in Ω mit $x_k \to x_0$ $(k \to \infty)$ gilt $f(x_k) \to a$ $(k \to \infty)$.*

Notation: $\lim\limits_{x \to x_0} f(x) = a$.

Bemerkung 4.1.2. Falls f einen Grenzwert besitzt an einer Stelle $x_0 \in \Omega$, so muss gelten $\lim\limits_{x \to x_0} f(x) = f(x_0)$. (Betrachte die konstante Folge $x_k = x_0 \in \Omega$, $k \in \mathbb{N}$.)

Definition 4.1.3. *Sei $\Omega \subset \mathbb{R}^d$, $f\colon \Omega \to \mathbb{R}^n$.*

i) *f heisst* **stetig an der Stelle** *$x_0 \in \Omega$, falls $a := \lim\limits_{x \to x_0} f(x)$ existiert (mit $a = f(x_0)$ gemäss Bemerkung 4.1.2).*

ii) *f heisst* **an der Stelle** *$x_0 \in \overline{\Omega}\setminus\Omega$* **stetig ergänzbar**, *falls $\lim\limits_{x \to x_0} f(x) =: a$ existiert. (In diesm Fall ist die durch $f(x_0) = a$ ergänzte Funktion f offenbar stetig an der Stelle x_0.)*

Beispiel 4.1.3. i) Sei $p\colon \mathbb{C} \to \mathbb{C}$ das Polynom

$$p(z) = a_0 + a_1 z + \cdots + a_n z^n,\quad z \in \mathbb{C},$$

mit Koeffizienten $a_l \in \mathbb{C}$, $l = 0, \ldots, n$. Nach Satz 3.3.2 gilt für $z_k \to z_0$ $(k \to \infty)$

$$p(z_k) = a_0 + a_1 z + \cdots + a_n z_k^n \to p(z_0) \ (k \to \infty).$$

Also besitzt p an jeder Stelle $z_0 \in \mathbb{C}$ den Grenzwert $p(z_0)$ und p ist stetig an jeder Stelle $z \in \mathbb{C}$.

ii) Eine rationale Funktion p/q mit Polynomen $p, q \colon \mathbb{C} \to \mathbb{C}$ ist stetig in jedem Punkt z des natürlichen Definitionsbereichs $\Omega = \{z \in \mathbb{C};\ q(z) \neq 0\}$.

iii) Sei $f(x) = \frac{x^2-1}{x-1}$, $x \neq 1$. Wie in Beispiel 4.1.1 gezeigt, können wir die Funktion f an der Stelle $x = 1$ durch $f(1) = 2$ stetig ergänzen.

iv) Sei $\Omega = \mathbb{R}\backslash\{0\}$, $f(x) = 1/x$, $x \neq 0$. Dann gilt für $x_k \to x_0 \neq 0$ nach Satz 3.3.2

$$f(x_k) = \frac{1}{x_k} \to \frac{1}{x_0} = f(x_0) \quad (k \to \infty).$$

An der Stelle $x_0 = 0$ besitzt f jedoch keinen Grenzwert. (Betrachte zum Beispiel $x_k = \frac{1}{k} \to 0$ $(k \to \infty)$ mit $f(x_k) = k \to \infty$ $(k \to \infty)$.)

v) Sei $\Omega = \mathbb{R}$, $\chi_{\mathbb{Q}} \colon \mathbb{R} \to \mathbb{R}$ die "charakteristische Funktion" von $\mathbb{Q}$ mit

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q}, \end{cases}$$

Dann besitzt $\chi_{\mathbb{Q}}$ an keiner Stelle $x_0 \in \mathbb{R}$ einen Grenzwert, denn zu jedem $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ gibt es gemäss Beispiel 4.1.2.iv) eine Folge $(x_k)_{k\in\mathbb{N}} \subset \mathbb{Q}$ mit $x_k \to x_0$ $(k \to \infty)$ und $\lim_{k\to\infty} \chi_{\mathbb{Q}}(x_k) = 1 \neq \chi_{\mathbb{Q}}(x_0) = 0$; analog für $x_0 \in \mathbb{Q}$.

vi) Die stückweise konstante Funktion $f \colon \mathbb{R}\backslash\{0\} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} a, & x < 0 \\ b, & x > 0 \end{cases}$$

ist stetig an jeder Stelle $x_0 \neq 0$; sie ist jedoch für $a \neq b$ an der Stelle $x_0 = 0$ nicht stetig ergänzbar.

vii) Sei $-\infty \leq a < b \leq \infty$, und sei $f \colon]a, b[\to \mathbb{R}$ monoton wachsend. Dann existieren für jedes $x_0 \in]a, b[$ die links- und rechtsseitigen Grenzwerte

$$f(x_0^+) := \lim_{x\to x_0,\ x>x_0} f(x), \quad f(x_0^-) := \lim_{x\to x_0,\ x<x_0} f(x),$$

und f ist stetig an der Stelle x_0 genau dann, wenn $f(x_0^-) = f(x_0^+) = f(x_0)$.

Beweis. Sei $x_0 \in]a, b[$. Falls $(x_k)_{k\in\mathbb{N}} \subset]a, b[$ mit

$$x_k < x_{k+1} \to x_0 \quad (k \to \infty),$$

so ist die Folge $(f(x_k))_{k\in\mathbb{N}}$ monoton wachsend und beschränkt. Gemäss Satz 3.3.1 existiert

$$\lim_{k\to\infty} f(x_k) =: a.$$

Wir zeigen, dass der Limes unabhängig ist von der gewählten Folge.

Behauptung $a = \lim\limits_{x \to x_0,\ x<x_0} f(x)$.

Beweis. Sei $(y_k)_{k\in\mathbb{N}} \subset]a,b[$ mit $y_k \to x_0$ $(k \to \infty)$, $y_k < x_0$ für alle $k \in \mathbb{N}$. Zu $\epsilon > 0$ gibt es $k_0 \in \mathbb{N}$ mit

$$a - \epsilon < f(x_k) < a, \quad \forall k \geq k_0.$$

Da $y_k \to x_0$ $(k \to \infty)$, $x_k < x_0$ $(k \in \mathbb{N})$, gibt es $k_1 \in \mathbb{N}$ mit

$$x_{k_0} < y_k < x_0, \quad \forall k \geq k_1.$$

Zusammen mit der Monotonie von f folgt

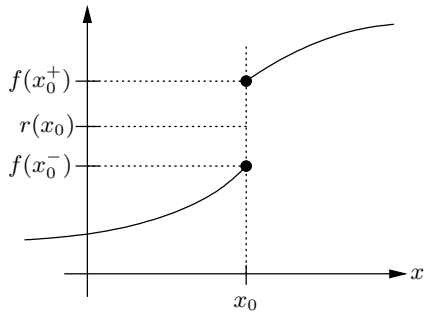
$$a - \epsilon < f(x_{k_0}) < f(y_k) < \lim_{k\to\infty} f(x_k) = a, \quad \forall k \geq k_1;$$

d.h.

$$f(y_k) \to a \quad (k \to \infty).$$

□

Analog existiert $f(x_0^+)$. Offenbar gilt $\lim\limits_{x\to x_0} f(x) = f(x_0)$ genau dann, wenn $f(x_0^-) = f(x_0^+) = f(x_0)$. □



Satz 4.1.1. *Sei $f\colon]a,b[\to \mathbb{R}$ monoton wachsend. Dann ist f in höchstens abzählbar vielen Punkten unstetig.*

Beweis. Gemäss Beispiel 4.1.3.vii) existiert für jede Unstetigkeitsstelle $x_0 \in]a,b[$ stets ein $r = r(x_0) \in \mathbb{Q}$ mit

$$f(x_0^-) < r < f(x_0^+).$$

Wegen der Monotonie von f gilt zudem für je zwei Unstetigkeitsstellen $x_0 < y_0$ von f auch stets $r(x_0) < r(y_0)$; die Abbildung

$$\{x_0 \in]a,b[;\, f \text{ ist unstetig an der Stelle } x_0\} \ni x_0 \mapsto r(x_0) \in \mathbb{Q}$$

ist also injektiv. Die Behauptung folgt. □

Definition 4.1.4. *Eine Funktion $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ heisst* **Lipschitz stetig** *mit* **Lipschitzkonstante** *L, falls gilt*

$$\|f(x) - f(y)\| \leq L \ \|x - y\|\,, \quad \forall x, y \in \Omega. \tag{4.1.1}$$

Beispiel 4.1.4. i) Die euklidsche Norm $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ ist Lipschitz-stetig mit Lipschitz-Konstante $L = 1$.

Beweis. Gemäss Satz 2.4.2.iii) gilt $\|x\| - \|y\| \leq \|x - y\|$ für alle $x, y \in \mathbb{R}^d$. Nach Vertauschen von x und y folgt die Ungleichung

$$\big|\, \|x\| - \|y\| \,\big| \leq \|x - y\| \ .$$

□

ii) Die Funktionen $\mathbb{R} \ni x \mapsto x^{\pm} = \max\{0, \pm x\}$ sind Lipschitz-stetig mit Lipschitz-Konstante $L = 1$.

iii) Die Addition $\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x + y \in \mathbb{R}^n$ ist Lipschitz-stetig mit Lipschitz-Konstante $L = \sqrt{2}$.

Beweis. Beachte die Äquivalenz $\mathbb{R}^n \times \mathbb{R}^n \sim \mathbb{R}^{2n}$ mit

$$\|(x, y)\|^2_{\mathbb{R}^n \times \mathbb{R}^n} = \|x\|^2_{\mathbb{R}^n} + \|y\|^2_{\mathbb{R}^n}\,, \quad \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Für $(x, y),\ (x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^n$ folgt

$$\begin{aligned} ||(x + y) - (x_0 + y_0)||^2_{\mathbb{R}^n} &\leq \big(\|x - x_0\|_{\mathbb{R}^n} + \|y - y_0\|_{\mathbb{R}^n}\big)^2 \\ &\leq 2\big(\|x - x_0\|^2_{\mathbb{R}^n} + \|y - y_0\|^2_{\mathbb{R}^n}\big) = 2\,\|(x, y) - (x_0, y_0)\|^2_{\mathbb{R}^n \times \mathbb{R}^n}\,, \end{aligned}$$

wobei wir die elementare Ungleichung $(a+b)^2 \leq 2a^2 + 2b^2$ für $a, b \in \mathbb{R}$ benutzen, welche unmittelbar aus Satz 2.2.2 folgt. □

Satz 4.1.2. *Sei $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ Lipschitz-stetig mit Lipschitz-Konstante $L \geq 0$. Dann ist f stetig (ergänzbar) an jeder Stelle $x_0 \in \overline{\Omega}$.*

Beweis. Sei $x_0 \in \Omega$ und sei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ eine Folge mit $x_k \to x_0$ $(k \to \infty)$. Mit (4.1.1) folgt

$$\|f(x_k) - f(x_0)\| \leq L \ \|x_k - x_0\| \to 0 \ \ (k \to \infty).$$

Also ist f an der Stelle $x_0 \in \Omega$ stetig.

Falls $x_0 \in \overline{\Omega}\backslash\Omega$, $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$, so folgt mit (4.1.1) analog

$$\|f(x_k) - f(x_l)\| \leq L \ \|x_k - x_l\| \to 0 \ \ (k, l \to \infty);$$

d.h. $(f(x_k))_{k\in\mathbb{N}}$ ist Cauchy-Folge, und gemäss Satz 3.6.2 existiert $a = \lim\limits_{k\to\infty} f(x_k)$. Der Limes ist unabhängig von der Folge (x_k). Für jede weitere Folge $(y_k)_{k\in\mathbb{N}}$ mit $y_k \to x_0$ $(k \to \infty)$ existiert ebenfalls der Limes $b = \lim\limits_{k\to\infty} f(y_k)$, und

$$\|a - b\| = \lim_{k\to\infty} \|f(x_k) - f(y_k)\| \leq \lim_{k\to\infty} \big(L\,\|x_k - y_k\|\big) = 0.$$

Also ist f an der Stelle $x_0 \in \overline{\Omega}\backslash\Omega$ stetig ergänzbar. □

Beispiel 4.1.5. Das Skalarprodukt $\mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto x \cdot y \in \mathbb{R}$ ist auf jeder Kugel $B_R(0) \subset \mathbb{R}^n$ Lipschitz-stetig mit Lipschitz-Konstante $L = \sqrt{2}R$.

Beweis. Für $x, y, x_0, y_0 \in B_R(0)$ schätze ab mit Satz 2.4.1

$$\begin{aligned} |x \cdot y - x_0 \cdot y_0| &\leq |(x - x_0) \cdot y| + |x_0 \cdot (y - y_0)| \\ &\leq \|x - x_0\| \, \|y\| + \|x_0\| \, \|y - y_0\| \leq R\big(\|x - x_0\| + \|y - y_0\|\big) \\ &\leq \sqrt{2}R \, \|(x, y) - (x_0, y_0)\| \,, \end{aligned}$$

analog zu Beispiel 4.1.4.iii). □

Beispiel 4.1.5 motiviert die folgende Definition.

Definition 4.1.5. *$f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ heisst* **lokal Lipschitz-stetig**, *falls zu jedem $x_0 \in \Omega$ eine Umgebung $U = B_r(x_0) \cap \Omega$ existiert, so dass die auf U eingeschränkte Funktion*

$$f|_U \colon U \ni x \mapsto f(x) \in \mathbb{R}^n$$

auf U Lipschitz-stetig ist.

Analog zu Satz 4.1.2 gilt

Satz 4.1.3. *Sei $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ (lokal) Lipschitz-stetig. Dann ist f stetig an jeder Stelle $x_0 \in \Omega$.*

Beweis. Zu $x_0 \in \Omega$ wähle eine Umgebung $U = B_r(x_0) \cap \Omega$ von x_0 mit $f|_U$ Lipschitz-stetig. Es sei L eine Lipschitz-Konstante für $f|_U$. Falls $(x_k)_{k\in\mathbb{N}}$ Folge in Ω ist mit $x_k \to x_0$ $(k \to \infty)$, so gilt für $k \geq k_0(U)$ auch $\|x_k - x_0\| < r$, $x_k \in U$; also

$$\|f(x_k) - f(x_0)\| \leq L \ \|x_k - x_0\| \to 0, \quad (k \geq k_0, \ k \to \infty).$$

□

Bemerkung 4.1.3. Später werden wir sehen, dass Funktionen $f \in C^1(\Omega)$ auf einer offenen Menge $\Omega \subset \mathbb{R}^d$ lokal Lipschitz stetig sind. Vergleiche Beispiel 7.2.2.iii).

4.2 Stetige Funktionen

Sei $\Omega \subset \mathbb{R}^d$, $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$.

Definition 4.2.1. *f heisst* **stetig auf Ω**, *falls f in jedem Punkt $x_0 \in \Omega$ stetig ist.*

Beispiel 4.2.1. i) Polynome sind stetige Funktionen auf $\mathbb{C}$.

ii) Rationale Funktionen p/q sind stetig auf $\Omega = \{z \in \mathbb{C};\ q(z) \neq 0\}$.

iii) Ist $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ lokal Lipschitz-stetig, so ist f stetig auf Ω.

Satz 4.2.1. *Sind $f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ und $g\colon \mathbb{R}^n \to \mathbb{R}^l$ stetig, so ist auch deren Komposition $g \circ f\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^l$ stetig.*

Beweis. Sei $x_0 \in \Omega$, und sei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$. Dann gilt zunächst

$$y_k := f(x_k) \to f(x_0) =: y_0 \quad (k \to \infty).$$

Da g insbesondere stetig an der Stelle y_0, folgt weiter

$$g(y_k) = g(f(x_k)) = (g \circ f)(x_k) \overset{(k\to\infty)}{\to} g(y_0) = (g \circ f)(x_0);$$

$g \circ f$ ist also stetig an der Stelle x_0. □

Mit Beispiel 4.1.4.iii) folgt

Satz 4.2.2. *Sind $f, g\colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ stetig, so sind auch die Funktionen $f+g$ und αf stetig, wobei $\alpha \in \mathbb{R}$ beliebig. Die stetigen Funktionen $f\colon \Omega \to \mathbb{R}^n$ bilden also einen $\mathbb{R}$-Vektorraum.*

Notation: $C^0(\Omega;\mathbb{R}^n) = \{f\colon \Omega \to \mathbb{R}^n;\ f \text{ ist stetig}\}$.

Unter geeigneten Annahmen an Ω kann man den Raum $C^0(\Omega;\mathbb{R}^n)$ mit einer Norm versehen mit den in Satz 2.4.2 genannten Eigenschaften; vergleiche Abschnitt 4.7.

Definition 4.2.2. *$K \subset \mathbb{R}^d$ heisst* **kompakt**, *falls jede Folge $(x_k)_{k\in\mathbb{N}} \subset K$ einen Häufungspunkt in K besitzt; d.h. falls eine Teilfolge $\Lambda \subset \mathbb{N}$ und ein $x_0 \in K$ existieren mit*

$$x_k \to x_0 \quad (k \to \infty,\ k \in \Lambda).$$

Beispiel 4.2.2. i) Das "abgeschlossene" Intervall $[0,1]$ ist kompakt.

Beweis. Jede Folge $(x_k)_{k\in\mathbb{N}}$ hat nach Satz 3.4.1 (Bolzano-Weierstrass) eine konvergente Teilfolge $x_k \to x_0$ $(k \to \infty,\ k \in \Lambda)$, und mit $0 \le x_k \le 1$ $(k \in \mathbb{N})$ folgt $0 \le x_0 \le 1$ mit Satz 3.3.2. □

ii) Das "offene" Intervall $]0,1[$ ist **nicht** kompakt.

Beweis. Die Folge $x_k = \frac{1}{k}$, $k \in \mathbb{N}$, mit $x_k \to 0$ $(k \to \infty)$ kann keinen weiteren Häufungspunkt $x_0 \in]0,1[$ haben. □

iii) $\mathbb{R}$ ist nicht kompakt. (Betrachte $x_k = k$, $k \in \mathbb{N}$.)

iv) Allgemein gilt, dass eine kompakte Menge $K \subset \mathbb{R}^d$ beschränkt sein muss. (Sonst gibt es zu jedem $k \in \mathbb{N}$ einen Punkt $x_k \in K$ mit $\|x_k\| \ge k$. Die Folge $(x_k)_{k\in\mathbb{N}}$ kann keinen Häufungspunkt haben.) Später werden wir sehen, dass für beschränkte $\Omega \subset \mathbb{R}^d$ die Menge $\overline{\Omega}$ stets kompakt ist; vergleiche Bemerkung 4.3.4.

v) $S^{d-1} := \{x \in \mathbb{R}^d;\ \|x\| = 1\}$ ist kompakt.

Beweis. Sei $(x_k)_{k\in\mathbb{N}} \subset S^{d-1}$. Nach Satz 3.6.3 gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, $x_0 \in \mathbb{R}^d$ mit $x_k \to x_0$ $(k \to \infty,\ k \in \Lambda)$. Da $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ gemäss Beispiel 4.1.4.i) stetig ist, folgt $1 = \|x_k\| \to \|x_0\|$ $(k \to \infty,\ k \in \Lambda)$, und $x_0 \in S^{d-1}$. □

Lemma 4.2.1. *Sei $K \subset \mathbb{R}$ kompakt. Dann ist K beschränkt und es gibt $a, b \in K$ mit*

$$a = \inf K = \min K, \quad b = \sup K = \max K.$$

Beweis. Wähle $a_k \in K$ mit $a_k \to \inf K \geq -\infty$ $(k \to \infty)$. Da K kompakt, gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, $a \in K$ mit $a_k \to a$ $(k \to \infty,\ k \in \Lambda)$, und

$$a = \lim_{k \to \infty,\ k \in \Lambda} a_k = \inf K = \min K.$$

Insbesondere ist $\inf K > -\infty$. Analog für b. □

Satz 4.2.3. *Sei $K \subset \mathbb{R}^d$ kompakt, $f\colon K \to \mathbb{R}^n$ stetig. Dann ist $f(K)$ kompakt. Insbesondere nehmen stetige Funktionen $f\colon K \to \mathbb{R}$ ihr Supremum und Infimum an.*

Beweis. Sei $y_k = f(x_k)$, $k \in \mathbb{N}$, Folge in $f(K)$. Da $K \subset \mathbb{R}^d$ nach Annahme kompakt ist, gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, $x_0 \in K$ mit

$$x_k \to x_0 \quad (k \to \infty,\ k \in \Lambda).$$

Da f stetig ist, folgt

$$y_k = f(x_k) \to f(x_0) \quad (k \to \infty,\ k \in \Lambda);$$

also ist $f(K)$ kompakt.

Der 2. Teil der Aussage folgt nun unmittelbar aus Lemma 4.2.1: Falls $f(K) \subset \mathbb{R}$ kompakt, gibt es $a = f(x_0) \in f(K)$, $b = f(x_1) \in f(K)$ mit

$$a = \min f(K) = \inf_{x \in K} f(x) = f(x_0), \quad b = \max f(K) = \sup_{x \in K} f(x) = f(x_1).$$

□

Beispiel 4.2.3. i) Sei $f\colon [a, b] \to \mathbb{R}$ stetig. Dann ist f beschränkt, und es existiert $x_0 \in [a, b]$ mit

$$f(x_0) = \max_{a \leq x \leq b} f(x).$$

ii) Analog existiert für kompaktes $K \subset \mathbb{R}^d$ und stetiges $f\colon K \to \mathbb{R}^n$ ein $x_0 \in K$ mit

$$\|f(x_0)\| = \max_{x \in K} \|f(x)\|.$$

Beweis. Die Funktion $F = \|\cdot\| \circ f\colon K \to \mathbb{R}$ ist stetig wegen Satz 4.2.1 und Beispiel 4.1.4.i). □

4.3 Ein wenig Topologie

Im vorangegangenen Abschnitt haben wir gesehen, dass die "Gestalt" einer Menge $\Omega \subset \mathbb{R}^d$ in interessanter Weise zusammenwirkt mit dem Begriff der Stetigkeit

von Funktionen $f\colon \Omega \to \mathbb{R}^n$. Was ist jedoch der "Abschluss" einer Menge Ω? Wann ist ein allgemeines Ω "kompakt"? - Auskunft auf diese Fragen gibt das Teilgebiet **Topologie** der Mathematik, mit dem wir uns nun beschäftigen.

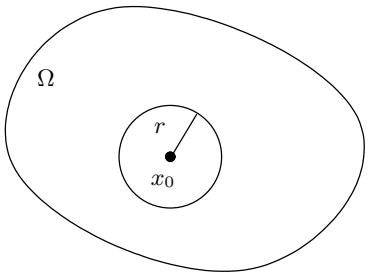
Definition 4.3.1. i) *Sei $x_0 \in \mathbb{R}^d$. Der* **offene Ball** *vom Radius $r > 0$ um x_0 ist die Menge*

$$B_r(x_0) = \{x \in \mathbb{R}^d;\ |x - x_0| < r\}.$$

ii) *$x_0 \in \Omega$ heisst* **innerer Punkt** *von Ω, falls*

$$\exists r > 0 :\ B_r(x_0) \subset \Omega.$$

iii) *$\Omega \subset \mathbb{R}^d$ heisst* **offen**, *falls jedes $x_0 \in \Omega$ ein innerer Punkt ist.*



Beispiel 4.3.1. i) Jeder offene Ball $B_R(a) \subset \mathbb{R}^d$ ist offen.

Beweis. Sei $x_0 \in B_R(a)$. Setze $r = R - |x_0 - a| > 0$. Dann gilt für $x \in B_r(x_0)$

$$|x - a| \le \underbrace{|x - x_0|}_{<r} + |x_0 - a| < R.$$

Also gilt $B_r(x_0) \subset B_R(a)$, und $B_R(a)$ ist offen. □

ii) Mit i) ist jedes offene Intervall $]a, b[= B_r(x; \mathbb{R})$ offen, wo $x = \frac{a+b}{2}$ und $r = \frac{b-a}{2} > 0$.

iii) Das halboffene Intervall

$$[a, b[= \{x \in \mathbb{R};\ a \le x < b\}$$

ist nicht offen, da a kein innerer Punkt ist.

Die folgenden Eigenschaften charakterisieren die durch die offenen Mengen erklärte "Topologie" auf $\mathbb{R}^d$.

Satz 4.3.1. *Es gilt*

i) *$\emptyset,\ \mathbb{R}^d$ sind offen,*

ii) *$\Omega_1,\ \Omega_2 \subset \mathbb{R}^d$ offen $\Rightarrow \Omega_1 \cap \Omega_2$ offen,*

iii) *$\Omega_\iota \subset \mathbb{R}^d$ offen, $\iota \in I \Rightarrow \bigcup_{\iota \in I} \Omega_\iota$ offen.*

Beweis. **i)** klar.

ii) Sei $x \in \Omega_1 \cap \Omega_2 \subset \Omega_i$, $i = 1, 2$. Da Ω_i offen, existiert $r_i > 0$ mit

$$B_{r_i}(x) \subset \Omega_i, \quad i = 1, 2.$$

Für $r = \min\{r_1,\ r_2\} > 0$ gilt dann

$$B_r(x) \subset B_{r_i}(x) \subset \Omega_i, \quad i = 1, 2,$$

also $B_r(x) \subset \Omega_1 \cap \Omega_2$.

iii) Sei $x \in \bigcup\limits_{\iota \in I} \Omega_\iota$. Dann gibt es $\iota_0 \in I$ mit $x \in \Omega_{\iota_0}$. Da Ω_{ι_0} offen, gibt es $r > 0$ mit

$$B_r(x) \subset \Omega_{\iota_0} \subset \bigcup_{\iota \in I} \Omega_\iota.$$

□

Bemerkung 4.3.1. Im allgemeinen ist der Durchschnitt unendlich vieler offener Mengen **nicht** offen; zum Beispiel gilt

$$\bigcap_{k=1}^{\infty} B_{1/k}(0) = \{0\}.$$

Definition 4.3.2. *$A \subset \mathbb{R}^d$ heisst* **abgeschlossen**, *falls $\mathbb{R}^d \backslash A$ offen ist.*

Beispiel 4.3.2. $[a, b]$ ist abgeschlossen, da $]-\infty, a[\cup]b, \infty[$ nach Beispiel 4.3.1 und Satz 4.3.1.iii) offen ist.

Satz 4.3.2. *Es gilt*

i) *$\emptyset$, $\mathbb{R}^d$ sind abgeschlossen;*

ii) *A_1, A_2 abgeschlossen $\Rightarrow$ $A_1 \cup A_2$ abgeschlossen,*

iii) *A_ι abgeschlossen für $\iota \in I$ $\Rightarrow$ $\bigcap\limits_{\iota \in I} A_\iota$ abgeschlossen.*

Beweis. Direkt aus Satz 4.3.1 mit den de Morganschen Regeln, Übung 1.4. Zum Beispiel sieht man ii), wie folgt.

Seien $A_{1,2}$ abgeschlossen, $\Omega_i = \mathbb{R}^d \backslash A_i$ also offen, $i = 1, 2$. Mit Satz 4.3.1.ii) folgt

$$(A_1 \cup A_2)^c = (A_1^c \cap A_2^c) = \Omega_1 \cap \Omega_2$$

ist offen; $A_1 \cup A_2$ ist also abgeschlossen. □

Definition 4.3.3. **i)** *Die Menge der inneren Punkte von Ω,*

$$int(\Omega) = \bigcup_{U \subset \Omega,\, U \text{ offen}} U =: \Omega^\circ$$

heisst **offener Kern** *oder das* **Innere** *(engl.: interior) von Ω.*

ii) *Die Menge*

$$clos(\Omega) = \bigcap_{A \supset \Omega,\, A \ abgeschlossen} A$$

heisst **Abschluss** *(engl.: closure) von* Ω.

iii) *Die Menge*

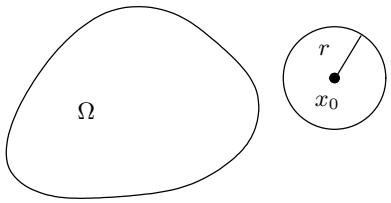
$$\partial\Omega = clos(\Omega) \backslash int(\Omega)$$

heisst **Rand** *von* Ω.

Bemerkung 4.3.2. i) Nach Satz 4.3.1.iii) ist $\Omega^\circ \subset \Omega$ offen. Offenbar ist $U = \Omega^\circ$ die **grösste** offene Menge $U \subset \Omega$, und Ω offen genau dann, wenn $\Omega = \Omega^\circ$.

ii) Analog ist nach Satz 4.3.2.iii) die Menge $A = clos(\Omega) \supset \Omega$ die **kleinste** abgeschlossene Menge $A \supset \Omega$, und Ω ist abgeschlossen genau dann, wenn $\Omega = clos(\Omega)$.

iii) Falls $B_r(x_0) \cap \Omega = \emptyset$ für ein $r > 0$, so folgt $x_0 \notin clos(\Omega)$. (Wähle $A = \mathbb{R}^n \backslash B_r(x_0) \supset \Omega$.) Da $clos(\Omega)$ abgeschlossen, $\mathbb{R}^d \backslash clos(\Omega)$ offen, gilt auch die Umkehrung.



Wir können nun den im vorigen Abschnitt eingeführten Abschluss einer Menge Ω identifizieren.

Satz 4.3.3. *Für* $\Omega \subset \mathbb{R}^d$ *gilt*

$$clos(\Omega) = \overline{\Omega} = \{x_0 \in \mathbb{R}^d;\ \exists (x_k)_{k\in\mathbb{N}} \subset \Omega,\ x_k \overset{(k\to\infty)}{\to} x_0\}.$$

Beweis. "$\subset$": Sei $x_0 \in clos(\Omega)$. Nach Bemerkung 4.3.2.iii) gibt es zu jedem $r > 0$ ein $x \in B_r(x) \cap \Omega$. Zu $r_k = \frac{1}{k}$ wähle $x_k \in B_{r_k}(x_0) \cap \Omega$, $k \in \mathbb{N}$. Offenbar gilt

$$\|x_k - x_0\| < r_k = \frac{1}{k} \to 0 \quad (k \to \infty);$$

also $x_0 \in \overline{\Omega}$.

"$\supset$": Sei $x_0 = \lim_{k\to\infty} x_k$ für eine Folge $(x_k)_{k\in\mathbb{N}} \subset \Omega$. Nimm widerspruchsweise an, dass $x_0 \notin clos(\Omega)$. Nach Bemerkung 4.3.2.iii) gibt es $r > 0$ mit

$$B_r(x_0) \cap \Omega = \emptyset.$$

Mit $x_k \to x_0$ $(k \to \infty)$ folgt jedoch $x_k \in B_r(x_0)$ für $k \geq k_0$, und $x_k \in \Omega$. $\square$

Beispiel 4.3.3. Mit Satz 4.3.3 und Beispiel 4.1.2.iii) erhalten wir

$$clos(B_R(0)) = \overline{B_R(0)} = \{x \in \mathbb{R}^d;\ \|x\| \leq R\}.$$

Was kann man aussagen über den Rand einer Menge?

Bemerkung 4.3.3. i) $\partial\Omega = \overline{\Omega}\backslash\Omega^\circ = \overline{\Omega} \cap (\mathbb{R}^n\backslash\Omega^\circ)$ ist abgeschlossen.

ii) Mit $\Omega^\circ \subset \Omega \subset \overline{\Omega}$ folgt $\overline{\Omega} = \Omega^\circ \cup \partial\Omega$, und die Zerlegung ist disjunkt.

iii) Somit erhalten wir das Kriterium

$$\Omega \subset \mathbb{R}^d \text{ abgeschlossen} \Leftrightarrow \Omega = \overline{\Omega} = \Omega^\circ \cup \partial\Omega \overset{(\Omega^\circ \subset \Omega)}{\Leftrightarrow} \partial\Omega \subset \Omega.$$

Beispiel 4.3.4. i) $\partial B_R(0) = \overline{B_R(0)}\backslash B_R(0) = \{x;\ \|x\| = R\}$, da $B_R(0) = B_R(0)^\circ$ offen.

ii) Sei $\Omega = \mathbb{Q} \subset \mathbb{R}$. Da jedes Intervall $B_r(x) =]x-r, x+r[\subset \mathbb{R}$ sowohl rationale als auch irrationale Punkte (zum Beispiel der Form $\sqrt{2} + q$ für $q \in \mathbb{Q}$) enthält, gilt

$$\mathbb{Q}^\circ = \emptyset, \quad (\mathbb{R}\backslash\mathbb{Q})^\circ = \emptyset.$$

Somit ist $A = \mathbb{R}$ die einzige abgeschlossene Menge $A \supset \mathbb{Q}$ (denn $\mathbb{R}\backslash A$ muss in $(\mathbb{R}\backslash\mathbb{Q})^\circ = \emptyset$ enthalten sein), und $\overline{\mathbb{Q}} = \mathbb{R}$. Es folgt

$$\partial\mathbb{Q} = \overline{\mathbb{Q}}\backslash\mathbb{Q}^\circ = \mathbb{R}.$$

Der Rand einer Menge Ω kann also erstaunlich gross sein.

Satz 4.3.4. *Für $\Omega \subset \mathbb{R}^d$ gilt*

$$\partial\Omega = \{x \in \mathbb{R}^d;\ \forall r > 0 :\ B_r(x) \cap \Omega \neq \emptyset \neq B_r(x)\backslash\Omega\}.$$

Beweis. "$\subset$": Sei $x \in \partial\Omega$, und sei $r > 0$. Falls $B_r(x) \cap \Omega = \emptyset$, so folgt $x \notin \overline{\Omega}$ nach Bemerkung 4.3.2.iii), also $x \notin \partial\Omega \subset \overline{\Omega}$ im Widerspruch zur Annahme. Falls $B_r(x) \subset \Omega$, so folgt $x \in \Omega^\circ$, also $x \notin \partial\Omega = \overline{\Omega}\backslash\Omega^\circ$, was ebenfalls der Annahme widerspricht.

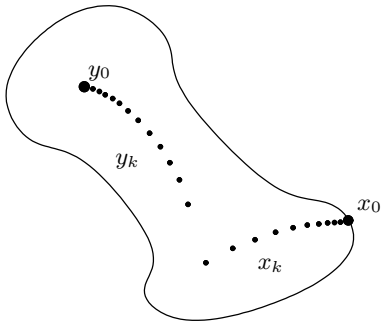
"$\supset$": Sei $x \notin \partial\Omega$, also entweder $x \notin \overline{\Omega}$ oder $x \in \Omega^\circ$. Falls $x \notin \overline{\Omega}$ gibt es nach Bemerkung 4.3.2.iii) ein $r > 0$ mit $B_r(x) \cap \Omega = \emptyset$. Falls $x \in \Omega^\circ$, gibt es $r > 0$ mit $B_r(x) \subset \Omega^\circ \subset \Omega$, also $B_r(x)\backslash\Omega = \emptyset$. □

Die Sätze 4.3.3 und 4.3.4 ergeben die folgene Charakterisierung abgeschlossener Mengen in $\mathbb{R}^d$ mittels Folgen.

Satz 4.3.5. (Folgenkriterium für Abgeschlossenheit) *Für $A \subset \mathbb{R}^d$ sind äquivalent:*

i) *A ist abgeschlossen,*

ii) $\forall (x_k)_{k\in\mathbb{N}} \subset A :\ x_k \to x_0\ (k \to \infty) \Rightarrow x_0 \in A.$



Beweis. *i) ⇒ ii):* Sei A abgeschlossen. Mit Satz 4.3.3 erhalten wir

$$A = clos(A) = \overline{A} = \{x_0;\ \exists (x_k)_{k\in\mathbb{N}} \subset A :\ x_k \to x_0\ (k \to \infty)\},$$

und ii) folgt.

ii) ⇒ i) Es gelte ii). Wegen Bemerkung 4.3.3 genügt es zu zeigen, dass $\partial A \subset A$. Sei $x_0 \in \partial A$. Mit Satz 4.3.4 erhalten wir für $r_k = \frac{1}{k}$ Punkte $x_k \in B_{r_k}(x_0) \cap A$ mit $x_k \to x_0$ $(k \to \infty)$, also $x_0 \in A$ gemäss ii). □

Satz 4.3.5 erlaubt nun eine einfache Charakterisierung kompakter Mengen in $\mathbb{R}^d$.

Satz 4.3.6. *Für $K \subset \mathbb{R}^d$ sind äquivalent*

i) *K ist (folgen)-kompakt im Sinne der Definition 4.2.2,*

ii) *K ist beschränkt und abgeschlossen.*

Beweis. *i) ⇒ ii):* Nach Beispiel 4.2.2.iv) ist jede kompakte Menge K beschränkt. Zum Beweis der Abgeschlossenheit von K Sei $(x_k)_{k\in\mathbb{N}} \subset K$ mit $x_k \to x_0$ $(k \to \infty)$. Dann ist x_0 der einzige Häufungspunkt von $(x_k)_{k\in\mathbb{N}}$. Nach Annahme ist K (folgen)-kompakt; also besitzt $(x_k)_{k\in\mathbb{N}}$ einen Häufungspunkt in K. Es folgt $x_0 \in K$, und nach Satz 4.3.4 ist K abgeschlossen.

ii) ⇒ i): Sei K beschränkt und abgeschlossen, und sei $(x_k)_{k\in\mathbb{N}}$ Folge in K. Nach Satz 3.6.3 gibt es $x_0 \in \mathbb{R}^d$ und eine Teilfolge $\Lambda \subset \mathbb{N}$ mit $x_k \to x_0$ $(k \to \infty,\ k \in \Lambda)$. Da K abgeschlossen, folgt mit Satz 4.3.5 $x_0 \in K$, wie gewünscht. □

Bemerkung 4.3.4. Für beschränktes Ω ist somit $\overline{\Omega}$ stets kompakt.

4.4 Äquivalente Normen

Die "Topologie" der offenen Mengen in $\mathbb{R}^d$ ist durch die offenen Bälle $B_r(x_0) \subset \mathbb{R}^d$ definiert. Neben der euklidischen Norm aus Satz 2.4.2 kann man jedoch auch andere Normen definieren, die ebenfalls die Eigenschaften in Satz 2.4.2.i)-iii) besitzen.

Definition 4.4.1. *Eine* **Norm** *auf* $\mathbb{R}^d$ *ist eine Abbildung* $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ *mit den Eigenschaften*

i) Definitheit: $\|x\| \geq 0,\ \|x\| = 0 \Leftrightarrow x = 0,$

ii) Positive Homogenität: $\|\alpha x\| = |\alpha|\,\|x\|,$

iii) Dreiecks-Ungleichung: $\|x + y\| \leq \|x\| + \|y\|$

für alle $x, y \in \mathbb{R}^d$, $\alpha \in \mathbb{R}$.

Beispiel 4.4.1. Für $1 \leq p < \infty$ definiert

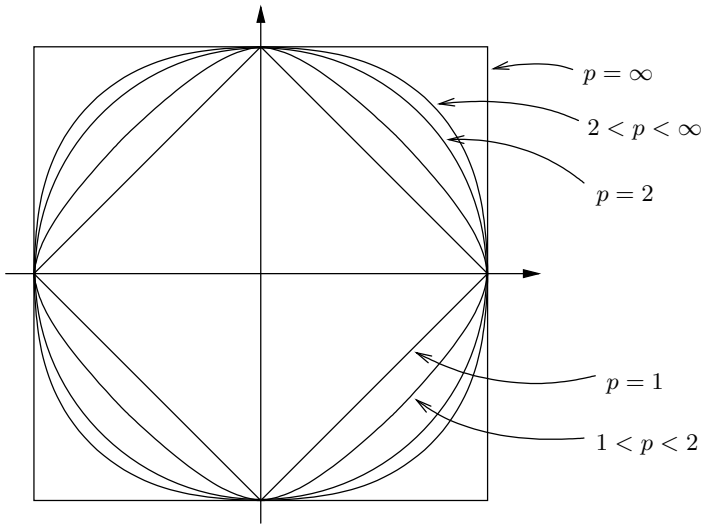
$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{d} |x_i|^p},\ x = (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

eine Norm auf $\mathbb{R}^d$, ebenso für $p = \infty$ der Ausdruck

$$\|x\|_\infty = \max_{1 \leq i \leq d} |x_i|\,,\ x = (x_1, \ldots, x_d) \in \mathbb{R}^d.$$

Offenbar gilt für alle $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $1 \leq p \leq \infty$:

$$\|x\|_\infty = \max_i |x_i| \leq \|x\|_p = \sqrt[p]{\sum_i |x_i|^p} \leq \sum_i |x_i| = \|x\|_1 \leq d\,\|x\|_\infty\,. \quad (4.4.1)$$



Definition 4.4.2. *Zwei Normen* $\|\cdot\|^{(1)}$, $\|\cdot\|^{(2)} : \mathbb{R}^d \to \mathbb{R}$ *heissen* **äquivalent**, *falls* $C > 0$ *existiert mit*

$$\frac{1}{C}\,\|x\|^{(1)} \leq \|x\|^{(2)} \leq C\ \|x\|^{(1)}\,,\quad \forall x \in \mathbb{R}^d. \quad (4.4.2)$$

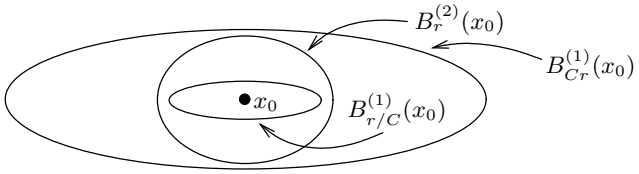
Beispiel 4.4.2. Je zwei der in Beispiel 4.4.1 definierten Normen $\|\cdot\|_p$, $\|\cdot\|_q$, $1 \leq p, q \leq \infty$ sind wegen (4.4.1) äquivalent.

Bemerkung 4.4.1. Äquivalente Normen $\|\cdot\|^{(1)}$, $\|\cdot\|^{(2)}$ definieren dieselben offenen Mengen.

Beweis. Für die Normkugeln $B_r^{(i)}(x_0) = \{x;\ \|x - x_0\|^{(i)} < r\}$ gilt mit (4.4.2)

$$B_{r/C}^{(1)}(x_0) \subset B_r^{(2)}(x_0) \subset B_{Cr}^{(1)}(x_0);$$

also ist $x_0 \in \Omega$ innerer Punkt von Ω bzgl. $\|\cdot\|^{(2)}$ genau dann, wenn $x_0 \in \Omega$ innerer Punkt ist bzgl. $\|\cdot\|^{(1)}$. □



Satz 4.4.1. *Je zwei Normen $\|\cdot\|^{(1)}$, $\|\cdot\|^{(2)} : \mathbb{R}^d \to \mathbb{R}$ sind äquivalent.*

Beweis. Es genügt, den Fall $\|\cdot\|^{(2)} = \|x\|_2$ zu betrachten, wobei $\|\cdot\|_2$ wie in Beispiel 4.4.1 die euklidische Norm bezeichnet. Der Kürze halber schreiben wir auch $\|\cdot\|$ anstelle von $\|\cdot\|^{(1)}$.

Behauptung 1 $\|\cdot\| : \mathbb{R}^d \to \mathbb{R}$ ist Lipschitz stetig bzgl. $\|\cdot\|_2$.

Beweis. Für $x, y \in \mathbb{R}^d$ mit $y - x = z = (z_1, \ldots, z_d) = \sum_{i=1}^{d} z_i e_i$ schätze ab

$$\begin{aligned}\|y - x\| = \|z\| &\leq \sum_{i=1}^{d} \|z_i e_i\| = \sum_{i=1}^{d} |z_i| \, \|e_i\| \\ &\leq C \sum_{i=1}^{d} |z_i| \leq Cd \max_i |z_i| \leq Cd \, \|z\|_2 = Cd \, \|y - x\|_2 .\end{aligned}$$

□

Nach Beispiel 4.2.3.iii) ist $S^{d-1} \subset \mathbb{R}^d$ kompakt. Gemäss Satz 4.2.3 gibt es x_{min}, $x_{max} \in S^{d-1}$ mit

$$\lambda := \|x_{min}\| = \inf_{x \in S^{d-1}} \|x\| \leq \sup_{x \in S^{d-1}} \|x\| = \|x_{max}\| =: \Lambda,$$

und $0 < \lambda \leq \Lambda$ wegen Definition 4.4.1. Wähle $C = \max\{\Lambda, \frac{1}{\lambda}\}$. Es folgt

$$\frac{1}{C} \leq \left\| \frac{x}{\|x\|_2} \right\| = \frac{\|x\|}{\|x\|_2} \leq C, \ \forall x \in \mathbb{R}^d \backslash \{0\}.$$

□

4.5 Topologisches Kriterium für Stetigkeit

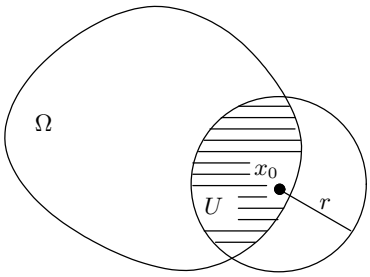
Sei $\Omega \subset \mathbb{R}^d,\ x_0 \in \Omega$.

Definition 4.5.1. i) *$U \subset \Omega$ heisst* **Umgebung** *von x_0 relativ zu Ω, falls $r > 0$ existiert mit*

$$B_r(x_0) \cap \Omega \subset U.$$

ii) *$U \subset \Omega$ heisst* **relativ offen**, *falls U Umgebung jedes Punktes $x_0 \in U$ ist; d.h. falls $U = E \cap \Omega$ für eine offene Menge $E \subset \mathbb{R}^d$.*

iii) *$A \subset \Omega$ heisst* **relativ abgeschlossen**, *falls $\Omega \backslash A$ relativ offen ist; d.h. falls $\Omega \backslash A = E \cap \Omega$ mit offenem $E \subset \mathbb{R}^d$ und daher $A = \Omega \backslash E = \Omega \cap F$ für ein abgeschlossenes $F = \mathbb{R}^d \backslash E \subset \mathbb{R}^d$.*



Beispiel 4.5.1. i) Sei $\Omega = [0,1[\subset \mathbb{R}$, und sei $0 < r < 1$. Dann ist $[0, r[$ relativ offen in Ω.

ii) Sei $\Omega =]0,1[\subset \mathbb{R}$, und sei $0 < r < 1$. Dann ist $]0, r]$ relativ abgeschlossen in Ω.

Satz 4.5.1. *Sei $f\colon \Omega \to \mathbb{R}^n$, $x_0 \in \Omega$. Es sind äquivalent*

i) *f ist stetig an der Stelle x_0 gemäss Definition 4.1.3 (Folgenkriterium);*

ii) $\forall \epsilon > 0\ \exists \delta > 0\ \forall x \in \Omega$*:*

$$\|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \epsilon$$

(Weierstrass'sches $\epsilon - \delta-$ Kriterium);

iii) *Für jede Umgebung V von $f(x_0)$ in $\mathbb{R}^n$ ist $U = f^{-1}(V)$ eine Umgebung von x_0 in Ω.*

Beweis. *i) $\Rightarrow$ ii)* (indirekt): Nimm an, für ein $\epsilon > 0$ und jedes $\delta > 0$ gibt es

$$x \in \Omega \cap B_\delta(x_0)\colon\ \|f(x) - f(x_0)\| \geq \epsilon.$$

Für $\delta_k = \frac{1}{k}$, $k \in \mathbb{N}$ erhalten wir eine Folge $(x_k)_{k \in \mathbb{N}} \subset \Omega$ mit

$$x_k \to x_0\ (k \to \infty),\ f(x_k) \not\to f(x_0)\ \ (k \to \infty)$$

im Widerspruch zu i).

ii) ⇒ iii): Sei V Umgebung von $f(x_0)$ in $\mathbb{R}^n$, und seien $\epsilon > 0$ mit $B_\epsilon(f(x_0)) \subset V$ und dazu $\delta > 0$ gemäss ii) gewählt. Mit ii) folgt $f(x) \in B_\epsilon(f(x_0)))$ für alle $x \in B_\delta(x_0) \cap \Omega$, also

$$B_\delta(x_0) \cap \Omega \subset f^{-1}(B_\epsilon(f(x_0))) \subset f^{-1}(V).$$

iii) ⇒ ii): Zu $V = B_\epsilon(f(x_0))$ gibt es nach iii) ein $\delta > 0$ mit $B_\delta(x_0) \cap \Omega \subset f^{-1}(V)$; d.h. ii).

ii) ⇒ i): Sei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$. Zu $\epsilon > 0$ wähle $\delta > 0$ gemäss ii), dazu k_0 mit

$$\|x_k - x_0\| < \delta, \quad \forall k \geq k_0.$$

Mit ii) folgt

$$\|f(x_k) - f(x_0)\| < \epsilon, \quad \forall k \geq k_0;$$

d.h. $f(x_k) \to f(x_0)$ $(k \to \infty)$. □

Beispiel 4.5.2. Mit Satz 4.5.1 erkennt man nun sofort, dass $f = \chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}$$

an keiner Stelle $x_0 \in \mathbb{Q}$ stetig ist, denn

$$f^{-1}(\,]\frac{1}{2}, \frac{3}{2}[\,) = \mathbb{Q}$$

ist wegen Beispiel 4.3.4.ii) keine Umgebung von x_0. Analog ist f in keiner Stelle $x_0 \notin \mathbb{Q}$ stetig, denn

$$f^{-1}(\,]-\frac{1}{2}, \frac{1}{2}[\,) = \mathbb{R}\backslash\mathbb{Q}$$

hat leeres Inneres; vergleiche Beispiel 4.1.3.v).

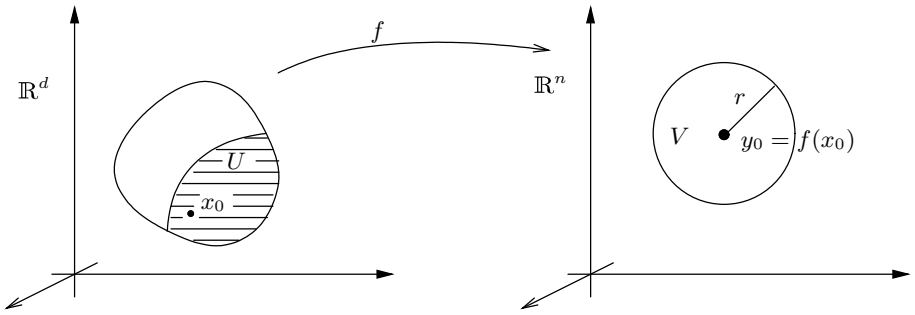
Aus Satz 4.5.1 ergibt sich nun das folgende, äusserst elegante topologische Kriterium für die Stetigkeit einer Funktion $f : \Omega \to \mathbb{R}^n$.

Satz 4.5.2. *Für $f : \Omega \to \mathbb{R}^n$ sind äquivalent*

i) *f ist stetig (in allen Punkten $x_0 \in \Omega$);*

ii) *Das Urbild $U = f^{-1}(V)$ jeder offenen Menge $V \subset \mathbb{R}^n$ ist relativ offen;*

iii) *Das Urbild $A = f^{-1}(B)$ jeder abgeschlossenen Menge $B \subset \mathbb{R}^n$ ist relativ abgeschlossen.*



Beweis. *i) ⇒ ii)*: Sei $V \subset \mathbb{R}^n$ offen, $x_0 \in U = f^{-1}(V)$. Da f stetig an der Stelle x_0, ist nach Satz 4.5.1 der Punkt x_0 innerer Punkt von U relativ zu Ω. Da $x_0 \in U$ beliebig, ist U somit relativ offen.

ii) ⇒ iii): Sei $F \subset \mathbb{R}^n$ abgeschlossen. Dann ist $V = \mathbb{R}^n \backslash F$ offen, mit ii) also $U = f^{-1}(V)$ relativ offen, und mit Definition 4.5.1.iii) folgt

$$A = f^{-1}(F) = f^{-1}(\mathbb{R}^n \backslash V) = \Omega \backslash f^{-1}(V) = \Omega \backslash U$$

ist relativ abgeschlossen.

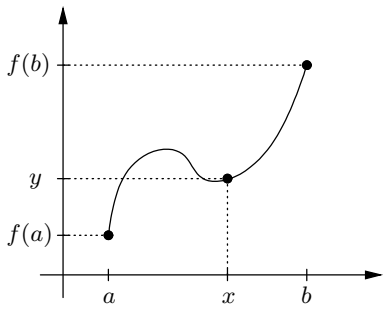
iii) ⇒ ii): analog.

ii) ⇒ i): Da jede Umgebung W eines Punktes $y_0 = f(x_0)$ eine offene Umgebung V von y_0 enthält, ist $f^{-1}(W) \supset f^{-1}(V)$ nach ii) Umgebung von x_0. Nach Satz 4.5.1 ist f also stetig an der Stelle x_0. □

4.6 Zwischenwertsatz und Folgerungen

Die Stetigkeit einer Funktion hat noch weitere Konsequenzen. Wir betrachten zunächst reelle Funktionen auf Intervallen.

Satz 4.6.1. *Seien $-\infty < a < b < \infty$, und sei $f\colon [a,b] \to \mathbb{R}$ stetig, $f(a) \leq f(b)$. Dann gibt es zu jedem $y \in [f(a), f(b)]$ ein $x \in [a,b]$ mit $f(x) = y$.*



Beweis. Wir benutzen das "Bisektionsverfahren". Definiere $a_1 = a$, $b_1 = b$.

Setze

$$a_2 = a = a_1,\ b_2 = \frac{a+b}{2},\ \text{falls } f(\frac{a+b}{2}) \geq y,$$

bzw. setze

$$a_2 = \frac{a+b}{2},\ b_2 = b = b_1,\ \text{falls } f(\frac{a+b}{2}) < y,$$

so dass $f(a_2) < y \leq f(b_2)$ und $|a_2 - b_2| = 2^{-1}|a-b|$. Allgemein seien $a_1, \ldots, a_k$ sowie $b_1, \ldots, b_k$ bereits definiert mit

$$a_1 \leq \cdots \leq a_k \leq b_k \leq \cdots \leq b_1$$

und

$$f(a_k) < y \leq f(b_k),\ |a_k - b_k| = 2^{1-k}|a-b|.$$

Sei $c = \frac{a_k+b_k}{2}$. Falls gilt $f(c) \geq y$, setzen wir

$$a_{k+1} = a_k,\quad b_{k+1} = c;$$

falls $f(c) < y$, setzen wir

$$a_{k+1} = c,\quad b_{k+1} = b_k.$$

Wir erhalten in jedem Fall $a_{k+1} \geq a_k$, $b_{k+1} \leq b_k$ mit $f(a_{k+1}) < y \leq f(b_{k+1})$, und

$$|a_{k+1} - b_{k+1}| = \frac{1}{2}|a_k - b_k| = 2^{-k}|a-b|.$$

Die Folgen $(a_k)_{k\in\mathbb{N}}$, $(b_k)_{k\in\mathbb{N}}$ sind monoton und beschränkt. Nach Satz 3.3.1 gibt es

$$\overline{a} = \lim_{k\to\infty} a_k \leq \overline{b} = \lim_{k\to\infty} b_k,$$

und mit Satz 3.3.2 folgt

$$\left|\overline{a} - \overline{b}\right| = \lim_{k\to\infty} |a_k - b_k| = 0.$$

D.h. $\overline{a} = \overline{b} =: x \in [a,b]$. Da f stetig, folgt mit Satz 3.3.2 schliesslich

$$y \leq \lim_{k\to\infty} f(b_k) = f(x) = \lim_{k\to\infty} f(a_k) \leq y;$$

also $f(x) = y$, wie gewünscht. □

Beispiel 4.6.1. i) Sei $p\colon \mathbb{R} \to \mathbb{R}$ ein Polynom von ungeradem Grad. Dann hat p eine Nullstelle.

Beweis. Beachte $|p(x)| \to \infty$ für $|x| \to \infty$. OBdA $p(x) \to \infty$ für $x \to \infty$. (Sonst betrachte $\widetilde{p} = -p$.) Da $\deg(p)$ ungerade, folgt $p(x) \to -\infty$ für $x \to -\infty$, und die Behauptung folgt mit Satz 4.6.1. □

ii) Jede 3×3 Matrix A mit Koeffizienten in $\mathbb{R}$ hat mindestens einen Eigenwert $\lambda \in \mathbb{R}$.

Beweis. Das charakteristische Polynom p von A hat Grad 3; die Nullstellen λ von p sind genau die Eigenwerte von A. (Siehe: "Lineare Algebra"). □

Definition 4.6.1. $f\colon [a,b] \to \mathbb{R}$ *heisst* **streng monoton wachsend**, *falls gilt*

$$a \le x < y \le b \Rightarrow f(x) < f(y). \tag{4.6.1}$$

Satz 4.6.2. *Sei* $f\colon [a,b] \to \mathbb{R}$ *stetig und streng monoton wachsend. Setze* $f(a) = c,\ f(b) = d$. *Dann ist* $f\colon [a,b] \to [c,d]$ *bijektiv, und* f^{-1} *ist stetig.*

Beweis. f ist injektiv mit $im(f) \subset [c,d]$ wegen (4.6.1) und surjektiv gemäss Satz 4.6.1, also bijektiv.

Zum Nachweis der Stetigkeit der Umkehrfunktion sei $y_k = f(x_k)$, $k \in \mathbb{N}$, mit $y_k \to y_0 \in [c,d]$. Es gilt $y_0 = f(x_0)$ für ein $x_0 \in [a,b]$. Die Folge $(x_k)_{k\in\mathbb{N}}$ hat einen Häufungspunkt x in der kompakten Menge $[a,b]$, und

$$f(x) = \lim_{k\to\infty} f(x_k) = y_0 = f(x_0).$$

Mit (4.6.1) folgt $x = x_0$. Analog hat auch jede Teilfolge von $(x_k)_{k\in\mathbb{N}}$ denselben Häufungspunkt x_0; insbesondere gilt

$$\limsup_{k\to\infty} x_k = \liminf_{k\to\infty} x_k = x_0,$$

also

$$f^{-1}(y_k) = x_k \to x_0 = f^{-1}(y_0) \quad (k \to \infty).$$

□

Ein analoger Satz gilt auf offenen Intervallen.

Satz 4.6.3. *Sei* $f\colon]a,b[\to \mathbb{R}$ *stetig und streng monoton wachsend mit monotonen Limites*

$$-\infty \le c := \lim_{x\downarrow a} f(x) < \lim_{x\uparrow b} f(x) =: d \le \infty.$$

Dann ist $f\colon]a,b[\to]c,d[$ *bijektiv, und* f^{-1} *ist stetig.*

Beweis. Dass $f\colon]a,b[\to]c,d[$ bijektiv, folgt wie in Satz 4.6.2.

Für $y_k = f(x_k) \in]c,d[$ mit $y_k \to y_0 \in]c,d[$ $(k \to \infty)$ gilt

$$c < \overline{c} := \inf_k y_k \le \sup_k y_k =: \overline{d} < d.$$

Seien

$$a < \overline{a} \le \overline{b} < b$$

mit

$$f(\overline{a}) = \overline{c}, \quad f(\overline{b}) = \overline{d}.$$

Auf die auf das abgeschlossene Intervall $[\overline{a},\overline{b}]$ eingeschränkte Funktion $\overline{f} = f|_{[\overline{a},\overline{b}]}\colon [\overline{a},\overline{b}] \to [\overline{c},\overline{d}]$ ist Satz 4.6.2 anwendbar. Es folgt

$$x_k = (\overline{f})^{-1}(y_k) = f^{-1}(y_k) \overset{(k\to\infty)}{\to} x_0 = (\overline{f})^{-1}(y_0) = f^{-1}(y_0),$$

also ist f^{-1} stetig. □

Beispiel 4.6.2. i) Sei $n \in \mathbb{N}$. Die Potenzfunktion $\mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ ist stetig gemäss Beispiel 4.1.3.i). Weiter gilt für $0 < x, y < \infty$:

$$y^n - x^n = (y-x)\underbrace{(y^{n-1} + y^{n-2}x + \cdots + x^{n-1})}_{>0},$$

also folgt

$$x < y \Rightarrow x^n < y^n;$$

d.h. die Potenzfunktion $x \mapsto x^n$ ist streng monoton wachsend auf $\mathbb{R}_+ =]0, \infty[$.

Mit Satz 4.6.3 folgt, dass die Wurzelfunktion $\mathbb{R}_+ \ni y \mapsto \sqrt[n]{y} \in \mathbb{R}_+$ stetig ist.

ii) Betrachte die Funktion $Exp\colon \mathbb{R} \to \mathbb{R}$, $\mathbb{R} \ni x \mapsto Exp(x) \in \mathbb{R}$.

Behauptung $Exp > 0$, Exp ist stetig und streng monoton wachsend mit $Exp(\mathbb{R}) =]0, \infty[$.

Beweis. Mit dem Additionstheorem gemäss Korollar 3.8.1 folgt zunächst

$$Exp(x) = \big(Exp(\frac{x}{2})\big)^2 \geq 0,$$

wegen

$$Exp(x) = 1/Exp(-x) \neq 0 \tag{4.6.2}$$

also

$$Exp(x) > 0, \ \forall x \in \mathbb{R}.$$

Weiter gilt für $|h| < 1$

$$|Exp(h) - 1| = \left|\sum_{k=1}^{\infty} \frac{h^k}{k!}\right| \leq \sum_{k=1}^{\infty} |h|^k = \frac{|h|}{1-|h|} \to 0 \quad (h \to 0),$$

also für $x = x_0 + h \to x_0$ mit Korollar 3.8.1

$$Exp(x) - Exp(x_0) = Exp(x_0)(Exp(h) - 1) \to 0, \tag{4.6.3}$$

und die Funktion Exp ist stetig. Da

$$Exp(h) - 1 = \sum_{k=1}^{\infty} \frac{h^k}{k!} > 0 \text{ für } h > 0,$$

ergibt (4.6.3) zudem die gewünschte Monotonie

$$Exp(x_0) < Exp(x) \text{ für } x_0 < x = x_0 + h.$$

Schliesslich gilt offenbar $Exp(x) \to \infty$ $(x \to \infty)$ und mit (4.6.2) also auch $Exp(x) \to 0$ $(x \to -\infty)$. □

Gemäss Satz 4.6.3 besitzt $Exp\colon \mathbb{R} \to]0, \infty[$ die stetige Umkehrfunktion

$$Log = \big(Exp|_{\mathbb{R}}\big)^{-1} :]0, \infty[\to \mathbb{R}.$$

Wegen der Identität

$$Exp(x) = e^x, \ \forall x \in \mathbb{Q}$$

gemäss Satz 3.9.2 stimmt *Log* überein mit dem **natürlichen Logarithmus** $Log = \log$. (Die vielfach gebräuchliche Notation ln verwenden wir nicht.) Mit

$$Exp(Log(x) + Log(y)) = Exp(Log(x)) \cdot Exp(Log(y)) = xy$$

folgt zudem das **Additionstheorem**

$$Log(xy) = Log(x) + Log(y), \ \forall x, y > 0. \tag{4.6.4}$$

Bemerkung 4.6.1. Die Eigenschaft (4.6.4) ermöglicht das vereinfachte Multiplizieren mittels Rechenschieber oder Logarithmentafel (Jost Bürgi (1552-1632)).

Der Zwischenwertsatz hat auch topologische Konsequenzen.

Satz 4.6.4. *Sei $I = [0,1] = \Omega_1 \cup \Omega_2$, wobei $\Omega_{1,2} \subset I$ relativ offen mit $\Omega_1 \cap I \neq \emptyset \neq \Omega_2 \cap I$. Dann ist $\Omega_1 \cap \Omega_2 \neq \emptyset$.*

Beweis (indirekt). Sei $I = \Omega_1 \cup \Omega_2$, wobei $\emptyset \neq \Omega_{1,2} \subset I$ relativ offen mit $\Omega_1 \cap \Omega_2 = \emptyset$. Definiere $f\colon [0,1] \to \mathbb{R}$ mit

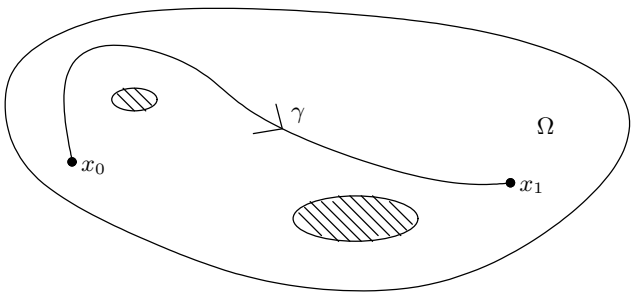
$$f(x) = \begin{cases} 1, & x \in \Omega_1 \\ 2, & x \in \Omega_2. \end{cases}$$

Dann ist f stetig gemaäss Satz 4.5.2, da alle möglichen Urbilder $\emptyset, \ \Omega_1, \ \Omega_2, \ [0,1]$ relativ offen sind. Für $a \in \Omega_1, \ b \in \Omega_2$ folgt daher mit Satz 4.6.1, dass für ein $x \in [a,b]$ gilt $f(x) = 1/2$, was jedoch der Definition von f widerspricht. □

Korollar 4.6.1. *Sei $E \subset \mathbb{R}^d$ sowohl offen als auch abgeschlossen. Dann gilt $E = \emptyset$ oder $E = \mathbb{R}^d$.*

Beweis. Andernfalls gibt es Punkte $x_0 \in E, \ x_1 \notin E$. Betrachte die stetige Funktion $f\colon [0,1] \to \mathbb{R}^d$ mit $f(t) = (1-t)x_0 + tx_1$. Die Mengen $\Omega_0 = f^{-1}(E)$, $\Omega_1 = f^{-1}(\mathbb{R} \setminus E) \subset [0,1]$ sind gemäss Satz 4.5.2 relativ offen und disjunkt mit $0 \in \Omega_0, \ 1 \in \Omega_1$, und $[0,1] = \Omega_0 \cup \Omega_1$ im Widerspruch zu Satz 4.6.4. □

Definition 4.6.2. *Allgemein heisst $\Omega \subset \mathbb{R}^d$ (stetig)* **wegzusammenhängend**, *falls zu je zwei Punkten $x_0, \ x_1 \in \Omega$ ein stetiger "Weg" $\gamma\colon [0,1] \to \Omega$ existiert mit $\gamma(0) = x_0, \ \gamma(1) = x_1$.*



Beispiel 4.6.3. i) $\mathbb{R}^d$ ist wegzusammenhängend.

ii) $B_R(0) \subset \mathbb{R}^d$ ist wegzusammenhängend, und allgemein jede **konvexe** Menge.

Korollar 4.6.2. *Sei $\Omega \subset \mathbb{R}^d$ wegzusammenhängend, und sei $E \subset \Omega$ sowohl relativ offen als auch relativ abgeschlossen. Dann gilt $E = \emptyset$ oder $E = \Omega$.*

Beweis. Setze $E_0 = E$, $E_1 = \Omega \backslash E$. Dann sind $E_i \subset \Omega$ relativ offen mit $E_0 \cap E_1 = \emptyset$. Nimm widerspruchsweise an $E_0 \neq \emptyset \neq E_1$. Wähle Punkte $x_i \in E_i,\ i = 0, 1$. Da Ω wegzusammenhängend, gibt es einen stetigen Weg $\gamma\colon [0,1] \to \Omega$ mit $\gamma(0) = x_0$, $\gamma(1) = x_1$. Dann sind $\Omega_i = \gamma^{-1}(E_i) \subset [0,1]$ relativ offen, disjunkt, und nichtleer, und $[0,1] = \Omega_0 \cup \Omega_1$ im Widerspruch zu Satz 4.6.4. □

4.7 Supremumsnorm

Sei $\Omega \subset \mathbb{R}^d$.

Definition 4.7.1. *$f\colon \Omega \to \mathbb{R}^n$ heisst auf $\overline{\Omega}$* **stetig ergänzbar**, *falls*

$$\lim_{x \to x_0,\ x \in \Omega} f(x) =: f(x_0)$$

existiert für alle $x_0 \in \overline{\Omega}$, d.h. falls eine stetige Funktion $F\colon \overline{\Omega} \to \mathbb{R}^n$ existiert mit $F|_\Omega = f$.

Beispiel 4.7.1. *Exp*$\colon \mathbb{R} \to \mathbb{R}$ ist die stetige Fortsetzung der Funktion $f(x) = e^x$, $x \in \mathbb{Q}$.

Satz 4.7.1. *Sei $\Omega \subset \mathbb{R}^d$ beschränkt, $f\colon \Omega \to \mathbb{R}^n$ stetig auf $\overline{\Omega}$ ergänzbar. Dann gilt*

$$\|f\|_{C^0} := \sup_{x \in \Omega} \|f(x)\| < \infty,$$

und $\|\cdot\|$ definiert eine Norm auf

$$C^0(\overline{\Omega}; \mathbb{R}^n) = \{f\colon \Omega \to \mathbb{R}^n;\ f \text{ ist auf } \overline{\Omega} \text{ stetig ergänzber}\},$$

die **Supremumsnorm**.

Beweis. Nach Satz 4.2.3 nimmt die stetige Funktion $F = \|\cdot\| \circ f$ auf der kompakten Menge $\overline{\Omega}$ ihr Supremum an einer Stelle $x_0 \in \overline{\Omega}$ an, wo

$$\|f\|_{C^0} = \sup_{x \in \Omega} \|f(x)\| = \|f(x_0)\| < \infty\ .$$

Definitheit und positive Homogenität der Norm sind offensichtlich; die Dreiecks-Ungleichung folgt mit

$$\begin{aligned} \|f + g\|_{C^0} &\leq \sup_{x \in \Omega} \big(\|f(x)\| + \|g(x)\|\big) \leq \sup_{x \in \Omega} \|f(x)\| + \sup_{y \in \Omega} \|g(y)\| \\ &= \|f\|_{C^0} + \|g\|_{C^0}\ . \end{aligned}$$

□

Notation. Zur besseren Unterscheidung der Norm im Funktionenraum C^0 von der Norm in $\mathbb{R}^n$ schreiben wir in Zukunft $|x|$ statt $\|x\|$ für die (euklidische) Norm von $x \in \mathbb{R}^n$.

Wann ist eine stetige Funktion $f\colon \Omega \to \mathbb{R}^n$ auf $\overline{\Omega}$ stetig ergänzbar?

Definition 4.7.2. *$f\colon \Omega \to \mathbb{R}^n$ heisst* **gleichmässig stetig**, *falls gilt*

$$\forall \epsilon > 0\ \exists \delta > 0\ \forall x, y \in \Omega :$$

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon. \tag{4.7.1}$$

Beispiel 4.7.2. Falls $f\colon \Omega \to \mathbb{R}^n$ Lipschitz stetig, so ist f gleichmässig stetig. (Wähle $\delta = \epsilon/L$, wobei $L > 0$ Lipschitz-Konstante von f ist.)

Satz 4.7.2. *Sei $f\colon \Omega \to \mathbb{R}^n$ gleichmässig stetig. Dann kann man f auf $\overline{\Omega}$ stetig ergänzen.*

Beweis. Sei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0 \in \overline{\Omega}$ $(k \to \infty)$. Zu $\epsilon > 0$ wähle $\delta > 0$ gemäss (4.7.1), dazu $k_0 \in \mathbb{N}$ mit

$$|x_k - x_l| < \delta, \quad \forall k \geq k_0.$$

Mit (4.7.1) folgt

$$|f(x_k) - f(x_l)| < \epsilon, \quad \forall k \geq k_0;$$

also ist $(f(x_k))_{k\in\mathbb{N}}$ Cauchy-Folge. Nach Satz 3.6.2 gibt es $a = \lim\limits_{k\to\infty} f(x_k)$, und a ist unabhängig von der Wahl von $(x_k)_{k\in\mathbb{N}}$. □

Umgekehrt gilt

Satz 4.7.3. *Sei $\Omega \subset \mathbb{R}^d$ beschränkt, $f\colon \Omega \to \mathbb{R}^n$ stetig und auf $\overline{\Omega}$ stetig ergänzbar. Dann ist f gleichmässig stetig.*

Beweis. Andernfalls gibt es $\epsilon > 0$ so, dass für jedes $\delta > 0$ Punkte $x, y \in \Omega$ existieren mit

$$|x - y| < \delta \wedge |f(x) - f(y)| \geq \epsilon. \tag{4.7.2}$$

Zu $\delta_k = \frac{1}{k}$ erhalten wir so $x_k,\ y_k \in \Omega$ mit (4.7.2), $k \in \mathbb{N}$. Da $\overline{\Omega}$ kompakt, gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, $x_0 \in \overline{\Omega}$ mit

$$x_k \to x_0 \quad (k \to \infty,\ k \in \Lambda).$$

Also gilt auch $y_k \to x_0$, da

$$|y_k - x_0| \leq \underbrace{|x_k - y_k|}_{<1/k} + |x_k - x_0| \to 0 \quad (k \to \infty,\ k \in \Lambda).$$

Es folgt

$$\epsilon \leq |f(y_k) - f(x_k)| \to |f(x_0) - f(x_0)| = 0 \quad (k \to \infty,\ k \in \Lambda),$$

im Widerspruch zur Voraussetzung. □

4.8 Punktweise und gleichmässige Konvergenz

Sei $\Omega \subset \mathbb{R}^d$, und seien $f,\ f_k \colon \Omega \to \mathbb{R}^n,\ k \in \mathbb{N}$.

Definition 4.8.1. i) *Die Folge* $(f_k)_{k\in\mathbb{N}}$ **konvergiert punktweise** *gegen* f, *falls gilt*

$$f_k(x) \to f(x) \quad (k \to \infty), \quad \forall x \in \Omega.$$

ii) *Die Folge* $(f_k)_{k\in\mathbb{N}}$ **konvergiert gleichmässig** *gegen* f, $f_k \overset{glm.}{\to} f$ $(k \to \infty)$, *falls*

$$\sup_{x\in\Omega} |f_k(x) - f(x)| \to 0 \quad (k \to \infty).$$

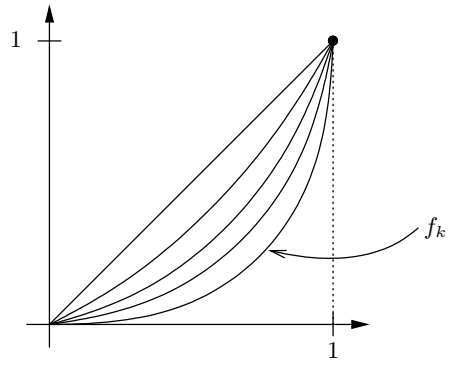
Beispiel 4.8.1. i) Sei $f_k(x) = x^k,\ k \in \mathbb{N},\ 0 \le x \le 1$. Offenbar gilt

$$f_k(x) \overset{(k\to\infty)}{\to} \begin{cases} 0, & x < 1, \\ 1, & x = 1; \end{cases}$$

die Folge $(f_k)_{k\in\mathbb{N}}$ ist also punktweise konvergent. Beachte, dass $f_k \colon [0,1] \to \mathbb{R}$ für jedes $k \in \mathbb{N}$ stetig ist, die Limesfunktion $f \colon [0,1] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & x < 1 \\ 1, & x = 1 \end{cases}$$

hingegen nicht.



ii) Seien $(a_k)_{k\in\mathbb{N}_0} \subset \mathbb{C}$, $p(z) = \sum\limits_{k=0}^{\infty} a_k z^k$ die Potenzreihe in $z \in \mathbb{C}$ mit Konvergenzradius

$$0 \le \rho = \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}} \le \infty.$$

Nimm an, $\rho > 0$. Dann konvergieren die Polynome

$$p_n(z) = \sum_{k=0}^{n-1} a_k z^k$$

gleichmässig gegen p auf $B_r(0)$ für jedes $r < \rho$.

Beweis. Wähle $r < s < \rho$. Für $|z| < r$ schätze ab

$$\begin{aligned}|p(z) - p_n(z)| &\leq \sum_{k=n}^{\infty} |a_k|\,|z|^k \leq \sum_{k=n}^{\infty} |a_k|\,r^k \\ &= \sum_{k=n}^{\infty} |a_k| \left(\frac{r}{s}\right)^k s^k \leq \left(\frac{r}{s}\right)^n \underbrace{\sum_{k=0}^{\infty} |a_k|\,s^k}_{<\infty} \to 0 \quad (n \to \infty),\end{aligned}$$

gleichmässig in z. □

Welche Konsequenzen hat die gleichmässige Konvergenz? Ist p insbesondere stetig in $B_\rho(0)$?

Satz 4.8.1. *Seien $f_k \colon \Omega \subset \mathbb{R}^d \to \mathbb{R}^n$ stetig, $k \in \mathbb{N}$. Weiter gelte $f_k \overset{glm.}{\to} f$ $(k \to \infty)$ für ein $f \colon \Omega \to \mathbb{R}^n$. Dann ist f stetig.*

Beweis. Fixiere $x_0 \in \Omega$. Zu $\epsilon > 0$ wähle $k_0 \in \mathbb{N}$ mit

$$\sup_{x \in \Omega} |f_k(x) - f(x)| < \epsilon, \quad \forall k \geq k_0.$$

Fixiere $k = k_0$. Da f_{k_0} stetig, gibt es $\delta > 0$, so dass

$$|x - x_0| < \delta \Rightarrow |f_{k_0}(x) - f_{k_0}(x_0)| < \epsilon$$

für alle $x \in \Omega$. Es folgt

$$|f(x) - f(x_0)| \leq |f(x) - f_{k_0}(x)| + |f_{k_0}(x) - f_{k_0}(x_0)| + |f_{k_0}(x_0) - f(x_0)| < 3\epsilon$$

für alle $x \in \Omega$ mit $|x - x_0| < \delta$. Nach Satz 4.5.1 ist f stetig an der Stelle x_0; da x_0 beliebig, folgt die Behauptung. □

Korollar 4.8.1. *Potenzreihen sind stetig im Innern ihres Konvergenzkreises.*

Beweis. Unmittelbar aus Satz 4.8.1 mit Beispiel 4.8.1.ii). □

Sind Cauchy-Folgen $(f_k)_{k \in \mathbb{N}}$ in $C^0(\overline{\Omega};\ \mathbb{R}^n)$ analog zu Satz 3.6.2 konvergent?

Satz 4.8.2. *Sei $(f_k)_{k \in \mathbb{N}}$ Folge in $C^0(\overline{\Omega};\ \mathbb{R}^n)$ mit*

$$\|f_k - f_l\|_{C^0} \to 0 \quad (k, l \to \infty).$$

Dann gibt es $f \in C^0(\overline{\Omega};\ \mathbb{R}^n)$ mit $f_k \overset{glm.}{\to} f$ $(k \to \infty)$.

Beweis. Für jedes $x_0 \in \overline{\Omega}$ ist wegen

$$|f_k(x_0) - f_l(x_0)| \leq \sup_{x \in \Omega} |f_k(x) - f_l(x)| = \|f_k - f_l\|_{C^0} \to 0 \ (k, l \to \infty)$$

die Folge $(f_k(x_0))_{k \in \mathbb{N}}$ Cauchy-Folge in $\mathbb{R}^n$. Also existiert gemäss Satz 3.6.2 für jedes $x_0 \in \overline{\Omega}$ der punktweise Limes

$$f(x_0) := \lim_{k \to \infty} f_k(x_0).$$

Weiter gilt

$$\begin{aligned} |f_k(x) - f(x)| &= \lim_{l\to\infty} |f_k(x) - f_l(x)| \\ &\leq \limsup_{l\to\infty} \|f_k - f_l\|_{C^0} \to 0 \quad (k \to \infty), \end{aligned}$$

gleichmässig in x; d.h. $f_k \overset{\text{glm.}}{\to} f$ $(k \to \infty)$, und Satz 4.8.1 ergibt, dass f stetig ist auf $\overline{\Omega}$. □

Der Raum $C^0(\overline{\Omega};\ \mathbb{R}^n)$ ist also **metrisch vollständig** bzgl. der Supremumsnorm, er ist ein **Banachraum**.

Wie $\mathbb{R}$ die rationalen Zahlen $\mathbb{Q}$ als abzählbare dichte Teilmenge enthält, so gibt es auch in $C^0(\overline{\Omega};\ \mathbb{R}^n)$ stets abzählbare dichte Teilmengen. Falls insbesondere $\Omega =]a, b[\subset \mathbb{R}$, so gilt beispielsweise der folgende Satz von Weierstrass.

Satz 4.8.3. (Weierstrass) *Sei $I =]a, b[$, und sei $f \in C^0(\overline{I})$ gegeben. Dann gibt es Polynome p_k, $k \in \mathbb{N}$, mit $p_k \overset{glm.}{\to} f$ $(k \to \infty)$.*

Da man in Satz 4.8.3 die Koeffizienten der Polynome p_k insbesondere auch rational wählen kann, erhält man sogar eine abzählbare Familie, die bzgl. $\|\cdot\|_{C^0}$ in $C^0(\overline{I})$ dicht liegt.

Kapitel 5

Differentialrechnung auf $\mathbb{R}$

5.1 Differential und Differentiationsregeln

Sei $\Omega \subset \mathbb{R}$ offen, $f: \Omega \to \mathbb{R}$, $x_0 \in \Omega$.

Definition 5.1.1. i) *f heisst* **differenzierbar** *an der Stelle x_0, falls der Grenzwert*

$$\lim_{x \to x_0,\ x \neq x_0} \frac{f(x) - f(x_0)}{x - x_0} =: f'(x_0) =: \frac{df}{dx}(x_0)$$

existiert. In diesem Fall heisst $f'(x_0)$ die **Ableitung** *(das Differential) von f an der Stelle x_0.*

ii) *Analog heisst $f = (f_1, \ldots, f_n): \Omega \to \mathbb{R}^n$ an der Stelle x_0 differenzierbar, falls jede der Komponentenfunktionen f_i an der Stelle x_0 differenzierbar ist, und $f'(x_0) = (f_1'(x_0), \ldots, f_n'(x_0))$.*

Bemerkung 5.1.1. Geometrisch entspricht der **Differenzenquotient** $\frac{f(x)-f(x_0)}{x-x_0}$ der Steigung der Sekanten durch die Punkte $(x, f(x))$, $(x_0, f(x_0))$ des Graphen $\mathcal{G}(f)$, das Differential $f'(x_0)$ der Steigung der Tangente an $\mathcal{G}(f)$ im Punkt $(x_0, f(x_0))$.

Definition 5.1.2. *$f: \Omega \to \mathbb{R}^n$ heisst* **auf Ω differenzierbar**, *falls f an jeder Stelle $x_0 \in \Omega$ differenzierbar ist.*

Beispiel 5.1.1. i) Sei $f(x) = mx + b$, $x \in \mathbb{R}$, mit Konstanten $m,\ b \in \mathbb{R}$. Es gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = m, \quad \forall x \neq x_0;$$

also ist f an jeder Stelle $x_0 \in \mathbb{R}$ differenzierbar mit $f'(x_0) = m$.

ii) Die Funktion $f(x) = |x|$, $x \in \mathbb{R}$, ist an der Stelle $x_0 = 0$ nicht differenzierbar,

da

$$\lim_{x\uparrow 0}\frac{f(x)-f(0)}{x-0}=\lim_{x\uparrow 0}\frac{|x|}{x}=-1\neq\lim_{x\downarrow 0}\frac{f(x)-f(x_0)}{x-0}=\lim_{x\downarrow 0}\frac{x}{x}=1.$$

iii) Sei $f(x)=Exp(x),\ x\in\mathbb{R}$. Mit Korollar 3.8.1 folgt für $x_0\neq x=x_0+h\in\mathbb{R}$

$$\begin{aligned}\frac{Exp(x_0+h)-Exp(x_0)}{h}&=\frac{Exp(x_0)(Exp(h)-1)}{h}\\&=Exp(x_0)\underbrace{\sum_{k=1}^{\infty}\frac{h^{k-1}}{k!}}_{\to 1\ (h\to 0)}\to Exp(x_0)\ (h\to 0);\end{aligned}$$

d.h. die Funktion $Exp\colon\mathbb{R}\to\mathbb{R}$ ist an jeder Stelle $x_0\in\mathbb{R}$ differenzierbar mit $Exp'(x_0)=Exp(x_0)$, oder

$$Exp'=Exp.$$

iv) Ebenso wie vektorwertige Funktionen werden Funktionen $f\colon\mathbb{R}\to\mathbb{C}\cong\mathbb{R}^2$ komponentenweise differenziert. Betrachte die Funktion

$$f(t)=Exp(it)=Cos(t)+i\ Sin(t),\quad t\in\mathbb{R}.$$

Analog zu iii) gilt für $t_0\neq t=t_0+h\in\mathbb{R}$

$$\frac{f(t)-f(t_0)}{t-t_0}=Exp(it_0)\frac{Exp(ih)-1}{h}\ .$$

Mit

$$\frac{Exp(ih)-1}{h}\to i\ \ (h\to 0)$$

folgt, dass f an jeder Stelle $t_0\in\mathbb{R}$ differenzierbar ist mit

$$\begin{aligned}f'(t_0)&=Cos'(t_0)+i\ Sin'(t_0)\\&=iExp(it_0)=-Sin(t_0)+i\ Cos(t_0).\end{aligned}$$

D.h.

$$Cos'=-Sin,\quad Sin'=Cos.$$

"Differenzierbarkeit" ist mehr als "Stetigkeit"; genauer gilt:

Satz 5.1.1. *Ist $f\colon\Omega\to\mathbb{R}$ differenzierbar an der Stelle $x_0\in\Omega$, so ist f an der Stelle x_0 auch stetig.*

Beweis. Für $(x_k)_{k\in\mathbb{N}}\subset\Omega$ mit $x_k\to x_0\ (k\to\infty)$ gilt gemäss Satz 3.3.2

$$f(x_k)-f(x_0)=\underbrace{\frac{f(x_k)-f(x_0)}{x_k-x_0}}_{\to f'(x_0)}\cdot\underbrace{(x_k-x_0)}_{\to 0}\to 0\ \ (k\to\infty,\ x_k\neq x_0),$$

bzw.

$$f(x_k)-f(x_0)=0,\ \text{falls } x_k=x_0.$$

Also folgt in jedem Fall $f(x_k)\to f(x_0)\ (k\to\infty)$, wie gewünscht. □

Bemerkung 5.1.2. i) Das Beispiel 5.1.1.ii) der Funktion $f(x) = |x|$, $x \in \mathbb{R}$, zeigt jedoch, dass stetige Funkionen nicht unbedingt überall differenzierbar sein müssen; vgl. auch Beispiel 5.4.3.iv).

ii) Es gibt sogar stetige Funktionen $f\colon \mathbb{R} \to \mathbb{R}$, die an **keiner** Stelle $x_0 \in \mathbb{R}$ differenzierbar sind ("Koch-Kurven").

Satz 5.1.2. *Seien $f, g\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar. Dann sind die Funktionen $f+g$, $f \cdot g$ und, falls $g(x_0) \neq 0$, auch die Funktion f/g an der Stelle x_0 differenzierbar, und es gilt*

i)
$$(f+g)'(x_0) = f'(x_0) + g'(x_0),$$

ii)
$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0),$$

iii)
$$(f/g)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

Beweis. **i)** Für $x \in \Omega,\ x \neq x_0$, folgt mit Satz 3.3.2

$$\begin{aligned}\frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \\ &\overset{(x \to x_0)}{\to} f'(x_0) + g'(x_0);\end{aligned}$$

also ist $f+g$ differenzierbar an der Stelle x_0 mit

$$(f+g)'(x_0) = f'(x_0) + g'(x_0).$$

ii) Analog folgt mit Satz 5.1.1 und

$$\begin{aligned}\frac{(fg)(x) - (fg)(x_0)}{x - x_0} &= \frac{(f(x) - f(x_0))g(x) + f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0} \\ &\overset{(x \to x_0)}{\to} f'(x_0)g(x_0) + f(x_0)g'(x_0)\end{aligned}$$

die gewünschte Aussage für $f \cdot g$.

iii) Mit Satz 5.1.1 folgt aus $g(x_0) \neq 0$, dass $g(x) \neq 0$ für alle x in einer Umgebung von x_0, und $g(x) \to g(x_0)$ $(x \to x_0,\ x \in \Omega)$. Sei $f \equiv 1$. Mit Satz 3.3.2 erhalten wir

$$\frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} = \frac{g(x_0) - g(x)}{x - x_0} \cdot \frac{1}{g(x)g(x_0)} \overset{(x \to x_0,\ x \neq x_0)}{\to} -\frac{g'(x_0)}{g^2(x_0)}.$$

Der allgemeine Fall folgt mit ii).

□

Beispiel 5.1.2. i) Für $n \in \mathbb{N}$ ist die Funktion $f(x) = x^n$, $x \in \mathbb{R}$, differenzierbar mit $f'(x) = nx^{n-1}$.

Beweis (Induktion). $n = 1$: Siehe Beispiel 5.1.1.i).

$n \to n+1$: Setze $f(x) = x^n$, $g(x) = x$. Nach Induktionsvoraussetzung sind f und g differenzierbar mit

$$f'(x) = nx^{n-1},\ g'(x) = 1.$$

Mit Satz 5.1.2.ii) folgt

$$\frac{dx^{n+1}}{dx}(x) = (fg)'(x) = f'(x)g(x) + f(x)g'(x) = (n+1)x^n.$$

□

ii) Polynome $p(x) = a_n x^n + \cdots + a_1 x + a_0$ sind auf $\mathbb{R}$ differenzierbar mit $p'(x) = na_n x^{n-1} + \cdots + a_1$.

iii) Rationale Funktionen $r(x) = \frac{p(x)}{q(x)}$ sind auf ihrem Definitionsbereich

$$\Omega = \{x \in \mathbb{R};\ q(x) \neq 0\}$$

differenzierbar, und

$$r' = \frac{p'q - pq'}{q^2}$$

ist wieder eine rationale Funktion auf Ω.

Satz 5.1.3. (Kettenregel) *Sei $f\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar, und sei $g\colon \mathbb{R} \to \mathbb{R}$ an der Stelle $y_0 = f(x_0)$ differenzierbar. Dann ist die Funktion $g \circ f\colon \Omega \to \mathbb{R}$ an der Stelle x_0 differenzierbar, und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beispiel 5.1.3. Für affine Funktionen

$$f(x) = mx + c,\quad g(y) = ly + d$$

gilt

$$(g \circ f)(x) = lmx + (lc + d),$$

und daher

$$(g \circ f)'(x_0) = lm = g'(f(x_0)) \cdot f'(x_0).$$

Beweis von Satz 5.1.3. Für $x \in \Omega$ mit $f(x) \neq f(x_0)$ schreibe

$$\frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0}.$$

Sei $(x_k)_{k\in\mathbb{N}} \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$, und sei $f(x_k) \neq f(x_0)$, $k \in \mathbb{N}$. Da nach Satz 5.1.1 mit $x_k \to x_0$ auch $f(x_k) \to f(x_0)$ $(k \to \infty)$, folgt

$$\lim_{k\to\infty} \frac{(g \circ f)(x_k) - (g \circ f)(x_0)}{x_k - x_0} = g'(f(x_0))f'(x_0). \tag{5.1.1}$$

Falls für eine Folge $x_k \to x_0$ $(k \to \infty)$ gilt

$$x_k \neq x_0,\ f(x_k) = f(x_0),\ k \in \mathbb{N},$$

so erhalten wir

$$f'(x_0) = \lim_{k\to\infty} \frac{f(x_k) - f(x_0)}{x_k - x_0} = 0,$$

und

$$\lim_{k\to\infty} \frac{g(f(x_k)) - g(f(x_0))}{x_k - x_0} = 0 = g'(f(x_0))f'(x_0).$$

Zusammen mit (5.1.1) liefert dies die gewünschte Konvergenz nun auch für **jede** Folge $(x_k) \subset \Omega$ mit $x_k \to x_0$ $(k \to \infty)$. □

Bemerkung 5.1.3. In der Notation

$$\frac{d(g \circ f)}{dx}(x_0) = \frac{dg}{df}(f(x_0)) \cdot \frac{df}{dx}(x_0)$$

kann man sich die Kettenregel leicht merken. ("df kann man kürzen.")

Beispiel 5.1.4. i) Die Funktion

$$x \mapsto (x^3 + 4x + 1)^2 = x^6 + 8x^4 + 2x^3 + 16x^2 + 8x + 1$$

ist von der Form $g \circ f$ mit

$$g(y) = y^2,\ f(x) = x^3 + 4x + 1.$$

Beispiel 5.1.2.i) und Satz 5.1.3 ergeben

$$\frac{d}{dx}(x^3 + 4x + 1)^2 = \underbrace{2(x^3 + 4x + 1)}_{=g'(f(x))} \cdot \underbrace{(3x^2 + 4)}_{=f'(x)} = 6x^5 + 32x^3 + 6x^2 + 32x + 8.$$

ii) Die Funktion $t \mapsto e^{\lambda t}$, wobei $\lambda \in \mathbb{R}$ fest, ist von der Form $g \circ f$ mit

$$g(x) = e^x,\ f(t) = \lambda t.$$

Mit Beispiel 5.1.2.i) und Beispiel 5.1.1.iii) folgt

$$\frac{d}{dt}(e^{\lambda t})\Big|_{t=t_0} = \underbrace{e^{\lambda t_0}}_{=g'(f(t_0))} \cdot \underbrace{\lambda}_{=f'(t_0)} .$$

iii) Analog gilt mit $g = f = Exp$ gemäss Beispiel 5.1.1.iii)

$$\frac{d}{dx}\left(e^{e^x}\right)\Big|_{x=x_0} = \underbrace{e^{e^{x_0}}}_{=g'(f(x_0))=e^{f(x_0)}} \cdot \underbrace{e^{x_0}}_{=f'(x_0)} .$$

5.2 Der Mittelwertsatz und Folgerungen

Im folgenden betrachten wir stets differezierbare Funktionen auf einem Intervall $\Omega =]a, b[\subset \mathbb{R}$.

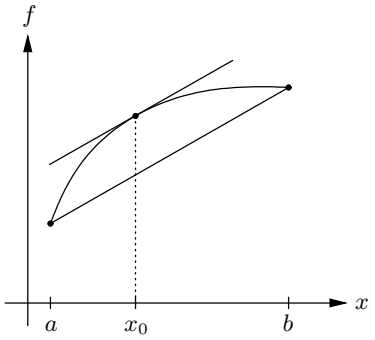
Satz 5.2.1. *Seien $-\infty < a < b < \infty$. Sei $f\colon [a,b] \to \mathbb{R}$ stetig und differenzierbar in $]a,b[$. Dann existiert $x_0 \in]a,b[$ mit*

$$f(b) = f(a) + f'(x_0)(b-a).$$

D.h.

$$f'(x_0) = \frac{f(b)-f(a)}{b-a}$$

ist die Steigung der Sekante durch die Punkte $(a,f(a))$, $(b,f(b)) \in \mathcal{G}(f)$.



Beweis. **i)** Zunächst betrachten wir den Fall, dass zusätzlich gilt

$$f(a) = f(b) = 0.$$

Nach Satz 4.2.3 und Beispiel 4.2.3.i) gibt es $\underline{x},\ \overline{x} \in [a,b]$ mit

$$f(\underline{x}) = \min_{a\le x\le b} f(x) \le 0 \le \max_{a\le x\le b} f(x) = f(\overline{x}).$$

Falls

$$f(\underline{x}) = 0 = f(\overline{x}),$$

so ist $f \equiv 0$; also $f'(x) = 0$ für alle $x \in [a,b]$, insbesondere $f'(\frac{a+b}{2}) = 0$.

Andernfalls gelte oBdA $f(\overline{x}) > 0$. (Sonst betrachte die Funktion $\widetilde{f} = -f$.) Dann gilt offenbar $a < \overline{x} < b$, und es folgt

$$0 \ge \lim_{x\downarrow\overline{x}} \frac{f(x)-f(\overline{x})}{x-\overline{x}} = f'(\overline{x}) = \lim_{x\uparrow\overline{x}} \frac{f(x)-f(\overline{x})}{x-\overline{x}} \ge 0,$$

also $f'(\overline{x}) = 0$.

ii) Für allgemeines f betrachte die Funktion $g\colon [a,b] \to \mathbb{R}$ mit

$$g(x) = f(x) - \Big(f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\Big).$$

Offenbar ist g stetig auf $[a,b]$ und in $]a,b[$ differenzierbar mit

$$g'(x) = f'(x) - \frac{f(b)-f(a)}{a-b},\ x \in]a,b[.$$

Weiter gilt $g(a) = 0 = g(b)$. Mit i) folgt die Existenz von $x_0 \in]a, b[$ mit $g'(x_0) = 0$. Die Behauptung folgt.

□

Als erste Anwendung folgt sofort:

Korollar 5.2.1. *Sei f wie in Satz 5.2.1.*

i) *Falls $f' \equiv 0$ auf $]a, b[$, so ist f konstant.*

ii) *Falls $f' \geq 0$ (bzw. > 0) auf $]a, b[$, so ist f (streng) monoton wachsend.*

Beweis. **i)** Für $a \leq x < y \leq b$ gibt es $x_0 \in]x, y[$ mit

$$\frac{f(y) - f(x)}{y - x} = f'(x_0) = 0.$$

ii) analog.

□

Beispiel 5.2.1. i) Für eine differenzierbare Funktion $f\colon \mathbb{R} \to \mathbb{R}$ gelte die Beziehung

$$f' = \lambda f, \text{ d.h. } f'(t) = \lambda f(t), \ \forall t \in \mathbb{R},$$

mit einer festen Zahl $\lambda \in \mathbb{R}$. Dann gilt

$$f(t) = f(0)e^{\lambda t}, \quad \forall t \in \mathbb{R}.$$

Beweis. Betrachte die differenzierbare Funktion g mit

$$g(t) = \frac{f(t)}{e^{\lambda t}} = e^{-\lambda t} f(t), \quad t \in \mathbb{R}.$$

Mit Satz 5.1.2 und Beispiel 5.1.4.ii) folgt

$$g'(t) = \frac{d}{dt}\big(e^{-\lambda t}\big) \cdot f(t) + e^{-\lambda t} f'(t) = e^{-\lambda t}\big(-\lambda f(t) + f'(t)\big) = 0$$

für alle $t \in \mathbb{R}$; also

$$g(t) \equiv g(0) = f(0)$$

gemäss Korollar 5.2.1.i), und

$$f(t) = e^{\lambda t} g(t) = f(0)e^{\lambda t}, \quad \forall t \in \mathbb{R}.$$

□

ii) Die Funktion $f\colon x \mapsto \frac{2x}{1+x^2}$ erfüllt gemäss Satz 5.1.2

$$f'(x) = \frac{2(1 + x^2) - 4x^2}{(1 + x^2)^2} = \frac{2(1 - x^2)}{(1 + x^2)^2} < 0$$

für $x > 1$; also ist $f\colon]1, \infty[\to]0, 1[$ streng monoton fallend.

Korollar 5.2.2. (Bernoulli-de l'Hospital) *Seien $f, g\colon [a,b] \to \mathbb{R}$ stetig und differenzierbar in $]a,b[$ mit $g'(x) \neq 0$, $\forall x \in]a,b[$. Weiter sei $f(a) = 0 = g(a)$, und es existiere der Grenzwert*

$$\lim_{x \downarrow a} \frac{f'(x)}{g'(x)} =: A.$$

Dann ist $g(x) \neq 0$ für alle $x > a$, und es gilt

$$\lim_{x \downarrow a} \frac{f(x)}{g(x)} = A.$$

Beweis. **i)** Falls $g(x) = 0$ für ein $x > a$, so gibt es nach Satz 5.2.1 ein $x_0 \in]a,x[$ mit $g'(x_0) = 0$ im Widerspruch zu unserer Annahme.

ii) Für festes $s > a$ betrachte die Funktion

$$h(x) = \frac{f(s)}{g(s)} g(x) - f(x), \quad x \in [a,s].$$

Die Funktion $h\colon [a,s] \to \mathbb{R}$ ist stetig und differenzierbar in $]a,s[$ mit $h(a) = 0 = h(s)$. Nach Satz 5.2.1 gibt es $x = x(s) \in]a,s[$ mit

$$0 = h'(x) = \frac{f(s)}{g(s)} g'(x) - f'(x);$$

d.h.

$$\frac{f(s)}{g(s)} = \frac{f'(x)}{g'(x)}.$$

Mit $s \to a$ folgt dann $x(s) \to a$, und

$$\frac{f(s)}{g(s)} \to \lim_{x \downarrow a} \frac{f'(x)}{g'(x)} = A.$$

□

Beispiel 5.2.2. i) Es gilt

$$\lim_{x \to 1} \frac{x^3 - 1}{x^2 - 1} = \lim_{x \to 1} \frac{3x^2}{2x} = \frac{3}{2}.$$

Man kann zur Probe den Faktor $x - 1$ in Zähler und Nenner kürzen:

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

ii) Mit Beispiel 5.1.1.iv) erhalten wir

$$\lim_{x \to 0} \frac{Sin(x)}{x} = \lim_{x \to 0} \frac{Cos(x)}{1} = 1.$$

iii) Man kann die Bernoullische Regel auch mehrmals anwenden. Mit ii) folgt so

$$\lim_{x\to 0}\frac{1-Cos(x)}{x^2}=\lim_{x\to 0}\frac{Sin(x)}{2x}=\frac{1}{2}.$$

Oft kann man in ähnlichen Fällen jedoch auch ohne Gebrauch der de l'Hospitalschen Regeln durch geschicktes Umformen zum Ziel gelangen.

Beispiel 5.2.3. i) Beachte, dass mit

$$\frac{Exp(x)}{x^k}\geq\frac{x^{k+1}}{(k+1)!}\cdot\frac{1}{x^k}=\frac{x}{(k+1)!}\to\infty\ (0<x\to\infty)$$

auch gilt

$$\lim_{x\to\infty}\left(x^k e^{-x}\right)=\lim_{x\to\infty}\frac{1}{\left(\frac{Exp(x)}{x^k}\right)}=0.$$

ii) Für beliebiges $\alpha>0$ definieren wir

$$x^\alpha:=Exp(\alpha Log(x))=e^{-y},\quad x>0,\ y=-\alpha Log(x).$$

Analog zu i) wollen wir nun den Limes des Ausdrucks $x^\alpha Log(x)$ für $x\downarrow 0$ bestimmen. Nach Substitution $y=-\alpha Log(x)\to\infty$ $(x\downarrow 0)$ erhalten wir

$$\lim_{x\downarrow 0}(x^\alpha Log(x))=-\frac{1}{\alpha}\lim_{y\to\infty}\left(e^{-y}y\right)=0.$$

iii) Da Exp stetig, folgt mit ii) nun auch

$$\lim_{x\downarrow 0}x^x=\lim_{x\downarrow 0}\Big(Exp(xLog(x))\Big)=Exp(0)=1;$$

vgl. Beispiel 3.2.1.iii).

Eine weitere Anwendung des Mittelwertsatzes erhalten wir durch Koppelung von Korollar 5.2.1.ii) mit Satz 4.6.3.

Satz 5.2.2. (Umkehrsatz) *Sei $f\colon]a,b[\to\mathbb{R}$ differenzierbar mit $f'>0$ auf $]a,b[$, und seien*

$$-\infty\leq c=\inf_{a<x<b}f(x)<\sup_{a<x<b}f(x)=d\leq\infty.$$

Dann ist $f\colon]a,b[\to]c,d[$ bijektiv, und die Umkehrfunktion $f^{-1}\colon]c,d[\to\mathbb{R}$ ist differenzierbar mit

$$\left(f^{-1}\right)'(f(x))=\left(f'(x)\right)^{-1},\ \forall x\in]a,b[,$$

bzw.

$$\left(f^{-1}\right)'(y)=\frac{1}{f(f^{-1}(y))},\quad\forall y\in]c,d[.$$

Beweis. Gemäss Korollar 5.2.1.ii) ist f streng monoton wachsend und zudem stetig nach Satz 5.1.1. Nach Satz 4.6.3 ist $f\colon]a,b[\to]c,d[$ bijektiv, und f^{-1} ist stetig.

Behauptung f^{-1} ist differenzierbar, $\left(f^{-1}\right)'(f(x_0)) = \frac{1}{f'(x_0)},\ \forall x_0$.

Beweis. Fixiere $y_0 = f(x_0)$. Sei $(y_k)_{k\in\mathbb{N}} \subset]c,d[$ mit

$$y_k = f(x_k) \to y_0\ (k\to\infty),\quad y_k \neq y_0\ (k\in\mathbb{N}).$$

Es folgt $x_k \neq x_0$ für alle k. Da f^{-1} stetig, gilt zudem

$$x_k = f^{-1}(y_k) \to x_0 = f^{-1}(y_0)\ (k\to\infty),$$

also

$$\frac{f^{-1}(y_k)-f^{-1}(y_0)}{y_k-y_0} = \frac{x_k-x_0}{f(x_k)-f(x_0)} = \frac{1}{\frac{f(x_k)-f(x_0)}{x_k-x_0}} \to \frac{1}{f'(x_0)}$$

für $k\to\infty$, wie gewünscht. □

Beispiel 5.2.4. i) $Exp\colon \mathbb{R}\to]0,\infty[$ ist differenzierbar mit $Exp' = Exp > 0$. Also ist $Log = Exp^{-1}\colon]0,\infty[\to\mathbb{R}$ differenzierbar mit

$$Log'(Exp(x_0)) = \frac{1}{Exp'(x_0)} = \frac{1}{Exp(x_0)},$$

oder -nach Substitution von $y = Exp(x_0)$-

$$Log'(y) = \frac{1}{y},\quad \forall y > 0.$$

ii) Wir können nun auch die in Beispiel 5.2.3 definierte allgemeine Potenzfunktion

$$x \mapsto x^\alpha = Exp(\alpha Log(x)),\quad 0 < x < \infty,$$

für beliebiges $\alpha > 0$ differenzieren. Mit der Kettenregel aus Satz 5.1.3 erhalten wir

$$\frac{dx^\alpha}{dx}\Big|_{x=x_0} = \underbrace{Exp'(\alpha Log(x_0))}_{=x_0^\alpha}\,\alpha\,\underbrace{Log'(x_0)}_{=\frac{1}{x_0}} = \alpha x_0^{\alpha-1}.$$

Stimmt diese Funktion für $\alpha = \frac{1}{n}$ überein mit der "klassischen" n-ten Wurzelfunktion gemäss Beispiel 4.6.2.i)?

iii) Die Potenzfunktion $f(x) = x^n$ ist für $x > 0$ differenzierbar mit $f'(x) = nx^{n-1} > 0$. Gemäss Satz 5.2.2 ist $f^{-1}(y) = \sqrt[n]{y}\colon]0,\infty[\to\mathbb{R}$ differenzierbar mit

$$\left(f^{-1}\right)'\Big|_{y=x_0^n} = \frac{1}{f'(x_0)} = \frac{1}{n}x_0^{1-n},$$

bzw.

$$\left(\sqrt[n]{y}\right)'\Big|_{y=y_0} = \frac{1}{n}y_0^{\frac{1}{n}-1},\ \forall y_0 > 0.$$

Mit ii) folgt

$$\frac{d}{dx}\big((\sqrt[n]{x}) - x^{1/n}\big) = 0 \text{ auf }]0,\infty[,$$

wobei $x^{1/n} = Exp(Log(x)/n)$, und mit Korollar 5.2.1.i) folgt

$$\sqrt[n]{x} - x^{1/n} \equiv \sqrt[n]{1} - 1^{1/n} = 0.$$

Die abstrakt definierte Potenzfunktion $x \mapsto x^\alpha,\ x > 0$, stimmt also für $\alpha = \frac{1}{n}$ mit der Wurzelfunktion überein; analog für $\alpha \in \mathbb{Q}$.

5.3 Die trigonometrischen Funktionen

Wir können nun endlich auch die lange vermutete Verbindung zwischen den über ihre Potenzreihe definierten Funktionen Sin und Cos und den trigonometrischen Funktionen sin und cos herstellen.

Satz 5.3.1. (Euler) *Für alle $\varphi \in \mathbb{R}$ gilt*

$$|Exp(i\varphi)|^2 = Cos^2(\varphi) + Sin^2(\varphi) = 1,$$

und

$$Exp(i\varphi) = Cos(\varphi) + iSin(\varphi) = \cos\varphi + i\sin\varphi = e^{i\varphi},$$

wobei $\cos\varphi,\ \sin\varphi$ Real-, bzw. Imaginärteil der Zahl $z = e^{i\varphi} \in \mathbb{C}$ mit $|z| = 1$ und Polarwinkel φ bezeichnen.

Beweis. **i)** Da $Cos(-\varphi) = Cos(\varphi),\ Sin(-\varphi) = -Sin(\varphi)$ für alle $\varphi \in \mathbb{R}$, gilt

$$\overline{Exp(i\varphi)} = Cos(\varphi) - iSin(\varphi) = Exp(-i\varphi),$$

also auch

$$|Exp(i\varphi)|^2 = Exp(i\varphi)\cdot\overline{Exp(i\varphi)} = Exp(i\varphi)\cdot Exp(-i\varphi) = 1, \quad \forall\varphi \in \mathbb{R}.$$

ii) Gemäss Beispiel 5.1.1.iv) gilt weiter

$$\frac{d}{d\varphi}Exp(i\varphi) = iExp(i\varphi)$$

mit

$$\left|\frac{d}{d\varphi}Exp(i\varphi)\right| = |Exp(i\varphi)| = 1;$$

d.h. die Kurve $\varphi \mapsto Exp(i\varphi)$ durchläuft den Enheitskreis im Gegenuhrzeigersinn mit Geschwindigkeit 1. Da $Exp(0) = 1$, stimmt das Argument φ des Punktes $Exp(i\varphi) \in \mathbb{C}$ überein mit der Bogenlänge am Einheitskreis.

□

Korollar 5.3.1. $Exp(z + 2\pi i) = Exp(z), \quad \forall z \in \mathbb{C}.$

Beweis. Gemäss Satz 5.3.1 gilt $Exp(2\pi i) = e^{2\pi i} = 1$, und die Behauptung folgt mit Korollar 3.8.1. □

Zyklometrische Funktionen (Arcus-Funktionen). Im folgenden schreiben wir *cos* statt *Cos*, etc. Mit Satz 5.2.2 können wir diese Funktionen auf geeigneten Intervallen auch umkehren.

i) Da $\sin' = \cos > 0$ in $]-\frac{\pi}{2}, \frac{\pi}{2}[$, besitzt $\sin\colon]-\frac{\pi}{2}, \frac{\pi}{2}[\to]-1, 1[$ gemäss Satz 5.2.2 eine differenzierbare Umkehrfunktion $\arcsin = \sin^{-1}\colon]-1, 1[\to]-\frac{\pi}{2}, \frac{\pi}{2}[$, und

$$\arcsin'(x) = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1,$$

wobei wir ausnutzen, dass $\cos = \sqrt{1-\sin^2}$ gemäss Satz 5.3.1.

ii) Analog besitzt $\cos\colon]0, \pi[\to]-1, 1[$ mit $\cos' = -\sin < 0$ in $]0, \pi[$ die differenzierbare Umkehrfunktion $\arccos\colon]-1, 1[\to]0, \pi[$, und

$$\arccos'(x) = \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1.$$

iii) Die Tangensfunktion $\tan = \frac{\sin}{\cos}\colon]-\frac{\pi}{2}, \frac{\pi}{2}[\to \mathbb{R}$ mit

$$\tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x) = \frac{1}{\cos^2(x)} > 0$$

besitzt die differenzierbare Umkehrfunktion $\arctan\colon \mathbb{R} \to]-\frac{\pi}{2}, \frac{\pi}{2}[$ mit

$$\arctan'(x) = \frac{1}{\tan'(\arctan x)} = \frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

Hyperbel- und Areafunktionen. Die Hyperbelfunktionen

$$\cosh x = \frac{e^x + e^{-x}}{2}\colon \mathbb{R} \to [1, \infty[,$$
$$\sinh x = \frac{e^x - e^{-x}}{2}\colon \mathbb{R} \to \mathbb{R},$$
$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}\colon \mathbb{R} \to]-1, 1[$$

erfüllen die Gleichungen

$$\sinh' = \cosh, \quad \cosh' = \sinh,$$

bzw.

$$\tanh' = \frac{\cosh^2 - \sinh^2}{\cosh^2} = \frac{1}{\cosh^2} = 1 - \tanh^2 > 0,$$

wobei wir die Beziehung ausnutzen

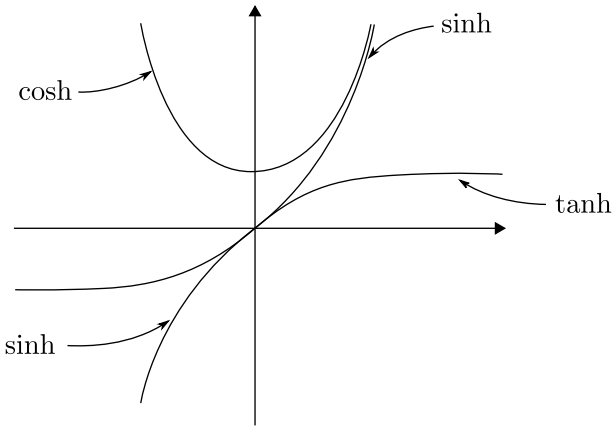
$$\cosh^2(x) - \sinh^2(x) = e^x \cdot e^{-x} = 1.$$

Insbesondere existieren die Areafunktionen

$$\begin{aligned} arsinh &= \sinh^{-1}\colon \mathbb{R} \to \mathbb{R}, \\ arcosh &= \cosh^{-1}\colon [1,\infty[\to \mathbb{R}, \\ artanh &= \tanh^{-1}\colon]-1,1[\to \mathbb{R} \end{aligned}$$

mit

$$\begin{aligned} arsinh'(x) &= \frac{1}{\cosh(arsinhx)} = \frac{1}{\sqrt{1+x^2}},\ x \in \mathbb{R}, \\ artanh'(x) &= \frac{1}{(1-\tanh^2)(artanh(x))} = \frac{1}{1-x^2}, \quad -1 < x < 1. \end{aligned}$$



5.4 Funktionen der Klasse C^1

Sei $\Omega \subset \mathbb{R}$ offen, $f\colon \Omega \to \mathbb{R}$ differenzierbar.

Definition 5.4.1. *f heisst* **von der Klasse C^1,** *falls die Ableitungsfunktion $x \mapsto f'(x)$ stetig ist.*

Notation: $C^1(\Omega) = \{f\colon \Omega \to \mathbb{R};\ f,\ f' \text{ stetig}\}$.

Beispiel 5.4.1. i) Die Funktionen Exp, Cos, Sin und Polynome sind in C^1.

ii) Sei die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Dann ist f stetig und an jeder Stelle $x \neq 0$ differenzierbar mit

$$f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}, \quad x \neq 0.$$

Weiter existiert

$$f'(0) = \lim_{x \to 0,\ x \neq 0} \frac{f(x)}{x} = \lim_{x \to 0} x \sin \frac{1}{x} = 0;$$

jedoch ist f' an der Stelle $x_0 = 0$ unstetig.

iii) Für $k \geq 3$ ist die Funktion

$$f(x) = \begin{cases} x^k \sin \frac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

von der Klasse C^1 auf $\mathbb{R}$.

Satz 5.4.1. *Sei $(f_k)_{k \in \mathbb{N}}$ eine Folge in $C^1(\Omega)$ mit*

$$f_k \overset{glm}{\to} f, \ f_k' \overset{glm}{\to} g \quad (k \to \infty),$$

wobei $f, g \colon \Omega \to \mathbb{R}$. Dann gilt $f \in C^1(\Omega)$ und $f' = g$.

Beweis. Nach Satz 4.8.1 sind f und g stetig. Die Aussage folgt somit aus

Behauptung f ist differenzierbar mit $f' = g$.

Beweis. Für $x_0, x \in \Omega$, $x \neq x_0$, gilt nach Satz 5.2.1

$$\begin{aligned} \left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| &= \lim_{k \to \infty} \left| \frac{f_k(x) - f_k(x_0)}{x - x_0} - g(x_0) \right| \\ &= \lim_{k \to \infty} |f_k'(x_k) - g(x_0)| \\ &\leq \lim_{k \to \infty} |f_k'(x_k) - g(x_k)| + \sup_{|y - x_0| < |x - x_0|} |g(y) - g(x_0)| \\ &\leq \sup_{|y - x_0| < |x - x_0|} |g(y) - g(x_0)| \to 0 \quad (x \to x_0). \end{aligned}$$

wobei $x_k = x_0 + \vartheta_k (x - x_0)$ mit geeignetem $0 < \vartheta_k < 1$. Also ist f an der Stelle x_0 differenzierbar mit $f'(x_0) = g(x_0)$. □

Offenbar gilt Satz 5.4.1 auch für vektorwertige Funktionen.

Beispiel 5.4.2. Die gleichmässige Konvergenz $f_k' \to g$ $(k \to \infty)$ ist notwendig für die Aussage von Satz 5.4.1, wie das folgende Beispiel zeigt. Sei

$$f_k(x) = \sqrt{\left(\frac{1}{k}\right)^2 + x^2}, \quad |x| < 1, \ k \in \mathbb{N}.$$

Mit der Abschätzung

$$|f_k(x) - |x|| = \left(\sqrt{\left(\frac{1}{k}\right)^2 + x^2} - |x|\right) \cdot \frac{\sqrt{\left(\frac{1}{k}\right)^2 + x^2} + |x|}{\sqrt{\left(\frac{1}{k}\right)^2 + x^2} + |x|}$$
$$= \frac{\left(\frac{1}{k}\right)^2}{\sqrt{\left(\frac{1}{k}\right)^2 + x^2} + |x|} \leq \frac{(1/k)^2}{1/k} \overset{(k\to\infty)}{\to} 0$$

erhalten wir gleichmässige Konvergenz $f_k \overset{\text{glm}}{\to} f$ für $|x| < 1$, wobei $f(x) = |x|$. Zudem konvergiert

$$f_k'(x) = \frac{x}{\sqrt{\left(\frac{1}{k}\right)^2 + |x|^2}} \to \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

punktweise; jedoch ist f nicht in C^1.

Potenzreihen. Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$, und sei

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad |x| < \rho = \frac{1}{\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|}} \leq \infty$$

die zugehörige Potenzreihe $f = \lim\limits_{n\to\infty} f_n$, wobei

$$f_n(x) = \sum_{k=0}^{n} a_k x^k, \quad x \in \mathbb{R}.$$

Nach Beispiel 5.1.2.ii) ist jedes f_n differenzierbar mit

$$f_n'(x) = \sum_{k=0}^{n} k a_k x^{k-1}.$$

Die Potenzreihe

$$g(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}$$

hat wegen $\sqrt[k]{k} \to 1$ $(k \to \infty)$ denselben Konvergenzradius ρ wie f, und wie in Beispiel 4.8.1.ii) folgt für $r < \rho$ gleichmässige Konvergenz

$$f_n \to f, \ f_n' \to g \quad \text{in } B_r(0) \ (n \to \infty) \ .$$

Satz 5.4.1 liefert somit das folgende Resultat.

Satz 5.4.2. *Eine Potenzreihe $f(x) = \sum\limits_{k=0}^{\infty} a_k x^k$ ist im Innern ihres Konvergenzkreises differenzierbar, und*

$$f'(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}.$$

Beispiel 5.4.3. i) Es gilt

$$Exp'(x) = \sum_{k=0}^{\infty} k\frac{x^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = Exp(x).$$

ii) Sei $f(x) = \sum\limits_{k=0}^{\infty} x^k = \frac{1}{1-x}$, $|x| < 1$. Dann folgt mit der Quotientenregel, bzw. mit Satz 5.4.2

$$f'(x) = \frac{1}{(1-x)^2} = \sum_{k=0}^{\infty} kx^{k-1}, \quad |x| < 1;$$

vgl. Beispiel 3.8.1.

iii) Analog zu Satz 5.4.2 zeigt man, die Zetafunktion

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x} = \sum_{n=1}^{\infty} e^{-x\cdot\log n}$$

ist für $x > 1$ differenzierbar mit

$$\zeta'(x) = -\sum_{n=1}^{\infty} \log n e^{-x\log n} = -\sum_{n=1}^{\infty} \frac{\log n}{n^x} =: \eta(x).$$

Beweis. Schreibe $\zeta = \lim\limits_{l\to\infty} \zeta_l$, wobei

$$\zeta_l(x) = \sum_{n=1}^{l} \frac{1}{n^x}, \quad \zeta_l'(x) = -\sum_{n=1}^{l} \frac{\log n}{n^x}.$$

Schätze ab für $x \geq s > r > 1$:

$$|\zeta(x) - \zeta_l(x)| \leq \sum_{n>l} \frac{1}{n^s} \to 0 \quad (l \to \infty)$$

$$|\eta(x) - \zeta_l'(x)| \leq \sum_{n>l} \frac{\log n}{n^s} \leq \sup_{n>l} \frac{\log n}{n^{s-r}} \cdot \sum_{n>l} \frac{1}{n^r} \to 0 \ (l \to \infty).$$

Die Behauptung folgt mit Satz 5.4.1. □

Sei nun $\Omega \subset \mathbb{R}$ offen und beschränkt. Analog zu Abschnitt 4.7 setzen wir

$$C^1(\overline{\Omega}; \mathbb{R}^n) = \{f \in C^1(\Omega; \mathbb{R}^n);\ f \text{ und } f' \text{ sind stetig auf } \overline{\Omega} \text{ ergänzbar}\}.$$

Für $f \in C^1(\overline{\Omega}; \mathbb{R}^n)$ gilt dann

$$\|f\|_{C^1(\overline{\Omega})} = \sup_{x\in\Omega} |f(x)| + \sup_{x\in\Omega} |f'(x)| = \|f\|_{C^0(\overline{\Omega})} + \|f'\|_{C^0(\overline{\Omega})} < \infty.$$

Offenbar ist $\|\cdot\|_{C^1(\overline{\Omega})}$ eine Norm auf $C^1(\overline{\Omega}; \mathbb{R}^n)$. Weiter liefern Satz 4.8.2 und Satz 5.4.1:

Satz 5.4.3. *Der Raum $C^1(\overline{\Omega}, \mathbb{R}^n)$ ist vollständig bzgl. $\|\cdot\|_{C^1(\overline{\Omega})}$, ein Banachraum.*

Beweis. Sei $(f_k)_{k\in\mathbb{N}} \subset C^1(\overline{\Omega})$ Cauchy-Folge. Dann sind $(f_k)_{k\in\mathbb{N}}$, $(f_k')_{k\in\mathbb{N}}$ Cauchy-Folgen in $C^0(\Omega, \mathbb{R}^n)$ mit Limites $f = \lim\limits_{k\to\infty} f_k$, $g = \lim\limits_{k\to\infty} f_k' \in C^0(\overline{\Omega})$ gemäss Satz 4.8.2, und $g = f'$ gemäss Satz 5.4.1; also $f \in C^1(\overline{\Omega})$, $f_k \overset{C^1\text{-glm}}{\to} f$ $(k \to \infty)$. □

Iterativ können wir auch höhere Ableitungen bilden. Sei $m \in \mathbb{N}$.

Definition 5.4.2. i) *f heisst* **auf Ω m-mal differenzierbar**, *falls f $(m-1)$-mal differenzierbar ist mit differenzierbarer $(m-1)$-ter Ableitung $f^{(m-1)}$.*

In diesem Fall heisst

$$f^{(m)} = \frac{df^{(m-1)}}{dx} = \frac{d^m f}{dx^m} \colon \Omega \to \mathbb{R}$$

die m-te **Ableitung** *von f.*

ii) *f ist* **von der Klasse** $C^m(\Omega)$, *falls f m-mal differenzierbar ist und falls die Funktionen $f = f^{(0)}$, $f' = f^{(1)}$, ... , $f^{(m)}$ stetig sind.*

Notation:

$$C^m(\Omega, \mathbb{R}^n) = \{f \colon \Omega \to \mathbb{R}^n;\ f \text{ ist } m\text{-mal diffbar, } f, \ldots, f^{(m)} \text{ stetig}\}.$$

Beispiel 5.4.4. i) Die Funktionen Exp, sin, cos, Polynome und rationale Funktionen sind in C^m für jedes $m \in \mathbb{N}$.

ii) Potenzreihen mit Konvergenzradius $\rho > 0$ sind in $C^m(B_\rho(0))$ für jedes m; die Ableitung erhält man durch gliedweises Differenzieren.

Falls $\Omega \subset \mathbb{R}$ offen und beschränkt, setzen wir

$$C^m(\overline{\Omega}; \mathbb{R}^n) = \{f \in C^m(\Omega; \mathbb{R}^n);\\ f = f^{(0)}, f' = f^{(1)}, \ldots, f^{(m)} \text{ sind stetig auf } \overline{\Omega} \text{ ergänzbar}\}$$

und definieren

$$\|f\|_{C^m} = \sum_{j=0}^{m} \left\| f^{(j)} \right\|_{C^0}, \quad \forall f \in C^m(\overline{\Omega}; \mathbb{R}^n).$$

Dann gilt ofenbar

$$C^m(\overline{\Omega}; \mathbb{R}^n) \hookrightarrow C^{m-1}(\overline{\Omega}; \mathbb{R}^n) \hookrightarrow \ldots \hookrightarrow C^0(\overline{\Omega}; \mathbb{R}^n)$$

und

$$\|f\|_{C^0} \le \|f\|_{C^1} \le \cdots \le \|f\|_{C^m}, \quad \forall f \in C^m(\overline{\Omega}; \mathbb{R}^n).$$

Analog zu Satz 5.4.3 gilt schliesslich

Satz 5.4.4. *Der Raum $C^m(\overline{\Omega};\mathbb{R}^n)$ ist vollständig bzgl. $\|\cdot\|_{C^m(\overline{\Omega})}$.*

Beweis. OBdA sei $n = 1$. Falls $(f_k)_{k\in\mathbb{N}}$ Cauchy-Folge im Raum $C^m(\overline{\Omega})$ ist, so ist $(f_k)_{k\in\mathbb{N}}$ Cauchy-Folge in $C^0(\overline{\Omega})$ mit $\lim\limits_{k\to\infty} f_k = f =: g_0$. Ebenso ist für jedes $0 < j \le m$ die Folge $(f_k^{(j)})_{k\in\mathbb{N}}$ Cauchy-Folge in $C^0(\overline{\Omega})$ mit $\lim\limits_{k\to\infty} f_k^{(j)} = g_j$. Da $f_k^{(j+1)} = df_k^{(j)}/dx$, folgt mit Satz 5.4.1 die Beziehung $g_{j+1} = g_j'$ für alle $j < m$, also $g_1 = f'$, $g_2 = f''$, …, $g_m = f^{(m)}$, und wir erhalten $f \in C^m(\overline{\Omega})$, $\|f_k - f\|_{C^m(\overline{\Omega})} \to 0 \ \ (k \to \infty)$, wie gewünscht. □

5.5 Taylor-Formel

Sei $\Omega =]a,b[$, $-\infty < a < b < \infty$, und $m \in \mathbb{N}$.

Satz 5.5.1. *Sei $f \in C^{m-1}([a,b])$ auf $]a,b[$ m-mal differenzierbar. Dann gibt es $\xi \in]a,b[$ mit*

$$f(b) = f(a) + f'(a)(b-a) + f''(a)\frac{(b-a)^2}{2} + \dots \\ \cdots + f^{(m-1)}(a)\frac{(b-a)^{m-1}}{(m-1)!} + f^{(m)}(\xi)\frac{(b-a)^m}{m!}.$$

Beweis. Wir führen den Satz zurück auf Satz 5.2.1. Betrachte die Funktion

$$g(x) = f(x) + f'(x)(b-x) + \dots \\ + f^{(m-1)}(x)\frac{(b-x)^{m-1}}{(m-1)!} + K\frac{(b-x)^m}{m!} - f(b), \tag{5.5.1}$$

wobei $K \in \mathbb{R}$ so gewählt ist, dass $g(a) = g(b) = 0$. Nach Annahme an f ist g stetig auf $[a,b]$, und in $]a,b[$ differenzierbar. Nach Satz 5.2.1 existiert ein $\xi \in]a,b[$ mit $g'(\xi) = 0$; d.h.

$$\begin{aligned} 0 &= f'(\xi) + \Big(f''(\xi)(b-\xi) - f'(\xi)\Big) + \Big(f'''(\xi)\frac{(b-\xi)^2}{2} - f''(\xi)(b-\xi)\Big) + \dots \\ &\quad + \Big(f^{(m)}(\xi)\frac{(b-\xi)^{m-1}}{(m-1)!} - f^{(m-1)}(\xi)\frac{(b-\xi)^{m-2}}{(m-2)!}\Big) - K\frac{(b-\xi)^{m-1}}{(m-1)!} \\ &= \big(f^{(m)}(\xi) - K\big)\frac{(b-\xi)^{m-1}}{(m-1)!}, \end{aligned}$$

da sich alle übrigen Terme paarweise aufheben. Da $b - \xi > 0$ folgt $K = f^{(m)}(\xi)$, und mit $g(a) = 0$ erhalten wir nach Einsetzen von $x = a$ in (5.5.1) die Behauptung. □

Bemerkung 5.5.1. Das **Taylor-Polynom m-ter Ordnung** für $f \in C^m$,

$$T_m f(x;a) = f(a) + f'(a)(x-a) + \cdots + f^{(m)}(a)\frac{(x-a)^m}{m!},$$

hat nach Satz 5.5.1 die folgende Approximationseigenschaft. Für $a < x < b$ gilt

$$f(x) - T_m f(x;a) = \left(f^{(m)}(\xi) - f^{(m)}(a)\right) \cdot \frac{(x-a)^m}{m!} =: r_m f(x;a)$$

für ein $\xi \in]a, x[$. Für den **Restterm** $r_m f$ erhalten wir die Abschätzung

$$|r_m f(x;a)| \leq \sup_{a<\xi<x} \left|f^{(m)}(\xi) - f^{(m)}(a)\right| \frac{(x-a)^m}{m!}.$$

Falls $f \in C^{m+1}$, können wir dies mit Satz 5.2.1 noch verbessern zu

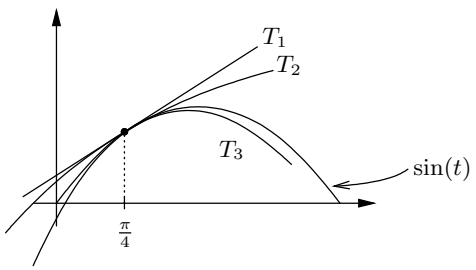
$$|r_m f(x;a)| \leq \sup_{a<\xi<x} \left|f^{(m+1)}(\xi)\right| \frac{(x-a)^{m+1}}{m!}.$$

Beispiel 5.5.1. i) Was ist der Sinus von $47° \hat{=} \frac{\pi}{4} + \frac{2\pi}{180}$? – Mit $\sin' = \cos$, $\sin'' = \cos' = -\sin$, usw., folgt aus Satz 5.5.1 z.B. bei Wahl von $m = 2$ für ein $\xi \in]\frac{\pi}{4}, \frac{\pi}{4} + \frac{\pi}{90}[$

$$\begin{aligned}\sin(\frac{\pi}{4} + \frac{\pi}{90}) &= \sin(\frac{\pi}{4}) + \cos(\frac{\pi}{4}) \cdot \frac{\pi}{90} - \sin(\frac{\pi}{4}) \frac{\pi^2}{2 \cdot 90^2} + r_2 \\ &= \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{\pi}{90} - \frac{\sqrt{2}}{2} \cdot \frac{\pi^2}{2 \cdot 90^2} + r_2,\end{aligned}$$

wobei

$$|r_2| \leq \frac{\pi^3}{2!\ 90^3} \approx 10^{-5}.$$



ii) Sei $p(x) = x^4 - x^2 + 1$. Mit Satz 5.5.1 können wir p im Punkt $x_0 = 1$

annähern durch

$$T_1p(x;1) = p(1) + \underbrace{p'(1)(x-1)}_{=(4x^3-2x)\big|_{x_0=1}(x-1)=2(x-1)} = 1 + 2(x-1) = 2x-1,$$

$$T_2p(x;1) = T_1p(x;1) + \underbrace{p''(1)\frac{(x-1)^2}{2}}_{=(12x^2-2)\big|_{x_0=1}\cdot\frac{x^2-2x+1}{2}=5(x^2-2x+1)} = 5x^2-8x+4,$$

$$T_3p(x;1) = T_2p(x;1) + \underbrace{p'''(1)\frac{(x-1)^3}{6}}_{=(24x)\big|_{x=1}\frac{x^3-3x^2+3x-1}{6}=4(x^3-3x^2+3x-1)} = 4x^3-7x^2+4x,$$

$$T_4p(x;1) = T_3p(x;1) + \underbrace{p^{(iv)}(1)}_{24}\frac{(x-1)^4}{24} = x^4-x^2+1 = p(x).$$

Lokale Extrema. Sei $\Omega \subset \mathbb{R}$ offen, $f\colon \Omega \to \mathbb{R}$.

Definition 5.1.1. *Ein $x_0 \in \Omega$ heisst (strikte)* **lokale Minimalstelle** *von f, falls in einer Umgebung U von x_0 gilt*

$$f(x) \geq f(x_0), \quad \forall x \in U \quad (\textit{bzw. } f(x) > f(x_0), \quad \forall x \in U\backslash\{x_0\}).$$

Falls f an einer lokalen Minimalstelle x_0 differenzierbar ist, so folgt wie im Beweis von Satz 5.2.1

$$0 \leq \lim_{x\downarrow x_0}\frac{f(x)-f(x_0)}{x-x_0} = f'(x_0) = \lim_{x\uparrow x_0}\frac{f(x)-f(x_0)}{x-x_0} \leq 0;$$

also $f'(x_0) = 0$. Allgemein gilt der folgende Satz.

Korollar 5.5.1. *Sei $f \in C^m(\Omega)$, $x_0 \in \Omega$ mit $f'(x_0) = \cdots = f^{(m-1)}(x_0) = 0$.*

i) *Falls $m = 2k+1$, x_0 lokale Minimalstelle, so folgt $f^{(m)}(x_0) = 0$.*

ii) *Falls $m = 2k$, und falls $f^{(m)}(x_0) > 0$, so ist x_0 strikte lokale Minimalstelle.*

Beweis. Nach Satz 5.5.1 existiert für $x \in \Omega$ ein ξ zwischen x und x_0 mit

$$f(x) = f(x_0) + f^{(m)}(\xi)\frac{(x-x_0)^m}{m!}.$$

i) Falls $m = 2k+1$, und falls x_0 lokales Minimum, so folgt

$$f^{(m)}(x_0) = \lim_{\xi\to x_0} f^{(m)}(\xi) = \begin{cases} \lim\limits_{x\downarrow x_0}\frac{f(x)-f(x_0)}{(x-x_0)^m}\cdot m! & \geq 0, \\ \lim\limits_{x\uparrow x_0}\frac{f(x)-f(x_0)}{(x-x_0)^m}\cdot m! & \leq 0; \end{cases}$$

also $f^{(m)}(x_0) = 0$.

ii) Falls $f^{(m)}(x_0) = \lim\limits_{\xi\to x_0} f^{(m)}(\xi) > 0$, $m = 2k$, so folgt für x nahe x_0, $x \neq x_0$, die Ungleichung $f(x) - f(x_0) > 0$; also ist x_0 ein striktes lokales Minimum. $\square$

Beispiel 5.5.2. i) Sei $f(x) = x^4 - x^2 + 1,\ x \in \mathbb{R}$. Nach Korollar 5.5.1.i) ist notwendig für das Vorliegen einer Extremalstelle im Punkt x_0 die Bedingung

$$f'(x_0) = 4x_0^3 - 2x_0 = 2(2x_0^2 - 1)x_0 = 0;$$

d.h.

$$x_0 \in \{-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\}.$$

Nach Korollar 5.5.1.ii) und mit

$$f''(x) = 12x^2 - 2 = \begin{cases} 4 > 0, & x = \pm 1/\sqrt{2} \\ -2 < 0, & x = 0 \end{cases}$$

liegt in $x_0 = 1/\sqrt{2}$, $x_0 = -1/\sqrt{2}$ jeweils ein striktes lokales Minimum, in $x_0 = 0$ ein striktes lokales Maximum vor.

ii) (Minimierungseigenschaft des arithmetischen Mittels) Seien $a_1, \ldots, a_n \in \mathbb{R}$. Gesucht ist $x_0 \in \mathbb{R}$ mit

$$f(x_0) = \sum_{k=1}^{n} (x_0 - a_k)^2 = \min_x f(x).$$

Beachte $f(x) \to \infty$ ($|x| \to \infty$); also existiert $x_0 \in \mathbb{R}$ mit $f(x_0) = \min f$.

Korollar 5.5.1.i) liefert die notwendige Bedingung:

$$f'(x_0) = 2\sum_{k=1}^{n} (x_0 - a_k) = 2nx_0 - 2\sum_{k=1}^{n} a_k = 0;$$

d.h. das arithmetische Mittel

$$x_0 = \frac{1}{n}\sum_{k=1}^{n} a_k$$

ist die einzig mögliche Minimalstelle. Zur Probe bestimmen wir noch

$$f'' \equiv 2n > 0.$$

Der Punkt x_0 ist also tatsächlich die gesuchte Minimalstelle.

Konvexe Funktionen. Sei $-\infty < a < b < \infty$.

Satz 5.5.2. *Sei $f \in C^2(]a,b[)$ mit $f'' \geq 0$. Dann gilt für alle $x_0, x_1 \in]a,b[$, $0 \leq t \leq 1$:*

$$f(tx_1 + (1-t)x_0) \leq tf(x_1) + (1-t)f(x_0). \tag{5.5.2}$$

Beweis. Fixiere $x_0, x_1 \in]a,b[$. Betrachte die Hilfsfunktion $g \in C^2([0,1])$:

$$g(t) = f(tx_1 + (1-t)x_0) - (tf(x_1) + (1-t)f(x_0))$$

mit

$$g(0) = g(1) = 0,\ g''(t) = f''(tx_1 + (1-t)x_0)(x_1 - x_0)^2 \geq 0,\ 0 \leq t \leq 1.$$

Nimm widerspruchsweise an,

$$\max_{0\le t\le 1} g(t) = g(t_{max}) > 0, \text{ wobei } 0 < t_{max} < 1.$$

Gemäss Korollar 5.5.1.i) folgt $g'(t_{max}) = 0$. Nach Satz 5.5.1 gibt es $\tau \in]t_{max}, 1[$ mit

$$0 = g(1) = g(t_{max}) + g'(t_{max})(1 - t_{max}) + g''(\tau)\frac{(1-t_{max})^2}{2} \ge g(t_{max}) > 0,$$

und es folgt der gewünschte Widerspruch. □

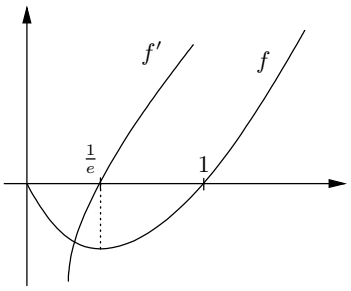
Definition 5.5.2. *Eine Funktion $f\colon]a,b[\to \mathbb{R}$ mit der Eigenschaft (5.5.2) heisst* **konvex**.

Beispiel 5.5.3. i) $\exp'' = \exp > 0$; also ist die Funktion exp konvex.

ii) Sei $f(x) = x\log x,\ x > 0$. Es gilt

$$f'(x) = \log x + 1,\ f''(x) = \frac{1}{x} > 0;$$

also ist f konvex gemäss Satz 5.5.2.



iii) Sei $f(x) = x^\alpha = \exp(\alpha\log x),\ x > 0$, wobei $\alpha > 1$ fest. Da

$$f'(x) = \alpha x^{\alpha-1},\ f''(x) = \alpha(\alpha-1)x^{\alpha-2} > 0,$$

ist f konvex gemäss Satz 5.5.2.

Die Eigenschaft (5.5.2) gilt analog auch für mehr als zwei Punkte.

Satz 5.5.3. (Jensen) *Sei $f\colon]a,b[\to \mathbb{R}$ konvex. Dann gilt für beliebige Punkte $x_1,\dots,x_N \in]a,b[$ und Zahlen $0 \le t_1,\dots,t_N \le 1$ mit $\sum_{i=1}^N t_i = 1$ die Ungleichung*

$$f\Big(\sum_{i=1}^N t_i x_i\Big) \le \sum_{i=1}^N t_i f(x_i).$$

Beweis (Induktion nach N). $N = 1$ ist klar; ebenso $N = 2$ nach Definition.

$N \to N+1$: OBdA sei $t_1 < 1$. (Sonst sind wir im Fall $N = 1$.) Setze

$$x_0 = \sum_{i=2}^{N+1} \frac{t_i}{1-t_1} x_i.$$

Da f konvex, erhalten wir

$$f\Big(\sum_{i=1}^{N+1} t_i x_i\Big) = f(t_1 x_1 + (1-t_1)x_0) \le t_1 f(x_1) + (1-t_1) f(x_0)$$

Da nach Induktions-Annahme gilt

$$f(x_0) = f\Big(\sum_{i=2}^{N+1} \frac{t_i}{1-t_1} x_i\Big) \le \sum_{i=2}^{N+1} \frac{t_i}{1-t_1} f(x_i),$$

folgt die Behauptung. □

Beispiel 5.5.4. (Vergleich von arithmetischem und geometrischem Mittel) Für alle $0 < x_1, \ldots, x_n < \infty$, $0 \le \alpha_1, \ldots, \alpha_n \le 1$ mit $\sum_{i=1}^{n} \alpha_i = 1$ gilt

$$\prod_{i=1}^{n} x_i^{\alpha_i} \le \sum_{i=1}^{n} \alpha_i x_i.$$

Beweis. Da die Funktion exp nach Beispiel 5.5.3.i) konvex ist, folgt die Aussage mit Satz 5.5.3 aus der Darstellung

$$\begin{aligned} \prod_{i=1}^{n} x_i^{\alpha_i} = \prod_{i=1}^{n} \exp(\alpha_i \log x_i) &= \exp\Big(\sum_{i=1}^{n} \alpha_i \log x_i\Big) \\ &\overset{\text{(Satz 5.5.3)}}{\le} \sum_{i=1}^{n} \alpha_i \exp(\log x_i) = \sum_{i=1}^{n} \alpha_i x_i. \end{aligned}$$

□

Insbesondere erhalten wir für $\alpha_i = \frac{1}{n}$, $1 \le i \le n$,

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \le \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Schliesslich können wir mit den Ideen aus Satz 5.5.1 einen Beweis des Fundamentalsatzes der Algebra gemäss Abschnitt 2.5 gewinnen.

Satz 5.5.4. *Jedes Polynom $p\colon \mathbb{C} \to \mathbb{C}$ vom Grad ≥ 1 hat (mindestens) eine Nullstelle.*

Beweis(indirekt). Sei $p(z) = a_n z^n + \cdots + a_0$ mit $a_n \ne 0$ ein Polynom vom Grad $n \ge 1$ ohne Nullstelle. OBdA sei $a_n = 1$. (Sonst betrachte $\tilde{p} = \frac{p}{a_n}$.) Sei

$$\mu = \inf_{z \in \mathbb{C}} |p(z)| \ge 0.$$

Da wir für genügend grosses $r_0 > 0$ für $|z| \geq r_0$ abschätzen können

$$|p(z)| = |z|^n \left(1 + a_{n-1} z^{-1} + \cdots + a_0 z^{-n}\right) \geq \frac{1}{2} |z|^n ,$$

können wir einen Radius $r_0 > 0$ wählen mit

$$|p(z)| \geq 1 + \mu$$

für $|z| \geq r_0$. Nach Satz 4.2.3 gibt es $z_0 \in B_{r_0}(0)$ mit

$$0 < |p(z_0)| = \min_{|z| \leq r_0} |p(z)| = \inf_{z \in \mathbb{C}} |p(z)| = \mu.$$

Entwickle p um z_0 analog zu Satz 5.5.1. (Wir machen keine Taylor-Näherung, verfahren jedoch in ähnlicher Weise.) Dies ergibt die Darstellung

$$p(z) = p((z - z_0) + z_0) = \sum_{k=0}^{n} b_k (z - z_0)^k,$$

von p als Polynom in $(z - z_0)$, wobei Einsetzen von $z = z_0$ die Identität $b_0 = p(z_0)$ ergibt.

Setze

$$k_0 = \min\{k;\ b_k \neq 0,\ k \geq 1\}.$$

Da $p \neq const$, ist $1 \leq k_0 \leq n$ wohldefiniert. Schreibe

$$p(z) = p(z_0) + b_{k_0} (z - z_0)^{k_0} + r_{k_0}(z; z_0)$$

mit dem Restterm

$$r_{k_0}(z; z_0) = \sum_{k=k_0+1}^{n} b_k (z - z_0)^k,$$

so dass

$$r_{k_0}(z; z_0)/(z - z_0)^{k_0} \to 0 \quad (z \to z_0).$$

Schliesslich wähle $s_0 > 0$, $0 \leq \varphi_0 < 2\pi$ mit

$$-\frac{p(z_0)}{b_{k_0}} = s_0 e^{i\varphi_0}.$$

Für $z = z_0 + s e^{i\varphi_0/k_0}$, $s > 0$ folgt

$$|p(z)| = \left|p(z_0) + b_{k_0} s^{k_0} e^{i\varphi_0} + r_{k_0}(z; z_0)\right| = |p(z_0)| \left|1 - s^{k_0} \Big(\frac{1}{s_0} - \frac{r_{k_0}(z; z_0)}{p(z_0) s^{k_0}}\Big)\right|.$$

Da $r_{k_0}(z; z_0)/s^{k_0} \to 0$ $(s \to 0)$, folgt der gewünschte Widerspruch

$$|p(z)| < |p(z_0)| = \mu, \text{ falls } 0 < s << 1.$$

□

5.6 Gewöhnliche Differentialgleichungen

Differentialgleichungen erscheinen in vielfältiger Weise in der physikalischen Beschreibung der Natur und in technischen Anwendungen, deren Eigenheiten sich auch in den Eigenschaften gewisser "natürlich" vorkommender Funktionen spiegeln.

Beispiel 5.6.1. i) Die Funktionen exp, sin, cos, tan stehen mit ihrer Ableitung in Beziehung:

$$\exp' = \exp,\ \sin'' = -\sin,\ \cos'' = -\cos,\ \tan' = 1 + \tan^2 .$$

Solche Beziehungen bezeichnet man allgemein als **Differentialgleichungen**. Die Funktion $f(x) = \tan x$ löst also für $-\pi/2 < x < \pi/2$ die Differentialgleichung

$$f' = 1 + f^2.$$

ii) Physikalische Prozesse lassen sich oft durch Differentialgleichungen beschreiben. Z.B. ist beim **radioaktiven Zerfall** die pro Zeitenheit zerfallende Masse proportional zur noch vorhandenen Masse $f(t)$ eines Stoffes; d.h. mit einer Zahl $\alpha > 0$ gilt

$$\dot{f} = \frac{df}{dt} = -\alpha f, \quad f(0) = f_0. \tag{5.6.1}$$

Gemäss Beispiel 5.2.1.i) ist die Lösung dieser Differentialgleichung stets von der Form

$$f(t) = f_0 e^{-\alpha t}, \quad t > 0,$$

wobei die Konstante $f(0)$ durch die Anfangsbedingung festgelegt ist.

iii) Die Auslenkung $f(t)$ eines **Federpendels** aus der Ruhelage $f = 0$ erfüllt nach dem Newton'schen und Hooke'schen Gesetz die Gleichung

$$\frac{d(m\dot{f})}{dt} = m\ddot{f} = -Kf,$$

wobei $m > 0$ die Masse des Pendels und $K > 0$ die Federkonstante bezeichnen; d.h. wir haben die Gleichung

$$\ddot{f} + \omega_0^2 f = 0, \quad \omega_0^2 = \frac{K}{m} > 0. \tag{5.6.2}$$

Nach i) sind die Funktionen

$$f(t) = a\cos(\omega_0 t) + b\sin(\omega_0 t)$$

für beliebige $a, b \in \mathbb{R}$ Lösungen von (5.6.2). Sind alle Lösungen von (5.6.2) von dieser Form?

iv) Beim mehrstufigen radioaktiven Zerfall eines Substanz s_1 in die stabile Substanz s_n über Zwischenstufen $s_2, \ldots, s_{n-1}$ mit Massen $f_i(t)$ und Zerfallsraten

$\alpha_i > 0$ erhalten wir das **System** von Differentialgleichungen

$$\begin{aligned} \dot{f}_1 &= -\alpha_1 f_1, \\ \dot{f}_2 &= \alpha_1 f_1 - \alpha_2 f_2, \\ &\vdots \\ \dot{f}_n &= \alpha_{n-1} f_{n-1}. \end{aligned} \tag{5.6.3}$$

Mit der Notation

$$F = F(t) = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R} \to \mathbb{R}^n, \quad A = \begin{pmatrix} -\alpha_1 & & \\ \alpha_1 & \ddots & 0 \\ & \ddots & -\alpha_{n-1} \\ 0 & & \alpha_{n-1} \quad 0 \end{pmatrix}$$

können wir (5.6.3) in der Form schreiben

$$\dot{F} = AF. \tag{5.6.4}$$

analog zu (5.6.1).

v) Führen wir im Fall (5.6.2) die Funktion

$$F = \begin{pmatrix} f \\ \dot{f} \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2$$

ein, so lässt sich auch diese Gleichung in der Form (5.6.4) schreiben mit

$$\dot{F} = \begin{pmatrix} \dot{f} \\ \ddot{f} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix} \cdot \begin{pmatrix} f \\ \dot{f} \end{pmatrix} = AF,$$

wo

$$A = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix}.$$

Definition 5.6.1. *Die Gleichung (5.6.4) ist die Standardform eines* **homogenen Systems linearer Differentialgleichungen** *mit konstanten Koeffizienten.*

Satz 5.6.1. (Existenz- und Eindeutigkeitssatz für (5.6.4)) *Sei A eine $n \times n$-Matrix mit Koeffizienten in $\mathbb{R}$ (oder $\mathbb{C}$), $F_0 \in \mathbb{R}^n$ (oder $\mathbb{C}^n$). Dann besitzt das* **Anfangswertproblem**

$$\frac{dF}{dt} = AF, \quad F(0) = F_0 \tag{5.6.5}$$

genau eine Lösung $F \in C^1(\mathbb{R}; \mathbb{R}^n)$ (bzw. $C^1(\mathbb{R}; \mathbb{C}^n)$).

Beispiel 5.6.2. Insbesondere ist die Funktion $F(t) = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}$ die einzige Lösung des Anfangswertproblems

$$\dot{F} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} F, \quad F(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{5.6.6}$$

Somit ist auch $f(t) = \cos t$ die einzige Lösung des Anfangswertproblems

$$\ddot{f} + f = 0, \quad f(0) = 1, \ \dot{f}(0) = 0, \tag{5.6.7}$$

denn jede Lösung f von (5.6.7) induziert eine Lösung $F = \begin{pmatrix} f \\ \dot{f} \end{pmatrix}$ von (5.6.6).

Beweis von Satz 5.6.1. i) (Eindeutigkeit.) Es seien $F, G \in C^1(\mathbb{R}; \mathbb{R}^n)$ Lösungen von (5.6.5). Dann löst die Funktion

$$H = F - G \in C^1(\mathbb{R}; \mathbb{R}^n)$$

die Gleichung

$$\frac{dH}{dt} = \frac{dF}{dt} - \frac{dG}{dt} = AF - AG = AH$$

mit Anfangswert

$$H(0) = 0.$$

Betrachte die Hilfsfunktion η mit

$$\eta(t) = |H(t)|^2 = \sum_{i=1}^{n} |H_i(t)|^2, \quad t \in \mathbb{R}.$$

Mit der Cauchy-Schwarzschen Ungleichung gemäss Satz 2.4.1 folgt

$$\begin{aligned} \frac{d\eta}{dt} &= 2 \left\langle H, \frac{dH}{dt} \right\rangle = 2 \langle H, AH \rangle \leq 2 |H| \cdot |AH| \\ &\leq C_1 |H|^2 = C_1 \eta, \quad \eta(0) = 0. \end{aligned}$$

Somit gilt

$$\frac{d}{dt}\big(e^{-C_1 t}\eta(t)\big) = \Big(\frac{d\eta}{dt} - C_1\eta\Big)e^{-C_1 t} \leq 0,$$

und Korollar 5.2.1.ii) ergibt

$$e^{-C_1 t}\eta(t) \leq e^{-C_1 \cdot 0}\eta(0) = 0, \quad \forall t \geq 0.$$

D.h. $\eta(t) = 0$ und somit auch $H(t) = 0$, $\forall t \geq 0$; analog für $t \leq 0$.

ii) (Existenz.) Analog zu Beispiel 5.6.1.ii) machen wir für die Lösung von (5.6.5) den Ansatz

$$F(t) = Exp(At)F_0, \quad t \in \mathbb{R},$$

wobei die Reihe (mit $A^0 = id$)

$$Exp(At) = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

in jeder Matrix-Norm analog zu Beispiel 3.7.2 für beliebige $t \in \mathbb{R}$ konvergiert. Weiter gilt analog zu Beispiel 5.4.3.i), dass $Exp(At) \in C^1(\mathbb{R}; \mathbb{R}^{n\times n})$ mit

$$\frac{d}{dt}Exp(At) = \sum_{k=0}^{\infty} \frac{d}{dt}\Big(\frac{A^k t^k}{k!}\Big) = AExp(At);$$

d.h.

$$\frac{dF}{dt} = AF, \quad F(0) = \underbrace{Exp(A \cdot 0)}_{=id} F_0 = F_0,$$

wie gewünscht. □

Definition 5.6.2. *Die Matrix-wertige Funktion*

$$t \mapsto \Phi(t) = Exp(At) \quad mit \quad \frac{d\Phi}{dt} = A\Phi,\ \Phi(0) = id$$

heisst **Fundamentallösung** *von (5.6.4) oder (5.6.5).*

Bemerkung 5.6.1. i) Für invertierbares A ist die Definition $A^0 = id$ offenbar sinnvoll. Eine beliebige Matrix A kann man durch invertierbare Matrizen approximieren, was die Vereinbarung $A^0 = id$ für beliebige Matrizen rechtfertigt.

ii) Ohne Vorbereitung kann man $\Phi(t) = Exp(At)$ nur mit Mühe berechnen. Falls man jedoch durch eine lineare Transformation $T\colon \mathbb{R}^n \to \mathbb{R}^n$ die Matrix A in Diagonalform

$$TAT^{-1} = \begin{pmatrix} \lambda_1 & & \\ 0 & \ddots & 0 \\ & & \lambda_n \end{pmatrix} =: \Lambda$$

bringen kann, so lässt sich diese Rechnung wesentlich vereinfachen. Es gilt nämlich

$$(TAT^{-1})^k = TAT^{-1}TAT^{-1}\dots TAT^{-1} = TA^kT^{-1}, \quad k \in \mathbb{N}_0,$$

also auch

$$T \cdot Exp(At)T^{-1} = Exp(TAT^{-1}t) = Exp(\Lambda t) = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix},$$

und wir erhalten

$$\Phi(t) = T^{-1} \begin{pmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{pmatrix} T.$$

iii) Falls insbesondere (5.6.4) die äquivalente Form ist für eine skalare Differentialgleichung n-ter Ordnung

$$f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0 f = 0 \tag{5.6.8}$$

mit

$$F = \begin{pmatrix} f \\ \dot{f} \\ \vdots \\ f^{(n-1)} \end{pmatrix} \in C^1(\mathbb{R}, \mathbb{R}^n),$$

so erwarten wir, dass f eine Linearkombination von Funktionen der Form $e^{\lambda_i t}$ ist. Der **Exponentialansatz** $f(t) = e^{\lambda t}$ für eine Lösung von (5.6.8) führt auf die Gleichung

$$(\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0)e^{\lambda t} = 0;$$

d.h. die Koeffizienten λ_i sind genau die Nullstellen des **charakteristischen Polynoms**

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$$

von (5.6.8).

Beispiel 5.6.3. i) Betrachte die Gleichung

$$f^{(4)} - 3f^{(2)} + 2f = 0. \tag{5.6.9}$$

Das charakteristische Polynom lautet

$$p(\lambda) = \lambda^4 - 3\lambda^2 + 2 = (\lambda^2 - 1)(\lambda^2 - 2).$$

Es hat die Nullstellen

$$\lambda_{1,2} = \pm 1, \quad \lambda_{3,4} = \pm\sqrt{2}.$$

Folglich lässt die allgemeine Lösung von (5.6.9) sich in der Form darstellen

$$f(t) = ae^t + be^{-t} + ce^{\sqrt{2}t} + de^{-\sqrt{2}t}.$$

Durch Vorgabe von $F(0) = \big(f(0), \ldots, f^{(3)}(0)\big)^t$ ist f gemäss Satz 5.6.2 eindeutig bestimmt, und man kann die Konstanten $a, \ldots, d$ aus den vorgegebenen Werten für $f(0), \ldots f^{(3)}(0)$ bestimmen.

ii) Die Differentialgleichung (5.6.2) hat das charakteristische Polynom

$$p(\lambda) = \lambda^2 + \omega_0^2$$

mit den Nullstellen $\lambda_{1,2} = \pm i\omega_0$. Die allgemeine Lösung $f \in C^1(\mathbb{R}; \mathbb{C})$ hat die Gestalt

$$f(t) = ae^{i\omega_0 t} + be^{-i\omega_0 t}, \quad t \in \mathbb{R},$$

wobei $a, b \in \mathbb{C}$ durch die Anfangsbedingungen

$$f(0) = f_0, \quad \dot{f}(0) = f_1$$

bestimmt sind. Da die Koeffizienten von (5.6.2) reell sind, sind mit $f \in C^1(\mathbb{R}; \mathbb{C})$ auch die Funktionen $Re(f)$, $Im(f) \in C^1(\mathbb{R}; \mathbb{R})$ Lösungen; d.h. die allgemeine reelle Lösung $f \in C^1(\mathbb{R}; \mathbb{R})$ hat die Gestalt

$$f(t) = a\cos(\omega_0 t) + b\sin(\omega_0 t), \quad t \in \mathbb{R}, \tag{5.6.10}$$

wobei $a, b \in \mathbb{R}$. Sind dies alle Lösungen? Was kann man aussagen im Fall von mehrfachen Nullstellen $\lambda_1 = \cdots = \lambda_k$ von p?

Satz 5.6.2. i) *Der Lösungsraum*

$$X = \{F \in C^1(\mathbb{R}; \mathbb{R}^n);\ \frac{dF}{dt} = AF\}$$

von (5.6.4) für eine reelle $n \times n$-Matrix A ist ein n-dimensionaler $\mathbb{R}$-Vektorraum.

ii) *Analog ist für eine $n \times n$-Matrix A mit Koeffizienten in $\mathbb{C}$ der Lösungsraum*

$$\tilde{X} = \{F \in C^1(\mathbb{R}, \mathbb{C}^n);\ \frac{dF}{dt} = AF\}$$

von (5.6.4) ein n-dimensionaler $\mathbb{C}$-Vektorraum.

Beweis. Dem Beweis liegt die einfache Idee zugrunde, dass die Abbildung $F_0 \mapsto \Phi(t)F_0 \in X$ eine injektive lineare Abbildung von $\mathbb{R}^n$ in den Raum $C^1(\mathbb{R};\mathbb{R}^n)$ ist, ein "linearer Isomorphismus", dessen Bild wiederum ein Vektorraum der Dimension n ist.

Seien $F_1,\ F_2 \in X$, $a_1, a_2 \in \mathbb{R}$. Dann gilt das "Superpositionsprinzip"

$$\frac{d(a_1F_1 + a_2F_2)}{dt} = a_1AF_1 + a_2AF_2 = A(a_1F_1 + a_2F_2);$$

d.h. $F = a_1F_1 + a_2F_2 \in X$, und X ist ein Vektorraum.

Die Lösungen $F_i \in C^1(\mathbb{R},\mathbb{R}^n)$ von (5.6.4) mit $F_i(0) = e_i$, $1 \le i \le n$, sind linear unabhängig, da die Vektoren $F_i(0)$ es sind; also gilt $dim_{\mathbb{R}}X \ge n$. Sind andererseits $F_1, \dots, F_{n+1} \in X$, so gibt es $(a_i)_{1\le i\le n+1} \in \mathbb{R}^{n+1}\backslash\{0\}$ mit

$$F_0 = \sum_{i=1}^{n+1} a_iF_i(0) = 0 \in \mathbb{R}^n.$$

Die Funktion $F = \sum\limits_{i=1}^{n+1} a_iF_i \in X$ löst somit das Anfangswertproblem (5.6.5) mit $F_0 = 0$. Mit Satz 5.6.1 folgt $F = 0$; d.h. $dim_{\mathbb{R}}X \le n$. Analog in $\mathbb{C}$. □

Korollar 5.6.1. *Seien $a_0, \dots, a_{n-1} \in \mathbb{R}$ (oder $\mathbb{C}$). Der Lösungsraum von (5.6.8)*

$$Z = \{f \in C^n(\mathbb{R});\ f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_0f = 0\}$$

ist ein n-dimensonaler $\mathbb{R}$-Vektorraum (oder $\mathbb{C}$-VR).

Beweis. Sei (5.6.4) das zu (5.6.8) äquivalente System erster Ordnung, X der zugehörige Lösungsraum nach Satz 5.6.2. Die Abbildung

$$Z \ni f \mapsto F = \left(f, f', \dots, f^{(n-1)}\right)^t \in X$$

ist ein linerarer Isomorphismus. □

Beispiel 5.6.4. Der $\mathbb{R}$-Lösungsraum von (5.6.2) ist 2-dimensional; die Lösungen $f_1(t) = \cos(\omega_0 t)$, $f_2(t) = \sin(\omega_0 t)$ sind linear unabhängig; also ist jede Lösung von (5.6.2) von der Form (5.6.10).

Mehrfache Nullstellen Seien $a_0, \dots, a_{n-1} \in \mathbb{R}$ oder $\mathbb{C}$. Mit $D = \frac{d}{dt}$ gilt

$$f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_0f = p(D)f, \tag{5.6.8}$$

wobei

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0$$

das charakteristische Polynom von (5.6.8)). Gemäss Satz 5.5.4 gilt

$$p(\lambda) = \prod_{i=1}^{l} (\lambda - \lambda_i)^{m_i},$$

wobei $\lambda_1, \ldots, \lambda_l \in \mathbb{C}$ die paarweise verschiedenen Nullstellen von p bezeichnen und $m_1, \ldots, m_l \in \mathbb{N}$ deren Vielfachheit; d.h.

$$p(D) = \prod_{i=1}^{l} (D - \lambda_i \cdot id)^{m_i}.$$

Beispiel 5.6.5. Die Gleichung

$$\ddot{f} - 2\dot{f} + f = 0 \tag{5.6.11}$$

hat das charakteristische Polynom

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

Es gilt

$$p(D)f = (D - id)^2 f = (D - id)(\dot{f} - f) = \ddot{f} - 2\dot{f} + f.$$

Satz 5.6.3. *Sei* $p(\lambda) = \prod_{i=1}^{l} (\lambda - \lambda_i)^{m_i}$ *mit* $\lambda_i \neq \lambda_j$ $(i \neq j)$. *Dann ist jede Lösung der zugehörigen Differentialgleichung* (5.6.8) *darstellbar als Linearkombination der* n *linear unabhängigen Funktionen*

$$f_{ik}(t) = t^k e^{\lambda_i t}, \quad 1 \leq i \leq l, \ 0 \leq k < m_i.$$

Beispiel 5.6.6. i) Jede Lösung f der Differentialgleichung (5.6.11) ist also von der Form

$$f(t) = (a + bt)e^t, \quad t \in \mathbb{R},$$

wobei $a, b \in \mathbb{R}$.

ii) Allgemein haben wir folgende Fälle für die Differentialgleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 0 \tag{5.6.12}$$

mit dem charakteristischen Polynom

$$p(\lambda) = \lambda^2 + 2\delta\lambda + \omega_0^2$$

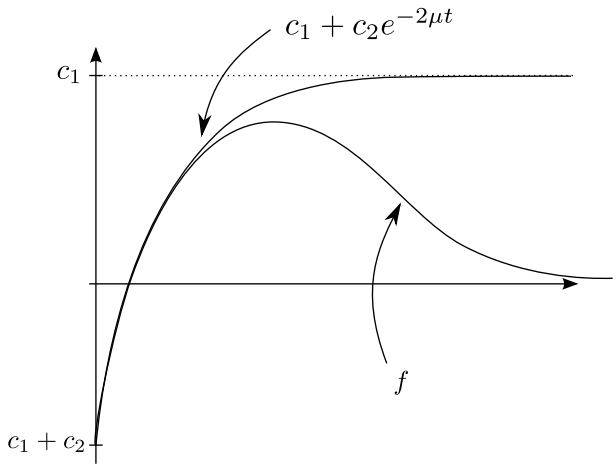
und den (im allgemeinen komplexen) Nullstellen

$$\lambda_{1,2} = -\delta \pm \sqrt{\delta^2 - \omega_0^2} = -\delta \pm i\sqrt{\omega_0^2 - \delta^2}.$$

a) $\delta^2 > \omega_0^2$ ("superkritische Dämpfung") Die allgemeine Lösung von (5.6.12) hat die Form

$$f(t) = c_1 e^{(\mu - \delta)t} + c_2 e^{-(\mu + \delta)t} = e^{(\mu - \delta)t}(c_1 + c_2 e^{-2\mu t}),$$

wobei $\mu = \sqrt{\delta^2 - \omega_0^2} < \delta$, $c_1, c_2 \in \mathbb{R}$.



b) $\delta^2 = \omega_0^2$ ("kritische Dämpfung") Die allgemeine Lösung ist

$$f(t) = c_1e^{-\delta t} + c_2te^{-\delta t} = (c_1 + c_2t)e^{-i\delta t}.$$

Die Lösungen in den Fällen a) und b) sind also stets exponentiell abfallend mit höchstens ener Nullstelle (Stossdämpfertest).

c) $\delta^2 < \omega_0^2$ ("sub-kritische Dämpfung") Schreibe $\lambda_{1,2} = -\delta \pm i\mu$, wobei wir diesmal $\mu = \sqrt{\omega_0^2 - \delta^2} > 0$ setzen. Die allgemeine Lösung von (5.6.12) in $\mathbb{C}$ lautet

$$f(t) = e^{-\delta t}\Big(c_1e^{i\mu t} + c_2e^{-i\mu t}\Big),$$

wobei $c_{1,2} \in C$. Falls $f = \overline{f}$ reell, so folgt

$$c_1e^{i\mu t} + c_2e^{-i\mu t} = \overline{c_1}e^{-i\mu t} + \overline{c_2}e^{i\mu t};$$

d.h. $c_1 = \overline{c_2} = a + ib$, $c_2 = a - ib$, und

$$f(t) = e^{-\delta t}(2a\cos(\mu t) - 2b\sin(\mu t))$$

beschreibt eine gedämpfte Schwingung.

Wir kommen nun zum

Beweis von Satz 5.6.3. **i)** Für $q \in C^m(\mathbb{R})$, $\lambda \in \mathbb{C}$ gilt

$$(D - \lambda id)(qe^{\lambda t}) = \dot{q}e^{\lambda t}.$$

Mit Induktion erhalten wir

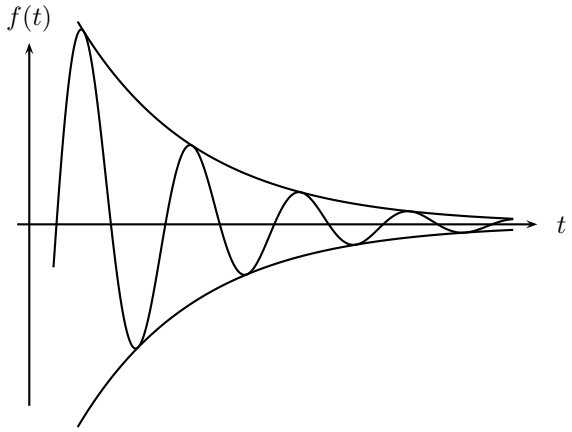
$$(D - \lambda id)^m(qe^{\lambda t}) = q^{(m)}e^{\lambda t},\ m \in \mathbb{N}.$$

Es folgt

$$(D - \lambda_i id)^{m_i}(t^ke^{\lambda_i t}) = \frac{d^{m_i}t^k}{dt^{m_i}}e^{\lambda_i t} = 0$$

falls $k < m_i$, und somit

$$p(D)f_{ik} = \prod_{j\neq i}(D - \lambda_j\ id)^{m_j}\Big((D - \lambda_i)^{m_i}f_{ik}\Big) = 0,\quad 1 \le i \le l,\ 0 \le k < m_i.$$



Beachte, dass die Reihenfolge der Operatoren $(D-\lambda_i)$ und $(D-\lambda_j)$ bei dieser Rechnung offenbar keine Rolle spielt.

ii) Zum Beweis der linearen Unabhängigkeit der Funktionen f_{ik} seien $b_{ik} \in \mathbb{C}$ mit $\sum\limits_{i,k} b_{ik} f_{ik} = 0$, wobei $b_{i_0 k_0} \neq 0$ für (mindestens) ein Paar von Indices i_0, k_0. OBdA sei k_0 so gewählt, dass $k = k_0$ maximal ist mit der Eigenschaft $b_{i_0 k} \neq 0$. Gemäss Teil i) gelten die Gleichungen

$$\prod_{i \neq i_0} (D-\lambda_i)^{m_i} \sum_{i \neq i_0, k} b_{ik} f_{ik} = 0, \ (D-\lambda_{i_0})^{k_0} \sum_{k<k_0} f_{i_0 k} = 0.$$

Die Rechnung

$$\begin{aligned} 0 &= \prod_{i \neq i_0} (D-\lambda_i)^{m_i} (D-\lambda_{i_0})^{k_0} \sum_{i,k} b_{ik} f_{ik} \\ &= \prod_{i \neq i_0} (D-\lambda_i)^{m_i} (D-\lambda_{i_0})^{k_0} b_{i_0 k_0} f_{i_0 k_0} = \prod_{i \neq i_0} (D-\lambda_i)^{m_i} b_{i_0 k_0} D^{k_0} t^{k_0} e^{\lambda_{i_0} t} \\ &= \prod_{i \neq i_0} (D-\lambda_i)^{m_i} b_{i_0 k_0} k_0! e^{\lambda_{i_0} t} = b_{i_0 k_0} k_0! \prod_{i \neq i_0} (\lambda_{i_0} - \lambda_i)^{m_i} \neq 0, \end{aligned}$$

liefert nun den gewünschten Widerspruch. Die n Funktionen f_{ik} sind also linear unabhängig.

□

5.7 Inhomogene Differentialgleichungen

Bisher haben wir nur homogene lineare Differentialgleichungen betrachtet. Sehr oft treten jedoch auch Zusatzterme in den Gleichungen auf.

Beispiel 5.7.1. i) Ein gedämpftes Federpendel wird mit der periodischen Kraft $b(t) = b_0 \cos(\omega t)$ mit der Frequenz $\omega > 0$ angetrieben. Mit dem Newtonschem

und Hookeschem Gesetz folgt die Gleichung

$$m\ddot{f} = -Kf - d\dot{f} + b.$$

Schreiben wir wieder $\omega_0^2 = \frac{K}{m}$ und setzen wir $2\delta = \frac{d}{m}$, $\beta_0 = \frac{b_0}{m}$ so erhalten wir

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = \beta_0 \cos(\omega t). \tag{5.7.1}$$

Als **partikuläre Lösung** dieser Gleichung erwarten wir eine Schwingung mit derselben Frequenz ω. Am leichtesten gelingt die Rechnung im Komplexen.

Für eine Lösung $f \in C^2(\mathbb{R};\mathbb{C})$ der Gleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = \beta_0 e^{i\omega t} \tag{5.7.2}$$

machen wir den **Ansatz**

$$f(t) = ce^{i\omega t}, \quad t \in \mathbb{R},$$

mit einer beliebigen Konstanten $c \in \mathbb{C}$. Einsetzen in (5.7.2) ergibt

$$\Big(\big((\omega_0^2 - \omega^2) + 2i\delta\omega\big)c - \beta_0\Big)e^{i\omega t} = 0$$

als Bestimmungsgleichung für c. Falls $\omega \neq \omega_0$, oder falls $\delta > 0$, kann diese Gleichung nach c aufgelöst werden, und wir erhalten

$$c = \frac{\beta_0}{(\omega_0^2 - \omega^2) + 2i\delta\omega}.$$

Schreiben wir

$$\frac{1}{\omega_0^2 - \omega^2 + 2i\delta\omega} = \frac{\omega_0^2 - \omega^2 - 2i\delta\omega}{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2} = Re^{i\varphi},$$

so können wir

$$R = \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2}}$$

als "Resonanzamplitude" und

$$\varphi = \arctan\Big(\frac{-2\delta\omega}{\omega_0^2 - \omega^2}\Big) \in]-\pi, 0].$$

als "Phasenverschiebung" gegenüber der von aussen wirkenden Kraft deuten. Schliesslich liefert

$$f_{part}(t) = \beta_0 R e^{i(\omega t + \varphi)}, \quad t \in \mathbb{R},$$

die gesuchte partikuläre Lösung von (5.7.2), bzw.

$$\tilde{f}_{part}(t) = Re(f_{part}(t)) = \beta_0 R \cos(\omega t + \varphi), \quad t \in \mathbb{R}$$

die gesuchte partikuläre Lösung von (5.7.1).

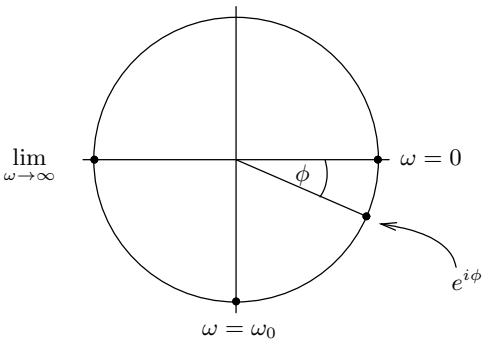


Abbildung 5.1: Phasenwinkel ϕ

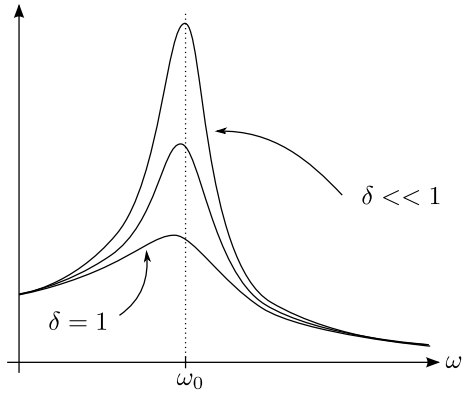


Abbildung 5.2: Resonanzamplitude R

ii) Eine partikuläre Lösung der Gleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 1 \tag{5.7.3}$$

kann man ebenfalls leicht erraten. Die "Kraft" der Grösse 1 suf der rechten Seite von (5.7.3) führt zu einer Verschiebung der Ruhelage des Pendels um $\frac{1}{\omega_0^2}$; d.h. neu entspricht die stationäre (zeitunabhängige) Lösung

$$f_{part}(t) = \frac{1}{\omega_0^2}, \quad t \in \mathbb{R},$$

dem Pendelgleichgewicht.

iii) Kommen beide Effekte aus i) und ii) zusammen, so ergibt dies die Gleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 1 + \beta_0 e^{i\omega t}. \tag{5.7.4}$$

Nach dem **Superpositionsprinzip** ergänzen sich die oben bestimmten partikulären Lösungen von (5.7.2) und (5.7.3) zu einer partikulären Lösung

$$f(t) = ce^{i\omega t} - \frac{1}{\omega_0^2}, \quad t \in \mathbb{R},$$

von (5.7.4), wobei $c \in \mathbb{C}$ wie in i) gewählt wird.

Wie findet man Lösungen zu vorgegebenen Anfangswerten? Wie findet man alle Lösungen? – Die Antworten auf diese Fragen formulieren wir wie vorher im Kontext von Systemen linearer Differentialgleichungen 1. Ordnung

$$\frac{dF}{dt} = AF + B, \tag{5.7.5}$$

wobei $B = B(t) \in C^0(\mathbb{R}; \mathbb{R}^n)$ (oder $\in C^0(\mathbb{R}; \mathbb{C}^n)$).

Satz 5.7.1. *Sei $F_{part} \in C^1(\mathbb{R};\mathbb{R}^n)$ eine beliebige ("partikuläre") Lösung von (5.7.5). Dann ist jede Lösung F von (5.7.5) von der Form*

$$F = F_{part} + F_{hom}, \tag{5.7.6}$$

wobei F_{hom} eine beliebige Lösung der homogenen Gleichung (5.6.4) ist. Insbesondere gibt es zu jedem $F_0 \in \mathbb{R}^n$ stets genau eine Lösung F von (5.7.5) mit $F(0) = F_0$. (Analog in $\mathbb{C}$.)

Beweis. i) Jedes F der Form (5.7.6) löst (5.7.5). Sind umgekehrt $F_1, F_2 \in C^1(\mathbb{R};\mathbb{R}^n)$ Lösungen von (5.7.5), so gilt

$$\frac{d(F_1 - F_2)}{dt} = A(F_1 - F_2) + B - B = A(F_1 - F_2);$$

d.h. jede Lösung von (5.7.5) ist von der Gestalt (5.7.6).

ii) Zu vorgegebenen Anfangswerten $F_0 \in \mathbb{R}^n$ sei F_{hom} die Lösung des Anfangswertproblems (5.6.5) mit $F_{hom}(0) = F_0 - F_{part}(0)$. Dann löst $F = F_{part} + F_{hom}$ das Anfangswertproblem (5.7.5) mit $F(0) = F_0$. Eindeutigkeit der Lösung F folgt mit i) und Satz 5.6.1. □

Beispiel 5.7.2. i) Radioaktiver Zerfall mit konstanter Zufuhr wird beschrieben durch die Modellgleichung

$$\dot{f} = -\alpha f + s, \tag{5.7.7}$$

wobei $\alpha, s > 0$ gegeben sind. Eine partikuläre Lösung ist offenbar die **Gleichgewichtslösung**

$$f_{part}(t) = \frac{s}{\alpha},\ t \in \mathbb{R}.$$

Mit der Darstellung $f_{hom}(t) = ce^{-\alpha t}$ der allgemeinen Lösung der homogenen Gleichung aus Beispiel 5.6.1.ii) erhalten wir die allgemeine Lösung

$$f(t) = ce^{-\alpha t} + \frac{s}{\alpha}$$

von (5.7.7), wobei $c \in \mathbb{R}$ durch $f(0)$ bestimmt wird.

ii) Die (nichtlineare) **logistische Gleichung**

$$y' = ay - by^2$$

für $y = y(t) > 0$ mit $a, b > 0$ kann nach Division durch y^2 auf die Form

$$\left(\frac{1}{y}\right)' = \frac{y'}{y^2} = \frac{a}{y} - b$$

gebracht werden. Für die Funktion

$$f(t) = \frac{1}{y(t)}$$

erhalten wir die Gleichung

$$f' = -af + b.$$

Mit i) folgt die Darstellung

$$f(t) = \frac{b}{a} + ce^{-at}, \quad c = f(0) - \frac{b}{a},$$

also

$$y(t) = \frac{1}{b/a + ce^{-at}} \to \frac{a}{b} \quad (t \to \infty).$$

Ein analoges Resultat gilt für allgemeine Differentialgleichungen vom "Bernoulli-Typ"

$$y' = ay - by^{\alpha}, \ \alpha > 1.$$

Wie findet man im allgemeinen Fall eine partikuläre Lösung zu (5.7.5)? Bereits im Fall $A = 0$ müssten wir in der Lage sein, zu vorgegebenem $B = B(t) \in C^0(\mathbb{R}; \mathbb{R}^n)$ die Gleichung $\frac{dF}{dt} = B$ zu lösen. Genau dies leistet die **Integration**, der wir uns nun zuwenden wollen.

Kapitel 6

Integration

6.1 Stammfunktionen

Seien $-\infty < a < b < \infty$, $f \in C^0(]a,b[)$.

Definition 6.1.1. *Ein* $F \in C^1(]a,b[)$ *heisst* **Stammfunktion** *zu* f*, falls gilt*

$$F' = \frac{dF}{dx} = f \quad \text{in }]a,b[.$$

Beispiel 6.1.1. i) Wegen $\log'(x) = \frac{1}{x}$, $x > 0$, ist $F(x) = \log(x)$ Stammfunktion zu $f(x) = \frac{1}{x}$, $x > 0$; ebenso $\widetilde{F}(x) = \log(x) + 1$, etc.

ii) Wegen $\arctan'(x) = \frac{1}{1+x^2}$ ist die Funktion $F(x) = \arctan(x) + c$, wobei $c \in \mathbb{R}$ beliebig, Stammfunktion von $f(x) = \frac{1}{1+x^2}$.

Allgemein ist mit F auch $F + c$ Stammfunktion zu vorgegebenem f, wobei $c \in \mathbb{R}$ beliebig. Umgekehrt gilt:

Satz 6.1.1. *Sind* $F_1,\ F_2 \in C^1(]a,b[)$ *Stammfunktionen zu* $f \in C^0(]a,b[)$*, so gilt* $F_1 - F_2 \equiv c \in \mathbb{R}$*.*

Beweis. $(F_1 - F_2)' = f - f = 0$, und die Behauptung folgt mit Korollar 5.2.1.i). □

Sei $f \in C^0(]a,b[)$ und $F \in C^1(]a,b[)$ eine Stammfunktion zu f.

Definition 6.1.2. *Für* $a < x_0 < x < b$ *heisst*

$$\int_{x_0}^{x} f(\xi)d\xi := F(x) - F(x_0)$$

das **Integral** *von* f *über* $[x_0, x]$*.*

Bemerkung 6.1.1. i) Wegen Satz 6.1.1 ist die Definition unabhängig von der Wahl der Stammfunktion.

ii) Das **unbestimmte Integral** $\int f(\xi)d\xi$ (ohne Grenzen) ist eine praktische und suggestive Notation für "die" Stammfunktion von f.

Aus den Beispielen des Abschnitts 5 ergeben sich sofort Stammfunktionen für eine Reihe von elementaren Funktionen wie in der folgenden Tabelle:

f	$\int f(\xi)\ d\xi$
$x^\alpha,\ \alpha \neq -1$	$\frac{x^{\alpha+1}}{\alpha+1}$
x^{-1}	$\log(x)$
$\exp(x)$	$\exp(x)$
$\cos(x)$	$\sin(x)$
$\vdots$	$\vdots$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$
$\frac{1}{1+x^2}$	$\arctan(x)$

Das Auffinden einer Stammfunktion -und damit die Integration- ist also eine "Umkehrung" der Differentiation (Englisch: "Anti-differential").

Weiter ergeben die Regeln des Abschnitts 5 unmittelbar die folgenden Regeln.

Satz 6.1.2. i) *(***Linearität.***) Seien $f, g \in C^0(]a,b[)$ mit Stammfunktionen $F, G \in C^1(]a,b[)$, und seien $\alpha, \beta \in \mathbb{R}$. Dann ist $\alpha F + \beta G \in C^1(]a,b[)$ Stammfunktion zu $\alpha f + \beta g$; d.h.*

$$\int (\alpha f + \beta g)\ dx = \alpha \int f\ dx + \beta \int g\ dx.$$

ii) *(***Partielle Integration.***) Seien $u, v \in C^1(]a,b[)$, und es existiere eine Stammfunktion F zu $f = uv' \in C^0(]a,b[)$. Dann besitzt die Funktion $u'v \in C^0(]a,b[)$ die Stammfunktion*

$$\int u'v\ dx = uv - \int uv'\ dx.$$

Beweis. Summen- und Produktregel gemäss Satz 5.1.2 ergeben

i) $(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g$

ii) $(uv - F)' = u'v + \underbrace{uv' - F'}_{=0} = u'v.$

□

Beispiel 6.1.2. i) Sei $p(x) = a_n x^n + \cdots + a_1 x + a_0$. Mit Satz 6.1.2.i) und $\int x^k\ dx = \frac{x^{k+1}}{k+1}$ folgt

$$\int p\ dx = \frac{a_n}{n+1} x^{n+1} + \cdots + a_0 x.$$

ii) Seien $u(x) = x$, $v(x) = \log(x)$ mit $uv'(x) = x \cdot \frac{1}{x} = 1$, $x > 0$. Mit Satz 6.1.2.ii) folgt

$$\begin{aligned}\int \log(x)\ dx &= \int u'v\ dx = x\log(x) - \int uv'\ dx \\ &= x\log(x) - \int 1\ dx = x\log(x) - x.\end{aligned}$$

iii) Nach k-facher partieller Integration erhalten wir

$$\begin{aligned}\int x^k \underbrace{e^{-x}}_{=-\frac{d}{dx}\left(e^{-x}\right)}\ dx &= -x^k e^{-x} + k\int x^{k-1}e^{-x}\ dx \\ &= -x^k e^{-x} - kx^{k-1}e^{-x} + k(k-1)\int x^{k-2}e^{-x}\ dx \\ &= \cdots = -\sum_{l=1}^{k}\frac{k!}{l!}x^l e^{-x} + k!\underbrace{\int e^{-x}\ dx}_{-e^{-x}}.\end{aligned}$$

Mit Beispiel 5.2.3 folgt die Identität

$$\Gamma(k+1) := \lim_{x\to\infty}\int_0^x x^k e^{-x}\ dx = k!, \quad k \in \mathbb{N}.$$

Satz 6.1.3. (Monotonie) *Seien $f, g \in C^0(]a,b[)$ mit Stammfunktionen $F, G \in C^1(]a,b[)$, und sei $f \le g$. Dann gilt für $a < x_0 < x_1 < b$:*

$$\int_{x_0}^{x_1} f\ dx \le \int_{x_0}^{x_1} g\ dx.$$

Beweis. OBdA $f = 0$. (Betrachte $\tilde{g} = g - f \ge 0$ und benutze Satz 6.1.2.i).) Mit Korollar 5.2.1.ii) folgt aus $G' = \frac{dG}{dx} = g \ge 0$ und Definition 6.1.2

$$\int_{x_0}^{x_1} g\ dx = G(x_1) - G(x_0) \ge 0 = \int_{x_0}^{x_1} f\ dx.$$

□

Beispiel 6.1.3. i) Da für $0 \le x \le \pi/2$ gilt

$$\sin^k(x) - \sin^{k+1}(x) = \underbrace{\sin^k(x)}_{\ge 0}\underbrace{(1-\sin(x))}_{\ge 0} \ge 0, \quad \forall k \in \mathbb{N}_0,$$

erhalten wir mit Satz 6.1.3 die Ungleichung

$$\int_0^{\pi/2}\sin^{k+1}(x)\ dx \le \int_0^{\pi/2}\sin^k(x)\ dx, \quad \forall k \in \mathbb{N}_0.$$

ii) Mit partieller Integration gilt andererseits

$$\begin{aligned}\int_0^{\pi/2} \sin^{k+1}(x)\,dx &= \int_0^{\pi/2} \underbrace{\sin(x)}_{=\frac{d}{dx}(-\cos(x))} \sin^k(x)\,dx \\ &= -\cos(x)\sin^k(x)\Big|_{x=0}^{\pi/2} + k\int_0^{\pi/2} \underbrace{\cos^2(x)}_{=1-\sin^2(x)} \sin^{k-1}(x)\,dx \\ &= k\int_0^{\pi/2} \sin^{k-1}(x)\,dx - k\int_0^{\pi/2} \sin^{k+1}(x)\,dx, \quad \forall k \in \mathbb{N};\end{aligned}$$

d.h.

$$\int_0^{\pi/2} \sin^{k+1}(x)\,dx = \frac{k}{k+1}\int_0^{\pi/2} \sin^{k-1}(x)\,dx, \ \forall k \in \mathbb{N}.$$

Iterativ erhalten wir so die Gleichungen

$$\begin{aligned}\int_0^{\pi/2} \sin^{2n}(x)\,dx &= \frac{2n-1}{2n}\cdot\frac{2n-3}{2n-2}\cdots\frac{1}{2}\int_0^{\pi/2} 1\,dx = \frac{(2n)!}{(2^n n!)^2}\cdot\frac{\pi}{2}, \\ \int_0^{\pi/2} \sin^{2n+1}(x)\,dx &= \frac{2n}{2n+1}\cdot\frac{2n-2}{2n-1}\cdots\frac{2}{3}\underbrace{\int_0^{\pi/2}\sin(x)\,dx}_{=-\cos(x)\big|_{x=0}^{\pi/2}=1} = \frac{(2^n n!)^2}{(2n+1)!}.\end{aligned}$$

Mit i) folgt

$$\frac{(2^n n!)^2}{(2n+1)!} \le \frac{(2n)!}{(2^n n!)^2}\cdot\frac{\pi}{2} \le \frac{(2^{n-1}(n-1)!)^2}{(2n-1)!} = \frac{1}{2n}\frac{(2^n n!)^2}{(2n)!}$$

oder

$$\frac{(2^n n!)^4}{(2n!)^2}\frac{2}{2n+1} \le \pi \le \frac{2}{2n}\frac{(2^n n!)^4}{(2n!)^2}$$

und somit

$$\pi = \lim_{n\to\infty}\frac{1}{n}\Big(\frac{2^{2n}(n!)^2}{(2n)!}\Big)^2 \quad \text{(Wallissches Produkt)}.$$

Einige Werte des Näherungsausdrucks:

$$\frac{1}{n}\Big(\frac{2^{2n}(n!)^2}{(2n)!}\Big)^2 = \begin{cases} 4, & n = 1, \\ 3.221\ldots, & n = 10, \\ 3.157\ldots, & n = 50, \\ 3.149\ldots, & n = 100. \end{cases}$$

Satz 6.1.4. (Gebietsadditivität des Integrals) *Sei $f \in C^0(]a,b[)$ mit Stammfunktion $F \in C^1(]a,b[)$, und seien $a < x_0 \le x_1 \le x_2 < b$. Dann gilt*

$$\int_{x_0}^{x_1} f(x)\,dx + \int_{x_1}^{x_2} f(x)\,dx = \int_{x_0}^{x_2} f(x)\,dx.$$

Beweis. Gemäss Definition 6.1.2 ist die linke Seite gegeben durch

$$(F(x_1) - F(x_0)) + (F(x_2) - F(x_1)) = F(x_2) - F(x_0),$$

was genau der rechten Seite entspricht. □

Beispiel 6.1.4. (Stirlingsche Formel) Für $n \in \mathbb{N}$ gilt

$$\log(n!) = \sum_{k=2}^{n} \log(k).$$

Da $F(x) = x\log(x) - x$ gemäss Beispiel 6.1.2.ii) Stammfunktion von $f(x) = \log(x)$, erhalten wir zudem für jedes $k \in \mathbb{N}$

$$\int_{k-1/2}^{k+1/2} \log(x)\ dx = (x\log(x) - x)\Big|_{x=k-1/2}^{k+1/2}.$$

Entwickeln wir gemäss Bemerkung 5.5.1, so folgt

$$\begin{aligned}(k \pm \frac{1}{2})\log(k \pm \frac{1}{2}) - (k \pm \frac{1}{2}) &= \\ &= (k \pm \frac{1}{2})\big(\log(k) \pm \frac{1}{2k} - \frac{1}{8k^2} + t_k^{\pm}\big) - \big(k \pm \frac{1}{2}\big) \\ &= k\log(k) - k \pm \frac{1}{2}\log(k) + \frac{1}{8k} + s_k^{\pm}\end{aligned} \tag{6.1.1}$$

mit

$$\big|t_k^{\pm}\big| \le \frac{1}{2k^3}, \quad \big|s_k^{\pm}\big| = \left|kt_k^{\pm} \pm \frac{1}{2}(t_k^{\pm} - \frac{1}{8k^2})\right| \le \frac{1}{k^2},$$

also

$$\int_{k-1/2}^{k+1/2} \log(x)\ dx = \log(k) - r_k$$

mit

$$|r_k| \le \big|s_k^{+}\big| + \big|s_k^{-}\big| \le \frac{2}{k^2},\ k \in \mathbb{N}. \tag{6.1.2}$$

Mit Satz 6.1.4 erhalten wir

$$\begin{aligned}\log(n!) &= \sum_{k=2}^{n}\Big(\int_{k-1/2}^{k+1/2} \log(x)\ dx + r_k\Big) \\ &= \int_{1}^{n+1/2} \log(x)\ dx + \sum_{k=2}^{n} r_k - \int_{1}^{3/2} \log(x)\ dx.\end{aligned}$$

Einsetzen von (6.1.1) liefert

$$\begin{aligned}\log(n!) &= \big(n + \frac{1}{2}\big)\log(n) - n + \frac{1}{8n} + s_n^{+} + 1 \\ &\quad + \sum_{k=2}^{n} r_k - \int_{1}^{3/2} \log(x)\ dx \\ &=: \big(n + \frac{1}{2}\big)\log(n) - n + a_n,\end{aligned}$$

wobei wegen (6.1.2) die Folge (a_n) konvergiert. Sei $a = \lim_{n\to\infty} a_n$. Exponenzieren ergibt

$$n! = \sqrt{n} \cdot n^n e^{-n} e^{a_n}$$

mit

$$b_n = e^{a_n} \to b = e^a \quad (n \to \infty).$$

Überraschenderweise kann man b bestimmen, da gemäss Satz 3.3.2 und Beispiel 6.1.3.ii) gilt

$$\begin{aligned} b &= \lim_{n\to\infty} b_n = \lim_{n\to\infty} \frac{b_n^2}{b_{2n}} = \lim_{n\to\infty} \Big(\frac{n!e^n}{\sqrt{n}n^n}\Big)^2 \cdot \Big(\frac{\sqrt{2n}2^{2n}n^{2n}}{(2n)!e^{2n}}\Big) \\ &= \lim_{n\to\infty} \sqrt{\frac{2}{n}} \frac{2^{2n}(n!)^2}{(2n)!} = \sqrt{2\pi}. \end{aligned}$$

Wir erhalten somit die Näherungsformel

$$n! \approx \sqrt{2\pi n} \cdot n^n e^{-n}.$$

Satz 6.1.5. (Substitutionsregel) *Seien $f, g \in C^1(]a, b[)$. Dann gilt für $a < x_0 < x_1 < b$:*

$$\int_{x_0}^{x_1} f'(g(x))g'(x)\ dx = f(g(x))\Big|_{x=x_0}^{x_1} = \int_{g(x_0)}^{g(x_1)} f'(y)\ dy.$$

Beweis. Mit der Kettenregel (Satz 5.1.3)

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

folgt die Behauptung unmittelbar aus Definition 6.1.2. □

Bemerkung 6.1.2. Wir können formal in Satz 6.1.5 die Variable $y = g(x)$ "substituieren" mit "$dy = g'(x)\ dx$".

Beispiel 6.1.5. i)

$$\int_0^2 \frac{x}{\sqrt{1+x^2}}\ dx \overset{(y=1+x^2)}{=} \frac{1}{2}\int_1^5 \frac{dy}{\sqrt{y}} = \sqrt{y}\Big|_{y=1}^5 = \sqrt{5} - 1.$$

ii)

$$\int_0^1 \frac{x}{1+x^2}\ dx \overset{(y=1+x^2)}{=} \frac{1}{2}\int_1^2 \frac{dy}{y} = \frac{1}{2}\log(y)\Big|_{y=1}^2 = \frac{\log 2}{2}.$$

iii)

$$\int_{-\pi/2}^{\pi/2} \cos^2(\phi)\ d\phi \overset{(x=\sin(\phi))}{=} \int_{-1}^{1} \underbrace{\sqrt{1-x^2}}_{=\sqrt{1-\sin^2(\phi)}=\cos(\phi)} \overbrace{dx}^{=\cos(\phi)\ d\phi}.$$

Andererseits folgt nach partieller Integration (mit $u = \sin\phi$, $v = \cos\phi$)

$$\begin{aligned}\int_{-\pi/2}^{\pi/2} \cos^2(\phi)\ d\phi &= \sin(\phi)\cos(\phi)\Big|_{\phi=-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2} \sin^2(\phi)\ d\phi \\ &= \int_{-\pi/2}^{\pi/2} (1-\cos^2(\phi))\ d\phi = \pi - \int_{-\pi/2}^{\pi/2} \cos^2(\phi)\ d\phi;\end{aligned}$$

d.h.

$$\int_{-1}^{1} \sqrt{1-x^2}\ dx = \int_{-\pi/2}^{\pi/2} \cos^2(\phi)\ d\phi = \pi/2.$$

iv)

$$\int_{x_0}^{x_1} \frac{g'(x)}{g(x)}\ dx \overset{(y=g(x))}{=} \int_{g(x_0)}^{g(x_1)} \frac{dy}{y} = \log(y)\Big|_{y=g(x_0)}^{g(x_1)} = \log\Big(\frac{g(x_1)}{g(x_0)}\Big)$$

für $0 < g \in C^1(]a,b[)$, $a < x_0 < x_1 < b$.

v) Insbesondere können wir nun das Anfangswertproblem

$$f' = af, \quad f(0) = f_0 > 0 \tag{6.1.3}$$

auch für eine zeitabhängige Funktion $a = a(x) \in C^0(\mathbb{R})$ mit Stammfunktion lösen. Solange $f(x) > 0$, gilt

$$\frac{f'}{f} = a \quad (\text{“Separation der Variablen”}),$$

und Beispiel iv) liefert

$$\log\Big(\frac{f(x)}{f(0)}\Big) = \int_0^x \frac{f'}{f}\ d\xi = \int_0^x a(\xi)\ d\xi;$$

d.h.

$$f(x) = f(0)\cdot\exp\Big(\int_0^x a(\xi)\ d\xi\Big)$$

analog zu Beispiel 5.6.1.ii).

vi) Die zu (6.1.3) gehörige inhomogene Gleichung

$$\dot f = af + b \tag{6.1.4}$$

mit $a, b \in C^0(\mathbb{R})$ kann man nun ebenfalls lösen. Für eine partikuläre Lösung von (6.1.4) machen wir den Ansatz

$$\begin{aligned}f(t) &= \int_0^t b(s)\ e^{\int_s^t a(\tau)\ d\tau}\ ds \\ &= e^{\int_0^t a(\tau)\ d\tau} \int_0^t b(s)\ e^{-\int_0^s a(\tau)\ d\tau}\ ds,\end{aligned} \tag{6.1.5}$$

sofern die Stammfunktionen existieren. Die rechte Seite in (6.1.5) kann man deuten als Überlagerung der Impulsantworten der Gleichung (6.1.4) auf die infinitesimalen Auslenkungen um den Wert $b(s)$ zu den Zeiten $0 \le s \le t$. Andererseits

erhält man denselben Ausdruck auch leicht aus dem Ansatz $f(t) = c(t)e^{\int_0^t a(\tau)\ d\tau}$ mit variablem $c = c(t$. Daher trägt die Darstellung (6.1.5) auch den Namen **Variation-der-Konstanten Formel**. Wir verifizieren

$$\dot{f}(t) = a(t)f(t) + e^{\int_0^t a(\tau)\ d\tau} \underbrace{\frac{d}{dt}\Big(\int_0^t b(s)e^{-\int_0^s a(\tau)\ d\tau}\ ds\Big)}_{=b(t)e^{-\int_0^t a(\tau)\ d\tau}}$$

$$= a(t)f(t) + b(t).$$

Im Fall $a \equiv -\alpha$, $b \equiv \beta \in \mathbb{R}$ erhalten wir

$$f(t) = e^{-\alpha t}\int_0^t \beta e^{t\alpha s}\ ds = e^{-\alpha t}\frac{\beta}{\alpha}\big(e^{\alpha t} - 1\big) = \frac{\beta}{\alpha}(1 - e^{-\alpha t});$$

vgl. Beispiel 5.7.2.i).

vii) Mittels Separation wie in Bespiel v) kann man auch gewisse nichtlineare Differentialgleichungen lösen. Beispielsweise geht die Gleichung

$$y' = 2xy^2, \quad y(0) = 1 \tag{6.1.6}$$

für eine Funktion $y = y(x) > 0$ nach Separation über in die Form

$$\frac{y'}{y^2} = 2x;$$

d.h.

$$-\Big(\frac{1}{y}\Big)' = (x^2)'.$$

Integration unter Beachtung der Anfangsbedingung ergibt nun sofort

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

also

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

Beachte, dass im Unterschied zu linearen Differentialgleichungen die Lösung y von (6.1.6) nur für $|x| < 1$ existiert.

Partialbruchzerlegung: Mit Hilfe von Satz 6.1.5 kann man die rationalen Funktionen $r(x) = \frac{p(x)}{q(x)}$ elementar integrieren.

Beispiel 6.1.6. Das unbestimmte Integral

$$\int \frac{dx}{1 - x^2} = artanh(x)$$

kann mittels der Zerlegung

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

über den Ansatz

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

mit $a = b = \frac{1}{2}$ elementar berechnet werden

$$\int \frac{dx}{1-x^2} = \frac{1}{2}\int \frac{dx}{1-x} + \frac{1}{2}\int \frac{dx}{1+x} = \frac{1}{2}\log\Big(\frac{1+x}{1-x}\Big) + c.$$

Beispiel 6.1.6 lässt sich verallgemeinern mit Hilfe des folgenden Satzes.

Satz 6.1.6. *Seien p, q Polynome in $\mathbb{R}$ mit*

$$\deg(p) < \deg(q) = n$$

und ohne gmeinsame Nullstelle. Sei weiter

$$q(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 = \prod_{i=1}^{l}(x - x_i)^{m_i},$$

wobei $x_1, \ldots, x_l$ die paarweise verschiedenen Nullstellen $x_i \in \mathbb{C}$ mit Vielfachheit $m_i \in \mathbb{N}$ bezeichnen. Dann gibt es es genau eine **Partialbruchzerlegung**

$$\frac{p(x)}{q(x)} = \sum_{i=1}^{l}\sum_{k=1}^{m_i}\frac{\gamma_{ik}}{(x-x_i)^k} \tag{6.1.7}$$

mit $\gamma_{ik} \in \mathbb{C}$, und für $1 \le i \le l$ erhalten wir für $k = m_i, m_{i-1}, \ldots, 1$ iterativ

$$\gamma_{ik} = \lim_{x\to x_i,\ x\neq x_i}\Big(\frac{p(x)}{q(x)}(x-x_i)^k - \sum_{j=k+1}^{m_i}\frac{\gamma_{ij}}{(x-x_i)^j}\Big). \tag{6.1.8}$$

Beweis. **i) Eindeutigkeit.** Offenbar folgt (6.1.8) aus (6.1.7).

ii) Existenz. Mache den Ansatz (6.1.7) und bringe die rechte Seite auf den Hauptnenner $q(x)$. Der Zähler ist ein Polynom $\tilde{p}$ vom Grad $< n$, und

$$\frac{\tilde{p}(x)}{q(x)} = \frac{p(x)}{q(x)}, \quad x \in \mathbb{R}.$$

Die Gleichung $\tilde{p}(x) = p(x)$ für $x \in \mathbb{R}$ liefert ein System von n Gleichungen für die n Koeffizienten $\gamma_{ik} \in \mathbb{C}$, $1 \le i \le l$, $1 \le k \le m_i$.

Da dieses Gleichungssystem linear ist und nach i) **höchstens** eine Lösung besitzt, folgt die Existenz einer Lösung mit der Rangformel.

□

Bemerkung 6.1.3. Da q nach Annahme reell ist, treten wegen der Identität

$$\prod_{i=1}^{l}(x-x_i)^{m_i} = q(x) = \overline{q(x)} = \prod_{i=1}^{l}(x-\overline{x}_i)^{m_i}$$

für $x \in \mathbb{R}$ nicht reelle Nullstellen in komplex konjugerten Paaren $x_i,\ x_{i+1} = \overline{x_i}$ auf mit Vielfachheiten $m_i = m_{i+1}$. Da für $x \in \mathbb{R}$ ebenfalls gilt

$$\sum_{i=1}^{l}\sum_{k=1}^{m_i}\frac{\gamma_{ik}}{(x-x_i)^k} = \frac{p(x)}{q(x)} = \overline{\Big(\frac{p(x)}{q(x)}\Big)} = \sum_{i=1}^{l}\sum_{k=1}^{m_i}\frac{\overline{\gamma_{ik}}}{(x-\overline{x_i})^k},$$

folgt wegen der Eindeutigkeit der Zerlegung (6.1.1) dann auch

$$\gamma_{(i+1)k} = \overline{\gamma_{ik}}, \quad 1 \le k \le m_i.$$

Wir können die entsprechenden Terme somit zusammenfassen

$$\begin{aligned}\frac{\gamma_{ik}}{(x-x_i)^k} + \frac{\gamma_{(i+1)k}}{(x-x_{i+1})^k} &= \frac{\gamma_{ik}(x-x_{i+1})^k + \gamma_{(i+1)k}(x-x_i)^k}{(x-x_i)^k(x-x_{i+1})^k} \\ &= \frac{2Re\big(\gamma_{ik}(x-\overline{x_i})^k\big)}{|x-x_i|^{2k}}, \quad 1 \le k \le m_i.\end{aligned}$$

Stammfunktionen lassen sich nun wieder elementar angeben.

Beispiel 6.1.7. i) Berechne $\int \frac{dx}{1-x^4}$. Zerlege dazu

$$1 - x^4 = (1-x^2)(1+x^2) = (1-x)(1+x)(1+x^2).$$

Satz 6.1.6 und Bemerkung 6.1.3 führen auf die Zerlegung

$$\frac{1}{1-x^4} = \frac{a}{1-x} + \frac{b}{1+x} + \frac{c+dx}{1+x^2}$$

mit

$$\begin{aligned}a &= \lim_{x\to 1} \frac{1-x}{1-x^4} = \frac{1}{(1+x)(1+x^2)}\Big|_{x=1} = \frac{1}{4}, \\ b &= \lim_{x\to -1} \frac{1+x}{1-x^4} = \frac{1}{(1-x)(1+x^2)}\Big|_{x=-1} = \frac{1}{4}, \\ c+dx &= \Big(\frac{1}{1-x^4} - \frac{1}{4}\big(\frac{1}{1-x} + \frac{1}{1+x}\big)\Big)(1+x^2) \\ &= \frac{1}{1-x^2} - \frac{1}{2}\cdot\frac{1+x^2}{1-x^2} = \frac{1}{2};\end{aligned}$$

d.h.

$$c = \frac{1}{2}, \ d = 0.$$

Es folgt

$$\int \frac{dx}{1-x^4} = \frac{1}{4}\log\frac{1+x}{1-x} + \frac{1}{2}\arctan(x) + c.$$

ii) Berechne

$$\int \frac{x^3 - 2x + 1}{x^2+1}\, dx.$$

Polynomdivision liefert

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

Nach Satz 6.1.1 und Bemerkung 6.1.3 kann $\frac{3x-1}{x^2+1}$ im Reellen nicht weiter zerlegt werden. Wir erhalten

$$\begin{aligned}\int \frac{x^3-2x+1}{x^2+1}\, dx &= \int x\, dx - \int \frac{3x}{x^2+1}\, dx + \int \frac{dx}{x^2+1} \\ &= \frac{x^2}{2} - \frac{3}{2}\log(1+x^2) + \arctan(x) + c.\end{aligned}$$

6.2 Das Riemannsche Integral

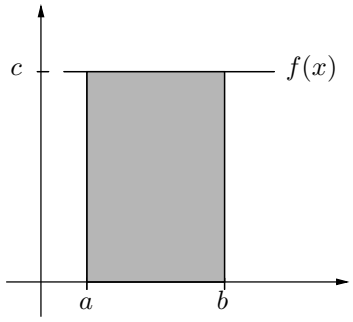
Wir wollen nun den Integralbegriff ausdehnen auf eine möglichst grosse Klasse von Funktionen $f\colon [a,b] \to \mathbb{R}$, wobei $-\infty < a < b < \infty$. Insbesondere wollen wir zeigen, dass jedes $f \in C^0([a,b])$ eine Stammfunktion besitzt.

Ausgangspunkt ist die folgende geometrische Interpretation der Stammfunktion.

Beispiel 6.2.1. i) Sei $f \equiv c$ für ein $0 < c \in \mathbb{R}$ mit Stammfunktion $F(x) = cx$, $x \in \mathbb{R}$. Dann stimmt für $a < b$ das Integral

$$\int_a^b f\ dx = \int_a^b c\ dx = cx\Big|_a^b = c(b-a)$$

überein mit dem elementargeometrisch definierten Flächeninhalt des Bereiches zwischen dem Intervall $[a,b]$ und dem Graphen von f.

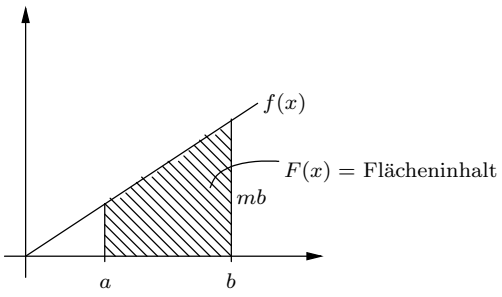


ii) Für $f \equiv c$ mit $c < 0$, $c \in \mathbb{R}$ gilt die Aussage i) analog, sofern wir den Flächeninhalt "mit der richtigen Orientierung" messen; vgl. später.

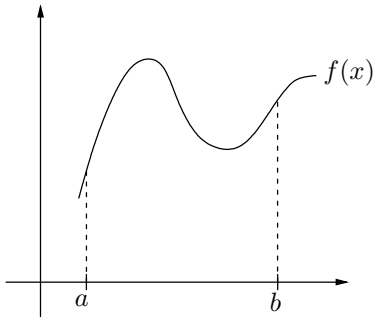
iii) Falls $f(x) = mx$ für ein $m \in \mathbb{R}$ mit Stammfunktion $F(x) = \frac{1}{2}mx^2$, $x \in \mathbb{R}$, so können wir für $a < b$ das Integral

$$\int_a^b mx\ dx = \frac{1}{2}mx^2\Big|_{x=a}^b = \frac{1}{2}mb^2 - \frac{1}{2}ma^2$$

ebenfalls interpretieren als den "orientierten Flächeninhalt zwischen $[a,b]$ und $\mathcal{G}(f)$".



Für welche Funktionen $f\colon [a,b] \to \mathbb{R}$ kann man den Flächeninhalt zwischen $[a,b]$ und $\mathcal{G}(f)$ messen? - Liefert dies den gewünschten Integralbegriff?



Definition 6.2.1. i) *Eine Funktion* $f\colon [a,b] \to \mathbb{R}$ *heisst* **Treppenfunktion**, *falls für eine Zerlegung von* $I = [a,b]$ *in disjunkte (abgeschlossene, offene, halboffene) Teilintervalle* $I_1, \ldots, I_K$ *mit Konstanten* $c_k \in \mathbb{R}$ *gilt*

$$f(x) = c_k \text{ für } x \in I_k,\ 1 \le k \le K;$$

d.h.

$$f = \sum_{k=1}^{K} c_k \chi_{I_k},$$

wobei

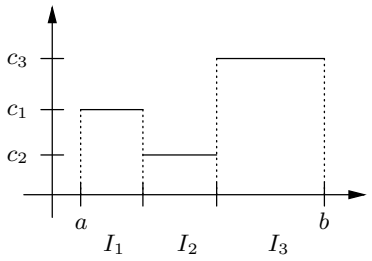
$$\chi_{I_k}(x) = \begin{cases} 1, & x \in I_k \\ 0, & x \notin I_k \end{cases}$$

die charakteristische Funktion von I_k *ist.*

ii) *Das Integral einer Treppenfunktion* $f = \sum\limits_{k=1}^{K} c_k \chi_{I_k}\colon [a,b] \to \mathbb{R}$ *ist*

$$\int_a^b \Big(\sum_{k=1}^{K} c_k \chi_{I_k} \Big)\, dx = \sum_{k=1}^{K} c_k \, |I_k| \,,$$

wobei $|I_k|$ *die Länge von* I_k *bezeichnet,* $1 \le k \le K$.



Bemerkung 6.2.1. i) Die konstante Funktion $f \equiv c$, $c \in \mathbb{R}$, kann man auch schreiben in der Form einer Treppenfunktion

$$f = \sum_{k=1}^{K} c_k \chi_{I_k} \quad \text{mit } c_k = c,\ 1 \le k \le K,$$

wobei $I_1, \dots, I_K$ disjunkte Zerlegung von $I = [a, b]$, und

$$\int_a^b f \, dx = c(b-a) = \sum_{k=1}^{K} c_k \, |I_k| \, .$$

ii) Analog kann man bei einer beliebigen Treppenfunktion $f = \sum\limits_{k=1}^{K} c_k \chi_{I_k}$ die Zerlegungsintervalle I_k weiter zerlegen ("verfeinern"), und das Integral ändert sich nicht.

Die vorherige Bemerkung ist wichtig für den Beweis der folgenden Aussage, die für das Weitere fundamental ist.

Lemma 6.2.1. *Sind $e, g \colon [a, b] \to \mathbb{R}$ Treppenfunktionen mit $e \le g$, dann gilt*

$$\int_a^b e \, dx \le \int_a^b g \, dx.$$

Beweis. Seien $e = \sum\limits_{k=1}^{K} c_k \chi_{I_k}$, $g = \sum\limits_{l=1}^{L} d_l \chi_{J_l}$ mit disjunkten Intervallen $I_1, \dots, I_K$ bzw. $J_1, \dots, J_L$, wobei

$$I = [a, b] = \bigcup_{k=1}^{K} I_k = \bigcup_{l=1}^{L} J_l.$$

Die Intervalle

$$I_{kl} = I_k \cap J_l, \quad 1 \le k \le K, \ 1 \le l \le L$$

sind dann ebenfalls disjunkt mit

$$I_k = \bigcup_{l=1}^{L} I_{kl}, \quad J_l = \bigcup_{k=1}^{K} I_{kl} \tag{6.2.1}$$

und

$$\bigcup_{k,l} I_{kl} = \bigcup_{k=1}^{K} \Big(\bigcup_{l=1}^{L} I_{kl} \Big) = \bigcup_{k=1}^{K} I_k = I.$$

Aus der Annahme $e \le g$ folgt die Ungleichung

$$c_k = e(x) \le g(x) = d_l, \quad \forall x \in I_{kl};$$

also

$$c_k \le d_l, \quad \text{falls } I_{kl} \neq \emptyset. \tag{6.2.2}$$

Da mit (6.2.1) auch gilt

$$|I_k| = \sum_{l=1}^{L} |I_{kl}| \, , \ |J_l| = \sum_{k=1}^{K} |I_{kl}| \, ,$$

erhalten wir mit Bemerkung 6.2.1

$$\int_a^b e\ dx = \sum_{k=1}^{K} c_k\,|I_k| = \sum_{k,l} c_k\,|I_{kl}|$$
$$\overset{(6.2.2)}{\leq} \sum_{k,l} d_l\,|I_{kl}| = \sum_{l=1}^{L} |J_l| = \int_a^b g\ dx,$$

wie gewünscht. □

Sei $f\colon [a,b] \to \mathbb{R}$ beschränkt; d.h.

$$\exists c \in \mathbb{R}\ \forall x \in [a,b]:\ |f(x)| \leq c.$$

Dann gibt es stets (mindestens) ein Paar von Treppenfunktionen $e, g\colon [a,b] \to \mathbb{R}$ mit $e \leq f \leq g$, und die folgende Definition ist sinnvoll.

Definition 6.2.2. i) *Für beschränktes $f\colon [a,b] \to \mathbb{R}$ bezeichnen*

$$\underline{\int_a^b} f\ dx = \sup\Big\{\int_a^b e\ dx;\ e \text{ Treppenfunktion},\ e \leq f\Big\},$$

bzw.

$$\overline{\int_a^b} f\ dx = \inf\Big\{\int_a^b g\ dx;\ g \text{ Treppenfunktion},\ f \leq g\Big\}$$

das **untere**, *bzw.* **obere** *Riemann-Intergral (R-Integral) von f.*

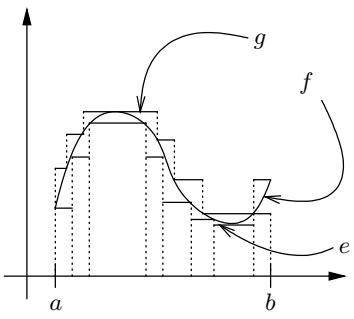
ii) *Ein solches f heisst über $[a,b]$* **Riemann-integrabel** *(R-integrabel), falls*

$$\underline{\int_a^b} f\ dx = \overline{\int_a^b} f\ dx =: A.$$

In diesem Fall heisst

$$A =: \int_a^b f\ dx$$

das **Riemann-Integral** *(R-Integral) von f.*



Bemerkung 6.2.2. i) Lemma 6.2.1 liefert für jedes beschränkte f die Unglei-

chung

$$\underline{\int_a^b} f\ dx \leq \overline{\int_a^b} f\ dx.$$

ii) Eine Funktion f ist R-integrabel genau dann, wenn zu jedem $\epsilon > 0$ Treppenfunktionen $e, g\colon [a,b] \to \mathbb{R}$ existieren mit $e \leq f \leq g$ und

$$\int_a^b g\ dx - \int_a^b e\ dx < \epsilon.$$

Beweis. **i)** Für Treppenfunktionen $e, g\colon [a,b] \to \mathbb{R}$ mit $e \leq f \leq g$ gilt nach Lemma 6.2.1

$$\int_a^b e\ dx \leq \int_a^b g\ dx;$$

also auch

$$\underline{\int_a^b} f\ dx = \sup_{e \leq f \text{ Trpfkt.}} \int_a^b e\ dx \leq \int_a^b g\ dx,$$

und die Behauptung folgt nach Übergang zum Infimum bzgl. $g \geq f$.

ii) Die Aussage ii) folgt nun unmittelbar aus der Ungleichung

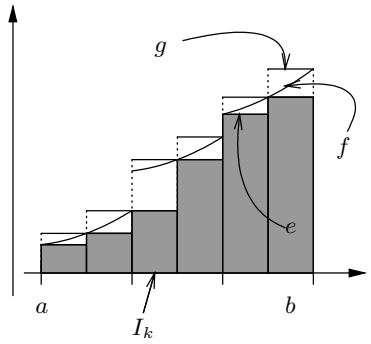
$$\int_a^b e\ dx \leq \underline{\int_a^b} f\ dx \overset{i)}{\leq} \overline{\int_a^b} f\ dx \leq \int_a^b g\ dx$$

für alle Treppenfunktionen $e, g\colon [a,b] \to \mathbb{R}$ mit $e \leq f \leq g$.

□

Wir können für eine grosse Zahl von Funktionen $f\colon [a,b] \to \mathbb{R}$ zeigen, dass sie R-integrabel sind.

Satz 6.2.1. *Sei $f\colon [a,b] \to \mathbb{R}$ monoton. Dann ist f über $[a,b]$ R-integrabel.*



Beweis. OBdA sei f monoton wachsend, also

$$f(a) \leq f(x) \leq f(y) \leq f(b) \quad \forall a \leq x \leq y \leq b.$$

Setze

$$c = \sup_{a \leq x \leq b} |f(x)| = \max\{|f(a)|, |f(b)|\} < \infty.$$

Für $K \in \mathbb{N}$ unterteile $[a,b]$ in K disjunkte Teilintervalle I_k mit Endpunkten

$$a_k = a + (k-1)\frac{b-a}{K},\ b_k = a_k + \frac{b-a}{K} = a_{k+1}$$

und Länge

$$|I_k| = \frac{b-a}{K},\quad 1 \le k \le K.$$

Dann sind $e = \sum_{k=1}^{K} c_k \chi_{I_k}$, $g = \sum_{k=1}^{K} d_k \chi_{I_k}$ mit

$$c_k = \inf_{I_k} f_k,\ d_k = \sup_{I_k} f_k$$

Treppenfunktionen mit $e \le f \le g$.

Weiter gilt

$$d_k = \sup_{I_k} f \le f(b_k) = f(a_{k+1}) \le \inf_{I_{k+1}} f = c_{k+1};$$

also

$$\begin{aligned}\int_a^b g\ dx - \int_a^b e\ dx &= \sum_{k=1}^{K} (d_k - c_k)\,|I_k| \\ &= \frac{b-a}{K}\sum_{k=1}^{K}(d_k - c_k) \\ &= \frac{b-a}{K}\Big(\underbrace{d_K - c_1}_{\le 2c} + \sum_{k=1}^{K-1}\underbrace{(d_k - c_{k+1})}_{\le 0}\Big) \\ &\le 2c\frac{b-a}{K} \to 0\quad (K \to \infty).\end{aligned}$$

Die Behauptung folgt somit aus Bemerkung 6.2.2.ii). □

Satz 6.2.2. *Sei $f\colon [a,b] \to \mathbb{R}$ stetig. Dann ist f über $[a,b]$ R-integrabel.*

Beweis. Da $[a,b]$ kompakt, ist f nach Satz 4.2.3 und 4.7.3 beschränkt und gleichmässig stetig. Zu $\epsilon > 0$ wähle $\delta > 0$ mit

$$|x-y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon,\quad \forall x, y \in [a,b]. \tag{6.2.3}$$

Für $K \in \mathbb{N}$ mit $\frac{b-a}{K} < \delta$ unterteile $[a,b]$ in disjunkte Teilintervalle I_k mit Endpunkten

$$a_k = a + (k-1)\frac{b-a}{K},\quad b_k = a_k + \frac{b-a}{K},$$

und Länge

$$|I_k| = \frac{b-a}{K},\ 1 \le k \le K,$$

wie in Satz 6.2.1, und setze

$$c_k = \inf_{I_k} f \le \sup_{I_k} f = d_k,\quad 1 \le k \le K.$$

Dann sind

$$e = \sum_{k=1}^{K} c_k \chi_{I_k}, \quad g = \sum_{k=1}^{K} d_k \chi_{I_k}$$

Treppenfunktionen mit $e \leq f \leq g$.

Da für $1 \leq k \leq K$ nach Konstruktion

$$\sup_{x,y \in I_k} |x - y| = |I_k| < \delta,$$

folgt mit (6.2.3) auch

$$d_k - c_k \leq \sup_{x,y \in I_k} |f(x) - f(y)| \leq \epsilon,$$

und wir können abschätzen

$$\int_a^b g \, dx - \int_a^b e \, dx = \sum_{k=1}^{K} (d_k - c_k) \, |I_k| \leq \epsilon \sum_{k=1}^{K} |I_k| = (b-a)\epsilon.$$

Die Behauptung folgt aus Bemerkung 6.2.2.ii). □

Das Integral einer Funktion $f \in C^0([a,b])$ kann man numerisch bequem approximieren. Es gilt

Satz 6.2.3. *Sei $f \in C^0([a,b])$. Dann gilt für eine beliebige Folge von Zerlegungen*

$$I = [a,b] = \bigcup_{k=1}^{K_n} I_k^n$$

von I in disjunkte Teilintervalle I_k^n, $1 \leq k \leq K_n$, mit **Feinheit**

$$\delta_n = \sup_{1 \leq k \leq K_n} |I_k^n| \to 0 \quad (n \to \infty)$$

und eine beliebige Auswahl von Punkten $x_k^n \in I_k^n$, $1 \leq k \leq K_n$, stets

$$\int_a^b \Big(\sum_{k=1}^{K_n} f(x_k^n) \chi_{I_k^n} \Big) \, dx = \sum_{k=1}^{K_n} f(x_k^n) \, |I_k^n| \to \int_a^b f \, dx \quad (n \to \infty)$$

Beweis. Zu $\epsilon > 0$ wähle $\delta \equiv \delta(\epsilon) > 0$ mit (6.2.3) wie in Satz 6.2.2, dazu $n_0 = n_0(\epsilon) \in \mathbb{N}$ mit

$$\delta_n < \delta, \quad \forall n \geq n_0.$$

Für $n \in \mathbb{N}$ setze weiter

$$c_k^n = \inf_{I_k^n} f \leq f(x_k^n) \leq \sup_{I_k^n} f = d_k^n, \quad 1 \leq k \leq K.$$

Wie in Satz 6.2.2 erhalten wir für $n \geq n_0$ die Abschätzung

$$d_k^n - c_k^n \leq \sup_{x,y \in I_k^n} |f(x) - f(y)| \leq \epsilon, \; 1 \leq k \leq K_n.$$

Definiere die Treppenfunktionen

$$e_n = \sum_{k=1}^{K_n} c_k^n \chi_{I_k^n} \le f_n = \sum_{k=1}^{K_n} f(x_k^n)\chi_{I_k^n} \le g_n = \sum_{k=1}^{K_n} d_k^n \chi_{I_k^n}.$$

Da f gemäss Satz 6.2.2 R-integrabel ist, können wir für $n \ge n_0(\epsilon)$ abschätzen

$$\begin{aligned} R_n &:= \int_a^b f\ dx - \int_a^b \Big(\sum_{k=1}^{K_n} f(x_k^n)\chi_{I_k^n}\Big)\ dx = \overline{\int_a^b} f\ dx - \int_a^b f_n\ dx \\ &\le \int_a^b g_n\ dx - \int_a^b e_n\ dx = \sum_{k=1}^{K_n} (d_k^n - c_k^n)\,|I_k^n| \le \epsilon \sum_{k=1}^{K_n} |I_k^n| = (b-a)\epsilon\ . \end{aligned}$$

Analog erhalten wir $R_n \ge -(b-a)\epsilon$, und die Behauptung folgt. □

Beispiel 6.2.2. i) Die stetige Funktion $f(x) = e^{x^2}$ ist über jedes Intervall $[a,b]$ R-integrabel; eine Stammfunktion lässt sich jedoch nicht elementar berechnen.

ii) Die Funktion $f = \chi_{\mathbb{Q}\cap[0,1]} \colon [0,1] \to \mathbb{R}$ ist **nicht** R-integrabel. Für jedes Intervall $I \subset [0,1]$ mit $|I| > 0$ gilt gemäss Beispiel 4.3.4.ii)

$$I \cap \mathbb{Q} \ne \emptyset \ne I \backslash \mathbb{Q}$$

und daher

$$0 = \inf_I f < \sup_I f = 1.$$

Für Treppenfunktionen $e = \sum\limits_{k=1}^{K} c_k \chi_{I_k}$, $g = \sum\limits_{l=1}^{L} d_l \chi_{J_l} \colon [0,1] \to \mathbb{R}$ mit $e \le f \le g$ folgt

$$\int_0^1 e\ dx = \sum_{k=1}^{n} c_k\,|I_k| \le 0,\ \int_0^1 g\ dx = \sum_{l=1}^{L} d_l\,|J_l| \ge 1;$$

also

$$\underline{\int_0^1} f\ dx \le 0 < 1 \le \int_0^1 f\ dx.$$

Mit Hilfe von Satz 6.2.3 können wir den Wert gewisser endlicher Summen approximativ berechnen, indem wir sie als "Riemann-Summen" deuten.

Beispiel 6.2.3. i) Für $\alpha > 0$ schreibe

$$\frac{\sum\limits_{k=1}^{n} k^\alpha}{n^{\alpha+1}} = \frac{1}{n}\sum_{k=1}^{n} \Big(\frac{k}{n}\Big)^\alpha$$

und deute $\frac{1}{n} = |I_k^n|$ für eine äquidistante Zerlegung von $I = [0,1]$ in n disjunkte Intervalle $I_1^n, \dots, I_n^n$. Setzen wir nun für festes n noch $\frac{k}{n} = x_k^n$, $1 \le k \le n$, und definieren wir $f(x) = x^\alpha$, $x \in I$, so können wir die Summen deuten als Riemann-Summen für f, und gemäss Satz 6.2.3 erhalten wir

$$\frac{\sum\limits_{k=1}^{n} k^\alpha}{n^{\alpha+1}} = \frac{1}{n}\sum_{k=1}^{n} \Big(\frac{k}{n}\Big)^\alpha \overset{(n\to\infty)}{\to} \int_0^1 x^\alpha\ dx = \frac{1}{\alpha+1}\ .$$

ii) Analog erhalten wir

$$\frac{1}{n+1}+\cdots+\frac{1}{2n}=\frac{1}{n}\sum_{k=n+1}^{2n}\frac{1}{\frac{k}{n}}\overset{(n\to\infty)}{\to}\int_1^2\frac{dx}{x}=\ln(2)\ .$$

6.3 Integrationsregeln, Hauptsatz

Analog zu Satz 6.1.3 für das Rechnen mit Stammfunktionen gilt

Satz 6.3.1. (Monotonie des R-Integrals) *Seien $f_{1,2}\colon [a,b]\to\mathbb{R}$ beschränkt und R-integrabel mit $f_1\le f_2$. Dann gilt*

$$\int_a^b f_1\ dx\le\int_a^b f_2\ dx.$$

Beweis. Jede Treppenfunktion $g\colon [a,b]\to\mathbb{R}$ mit $f_2\le g$ erfüllt auch $f_1\le g$; also

$$\int_a^b f_1\ dx=\overline{\int_a^b}f_1\ dx\le\int_a^b g\ dx.$$

Nach Übergang zum Infimum bzgl. derartiger Treppenfunktionen $g\ge f_2$ erhalten wir

$$\int_a^b f_1\ dx\le\overline{\int_a^b}f_2\ dx=\int_a^b f_2\ dx,$$

wie gewünscht. □

Bemerkung 6.3.1. Mit der Interpretation des Integrals als (orientierter) Flächeninhalt ist Satz 6.3.1 auch geometrisch evident.

Satz 6.3.2. (Linearität des R-Integrals) *Seien $f,f_1,f_2\colon [a,b]\to\mathbb{R}$ R-integrabel, und sei $\alpha\in\mathbb{R}$. Dann sind die Funktionen αf, f_1+f_2 über $[a,b]$ R-integrabel, und*

$$\int_a^b(\alpha f)\ dx=\alpha\int_a^b f\ dx,$$
$$\int_a^b(f_1+f_2)\ dx=\int_a^b f_1\ dx+\int_a^b f_2\ dx.$$

Beweis. Die Behauptung gilt offenbar für Treppenfunktionen. Der allgemeine Fall lässt sich darauf zurückführen.

i) Sei $\alpha \geq 0$. Für Treppenfunktionen e, g mit $e \leq f \leq g$ gilt $\alpha e \leq \alpha f \leq \alpha g$; also

$$\begin{aligned}
\overline{\int_a^b}(\alpha f)\,dx &\leq \inf\Big\{\underbrace{\int_a^b(\alpha g)\,dx}_{=\alpha\int_a^b g\,dx};\ g \text{ Treppenfunktion, } g \geq f\Big\} \\
&= \alpha\inf\left\{\int_a^b g\,dx;\ g \text{ Treppenfunktion, } g \geq f\right\} = \alpha\overline{\int_a^b} f\,dx \\
&\overset{(f \text{ R-int.})}{=} \alpha\underline{\int_a^b} f\,dx = \alpha\sup\left\{\int_a^b e\,dx;\ e \text{ Treppenfunktion, } e \leq f\right\} \\
&= \sup\Big\{\underbrace{\alpha\int_a^b e\,dx}_{=\int_a^b(\alpha e)\,dx};\ e \text{ Treppenfunktion, } e \leq f\Big\} \leq \underline{\int_a^b}(\alpha f)\,dx
\end{aligned}$$

Nach Bemerkung 6.2.2.i) ist die Funktion αf daher R-integrabel mit

$$\int_a^b(\alpha f)dx = \alpha\int_a^b f dx.$$

Analog für $\alpha = -1$ (und damit auch allgemein $\alpha < 0$).

ii) Für Treppenfunktionen e_i, g_i mit $e_i \leq f_i \leq g_i$, $i = 1, 2$ gilt:

$$e_1 + e_2 \leq f_1 + f_2 \leq g_1 + g_2$$

und $e_1 + e_2$, beziehungsweise $g_1 + g_2$ sind Treppenfunktionen. Es folgt:

$$\begin{aligned}
\overline{\int_a^b}(f_1+f_2)\,dx &\leq \inf\Big\{\underbrace{\int_a^b(g_1+g_2)\,dx}_{=\int_a^b g_1 dx+\int_a^b g_2 dx};\ g_i \text{ Treppenfkt., } g_i \geq f_i,\ i=1,2\Big\} \\
&= \inf\left\{\int_a^b g_1\,dx;\ g_1 \text{ Treppenfunktion, } g_1 \geq f_1\right\} \\
&\quad + \inf\left\{\int_a^b g_2\,dx;\ g_2 \text{ Treppenfunktion, } g_2 \geq f_2\right\} \\
&= \overline{\int_a^b} f_1\,dx + \overline{\int_a^b} f_2\,dx = \underline{\int_a^b} f_1\,dx + \underline{\int_a^b} f_2\,dx \\
&= \sup\left\{\int_a^b e_1\,dx;\ e_1 \text{ Treppenfunktion, } e_1 \leq f_1\right\} \\
&\quad + \sup\left\{\int_a^b e_2\,dx;\ e_2 \text{ Treppenfunktion, } e_2 \leq f_2\right\} \\
&= \sup\left\{\int_a^b(e_1+e_2)\,dx;\ e_i \text{ Treppenfkt., } e_i \leq f_i,\ i=1,2\right\} \\
&\leq \underline{\int_a^b}(f_1+f_2)\,dx \leq \overline{\int_a^b}(f_1+f_2)\,dx\ .
\end{aligned}$$

Also ist $f_1 + f_2$ R-integrabel mit

$$\int_a^b (f_1 + f_2)dx = \int_a^b f_1 dx + \int_a^b f_2 dx \ .$$

□

Korollar 6.3.1. *Für $f \in C^0([a,b])$ gilt*

$$\left| \int_a^b f \ dx \right| \leq \int_a^b |f| \ dx \leq \|f\|_{C^0} (b-a).$$

Beweis. Mit $\pm f \leq |f| \leq \|f\|_{C^0}$ folgt die Behauptung aus Satz 6.3.1 und 6.3.2. □

Korollar 6.3.2. *Seien $f, f_k \in C^0([a,b])$ mit $f_k \overset{glm}{\to} f$ $(k \to \infty)$. Dann gilt*

$$\begin{aligned} \left| \int_a^b f_k \ dx - \int_a^b f \ dx \right| &\leq \int_a^b |f_k - f| \ dx \\ &\leq (b-a) \, \|f_k - f\|_{C^0} \to 0 \quad (k \to \infty). \end{aligned}$$

Beweis. Unmittelbar aus Satz 6.3.2 und Korollar 6.3.1. □

Gemäss Beispiel 4.8.1.ii) sind Potenzreihen

$$p(x) = \sum_{k=0}^{\infty} c_k x^k$$

für jedes

$$r < \rho = \frac{1}{\overline{\lim\limits_{k\to\infty}} \sqrt[k]{|c_k|}}$$

in $B_r(0)$ gleichmässig konvergent. Korollar 6.3.2 ergibt somit für $-\rho < a < b < \rho$ die Darstellung

$$\begin{aligned} \int_a^b p(x) \ dx &= \int_a^b \Big(\lim_{n\to\infty} \sum_{k=0}^{n} c_k x^k \Big) \ dx = \lim_{n\to\infty} \int_a^b \sum_{k=0}^{n} c_k x^k \ dx \\ &= \lim_{n\to\infty} \Big(\sum_{k=0}^{n} c_k \int_a^b x^k \ dx \Big) = \sum_{k=0}^{\infty} \frac{c_k}{k+1} (b^{k+1} - a^{k+1}). \end{aligned}$$

Korollar 6.3.3. *Potenzreihen dürfen im Innern ihres Konvergenzkreises gliedweise integriert werden.*

Beispiel 6.3.1. Für $0 \leq b < 1$ erhalten wir

$$\begin{aligned}\log(1+b) &= \int_1^{1+b} \frac{dx}{x} = \int_1^{1+b} \frac{dx}{1-(1-x)} \\ &= \int_1^{1+b} \Big(\sum_{k=0}^{\infty} (1-x)^k\Big)\, dx = \sum_{k=0}^{\infty} \frac{-(1-x)^{k+1}}{k+1}\Big|_{x=1}^{1+b} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k b^{k+1}}{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} b^k}{k};\end{aligned}$$

d.h., wir erhalten die Taylor-Reihe von log um $x_0 = 1$. Beachte, dass nach Beispiel 3.5.1.iii) die alternierende Reihe $\sum\limits_{k=0}^{\infty} (-1)^{k-1} \frac{b^k}{k}$ sogar für alle $0 \leq b \leq 1$ konvergiert mit Fehlerabschätzung

$$\left|\log(1+b) - \sum_{k=1}^{n} \frac{(-1)^{k-1} b^k}{k}\right| \leq \frac{b^{n+1}}{n+1}, \quad n \in \mathbb{N}.$$

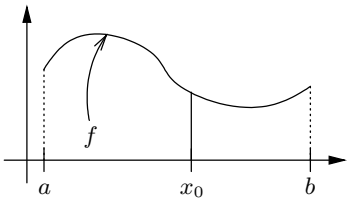
Grenzübergang $b \to 1$ und anschliessend $n \to \infty$ liefert die Summenformel

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = \log 2$$

für die alternierende harmonische Reihe.

Satz 6.3.3. (Gebietsadditivität) *Sei $f\colon [a,b] \to \mathbb{R}$ R-integrabel über $[a,b]$, und sei $x_0 \in [a,b]$. Dann sind die Funktionen $f\big|_{[a,x_0]}$, bzw. $f\big|_{[x_0,b]}$ R-integrabel, und es gilt*

$$\int_a^b f\ dx = \int_a^{x_0} f\ dx + \int_{x_0}^b f\ dx.$$



Beweis. **i)** Offenbar gilt die Aussage für Treppenfunktionen.

ii) Sei $f\colon [a,b] \to \mathbb{R}$ R-integrabel. Zu $\epsilon > 0$ wähle Treppenfunktionen e, g mit $e \leq f \leq g$ und

$$\int_a^b g\ dx - \int_a^b e\ dx < \epsilon.$$

Dann gilt $e \le f \le g$ auf $[a, x_0]$, bzw. auf $[x_0, b]$ und nach i) weiter

$$\underbrace{\Big(\int_a^{x_0} g\ dx - \int_a^{x_0} e\ dx\Big)}_{\ge 0} + \underbrace{\Big(\int_{x_0}^{b} g\ dx - \int_{x_0}^{b} e\ dx\Big)}_{\ge 0}$$
$$= \int_a^b g\ dx - \int_a^b e\ dx < \epsilon.$$

Also ist f über $[a, x_0]$ sowie über $[x_0, b]$ nach Bemerkung 6.2.2.ii) R-integrabel.

Weiter gilt

$$A := \int_a^b f\ dx - \Big(\int_a^{x_0} f\ dx + \int_{x_0}^{b} f\ dx\Big)$$
$$\le \int_a^b g\ dx - \underbrace{\Big(\int_a^{x_0} e\ dx + \int_{x_0}^{b} e\ dx\Big)}_{=\int_a^b e\ dx} < \epsilon;$$

analog $A > -\epsilon$.

Da $\epsilon > 0$ beliebig, folgt die Behauptung.

□

Aus Satz 6.3.3 folgt nun sofort der "Hauptsatz der Differential- und Integralrechnung":

Satz 6.3.4. *Sei $f \in C^0([a, b])$. Setze*

$$F\colon x \mapsto \int_a^x f\ dx, \quad x \in [a, b],$$

Dann gilt $F \in C^1(]a, b[)$ mit $F' = f$.

Beweis. Fixiere $x_0 \in]a, b[$, und wähle $\epsilon > 0$ beliebig. Wähle $\delta > 0$ mit

$$\forall x \in [a, b]:\ |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon.$$

Sei $x_0 < x < x_0 + \delta$. Mit Satz 6.3.3 folgt:

$$F(x) - F(x_0) = \int_{x_0}^{x} f\ d\xi.$$

Schätze ab mittels Korollar 6.3.1:

$$\left|\int_{x_0}^{x} f\ d\xi - (x - x_0) \cdot f(x_0)\right| = \left|\int_{x_0}^{x} \big(f(\xi) - f(x_0)\big)\ d\xi\right|$$
$$\le |x - x_0| \sup_{|y - x_0| < \delta} |f(y) - f(x_0)| \le \epsilon\, |x - x_0|\,.$$

Es folgt:

$$\sup_{x_0 < x < x_0 + \delta} \left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| \le \epsilon.$$

Analog für $x_0 - \delta < x < x_0$. □

Korollar 6.3.4. *Für Funktionen mit Stammfunktion ("elementar integrierbare Funktionen") ist deren R-Integral durch deren Stammfunktion gegeben.*

Beweis. Sei $f = F'$ mit $F \in C^1([a,b])$. Dann gilt nach Satz 6.3.4

$$\frac{d}{dx}\Big(F(x) - \int_a^x f\ dx\Big) = f - f = 0,$$

also wegen Korollar 5.2.1.i)

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

□

6.4 Uneigentliches Riemann-Integral

Sei $f\colon]a,b[\to \mathbb{R}$ über jedes kompakte Intervall $[c,d] \subset]a,b[$ R-integrabel.

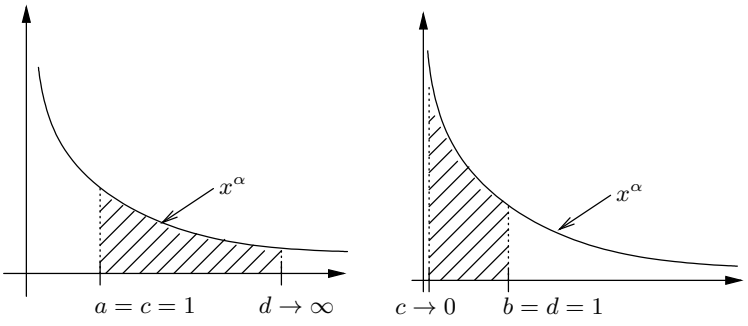
Definition 6.4.1. *f heisst über $]a,b[$* **uneigentlich R-integrabel**, *falls*

$$\int_a^b f\ dx := \lim_{c\downarrow a,\ d\uparrow b} \int_c^d f\ dx$$

existiert.

Beispiel 6.4.1. i) Für $\alpha < -1$ existiert

$$\int_1^\infty x^\alpha\ dx = \lim_{d\to\infty} \int_1^d x^\alpha\ dx = \lim_{d\to\infty} \frac{d^{\alpha+1}-1}{\alpha+1} = \frac{1}{|\alpha|-1}.$$



ii) Für $\alpha > -1$ existiert

$$\int_0^1 x^\alpha\ dx = \lim_{c\downarrow 0} \int_c^1 x^\alpha\ dx = \lim_{c\downarrow 0} \frac{1-c^{\alpha+1}}{\alpha+1} = \frac{1}{\alpha+1}.$$

iii)

$$\int_0^1 \frac{dx}{x} = \lim_{c\downarrow 0} \log(\frac{1}{c}),\ \int_1^\infty \frac{dx}{x} = \lim_{d\to\infty} \log d$$

existieren nicht.

iv)

$$\int_0^\infty e^{-t}\ dt = \lim_{d\to\infty} (1 - e^{-d}) = 1.$$

v) Für alle $\alpha > 0$ existiert

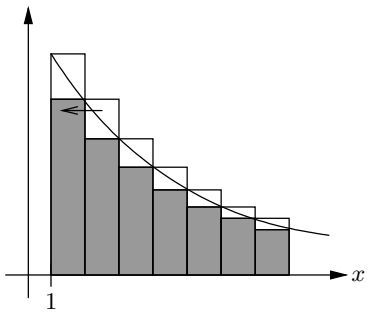
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\ dt;$$

vgl. Beispiel 6.1.2.iii).

Die Konvergenz gewisser Reihen lässt sich auf die Konvergenz von uneigentlichen Integralen zurückführen.

Satz 6.4.1. *Sei $f\colon [1,\infty[\to \mathbb{R}_+$ monoton fallend. Dann konvergiert die Reihe $\sum_{k=1}^{\infty} f(k)$ genau dann, wenn $\int_1^\infty f\ dx$ konvergiert, und in diesem Fall gilt*

$$0 \le \sum_{k=1}^{\infty} f(k) - \int_1^\infty f\ dx \le f(1).$$



Beweis. Die Treppenfunktionen

$$e = \sum_{k=1}^{\infty} f(k+1)\chi_{[k,k+1[},\ g = \sum_{k=1}^{\infty} f(k)\chi_{[k,k+1[}$$

erfüllen wegen der Monotonie von f die Ungleichung $e \le f \le g$; somit folgt

$$\begin{aligned}\int_1^n e\ dx &= \sum_{k=1}^{n-1} f(k+1) = \sum_{k=1}^{n} f(k) - f(1) \le \int_1^n f\ dx \\ &\le \int_1^n g\ dx = \sum_{k=1}^{n-1} f(k) = \sum_{k=1}^{n} f(k) - f(n)\ .\end{aligned}$$

Wir erhalten also

$$0 < f(n) \leq \sum_{k=1}^{n} f(k) - \int_1^n f \, dx \leq f(1) \, .$$

Die Behauptung folgt. □

Beispiel 6.4.2. i) $\zeta(s) = \sum\limits_{k=1}^{\infty} \frac{1}{k^s}$ existiert für alle $s > 1$ gemäss Satz 6.4.1 und Beispiel 6.4.1.i); vgl. Beispiel 3.7.4.

ii) Die Reihe $\sum\limits_{k=2}^{\infty} \frac{1}{k(\log k)^s}$ konvergiert gemäss Satz 6.4.1 für alle $s > 1$, da nach Beispiel 6.4.1.i)

$$\int_2^{\infty} \frac{dx}{x \log^s x} \overset{(y=\log x)}{=} \int_{\log 2}^{\infty} \frac{dy}{y^s} < \infty.$$

6.5 Differentialgleichungen

Sei $f\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ stetig, $u_0 \in \mathbb{R}^n$. Gesucht ist eine Lösung $u \in C^1([0,T[,\mathbb{R}^n)$ der Differentialgleichung

$$\dot{u} = \frac{du}{dt} = f(t,u(t)),\ 0 \leq t < T, \tag{6.5.1}$$

mit Anfangsbedingung

$$u(0) = u_0. \tag{6.5.2}$$

Beispiel 6.5.1. i) Die allgemeine lineare Differentialgleichung (5.6.4) lässt sich in der Form (6.5.1) schreiben mit $f(t,y) = Ay$, $y \in \mathbb{R}^n$.

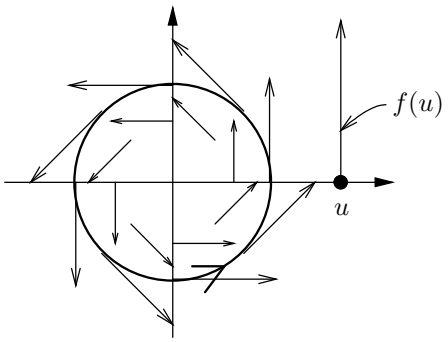
ii) Die Form (6.5.1) umfasst aber auch nichtlineare Gleichungen wie in Beispiel 5.7.2.ii) oder inhomogene Gleichungen.

Geometrisch können wir die Lösungen u von (6.5.1) als "Integralkurven" des durch f gegebenen "Richtungsfeldes" deuten.

Beispiel 6.5.2. i) Sei $n = 2$, $\mathbb{R}^2 \cong \mathbb{C}$, $f\colon \mathbb{C} \to \mathbb{C}$ gegeben durch $f(u) = iu$, $u \in \mathbb{C}$. Die Lösung $u \in C^1(\mathbb{R};\mathbb{C})$ von

$$\dot{u} = f(u) = iu, \quad u(0) = u_0$$

ist $u(t) = u_0 e^{it}$, $t \in \mathbb{R}$; sie beschreibt einen Kreis um $0 \in \mathbb{C}$ mit Radius $|u_0|$.



ii) Für $n = 1$ führt man mit Vorteil die Zeit $t = u^0$ als zusätzliche Variable ein. Für $U(t) = (t, u(t))$ ergibt (6.5.1) die Gleichung

$$\dot{U} = \begin{pmatrix} \frac{dt}{dt} \\ \frac{du}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ f(t, u(t)) \end{pmatrix} = F(U(t))$$

mit

$$F(t, u) = \begin{pmatrix} 1 \\ f(t, u) \end{pmatrix}.$$

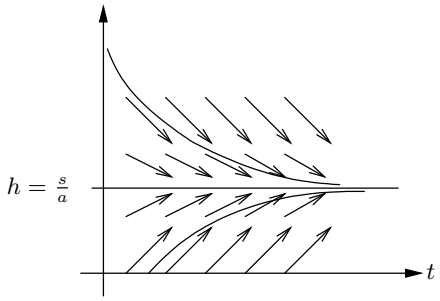
Im Falle $f(t, u) = s - au$ mit Konstanten $s, a > 0$ erhält man ein Richtungsfeld, das sehr schön die Konvergenz jeder Lösung $u \in C^1(\mathbb{R})$ der Gleichung

$$\dot{u} = s - au$$

gegen die Gleichgewichtslösung

$$u_{stat}(t) = \frac{s}{a}$$

zeigt für $t \to \infty$; vgl. Beispiel 5.7.2.i).



Fragen Was für lineare Differentialgleichungen selbverständlich war, muss im allgemeinen Fall des Anfangswertproblems (6.5.1), (6.5.2) nicht mehr gelten. Einige Fragen drängen sich auf:

i) Gibt es zu jedem $u_0 \in \mathbb{R}^n$ stets eine "lokale" Lösung $u \in C^1([0, T]; \mathbb{R}^n)$ des Anfangswertproblems (6.5.1), (6.5.2) für genügend kleines $T > 0$?

ii) Ist diese Lösung eindeutig durch ihre Anfangswerte bestimmt?

iii) Kann man sie für alle $t > 0$ fortsetzen?

iv) Was kann man in den Fällen aussagen, wo eine Fortsetzung nicht möglich ist?

Beispiel 6.5.3. i) Sei $n = 1$, $f(u) = u^2$, $u_0 > 0$. Nach Separation erhält man zu dem Anfangswertproblem

$$\dot{u} = u^2, \quad u(0) = u_0 \tag{6.5.3}$$

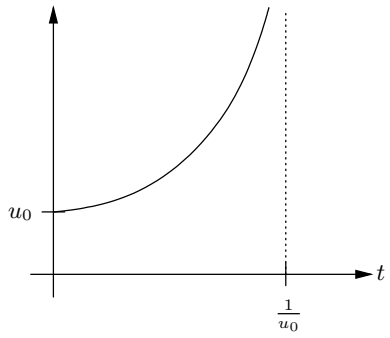
die äquivalente Form

$$\frac{d}{dt}\Big(-\frac{1}{u}\Big) = \frac{\dot{u}}{u^2} = 1, \quad u(0) = u_0$$

mit der eindeutigen Lösung

$$u(t) = \frac{1}{1/u_0 - t}, \quad 0 \le t < 1/u_0.$$

Beachte, dass $u(t) \to \infty$ für $t \uparrow 1/u_0$.



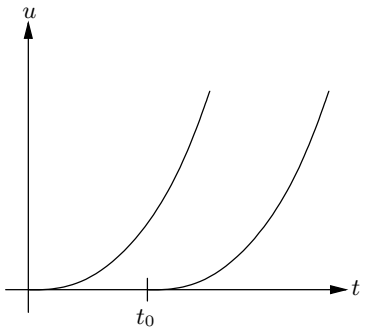
ii) Sei $n = 1$. Das Anfangswertproblem

$$\dot{u} = 2\sqrt{|u|}, \quad u(0) = 0, \tag{6.5.4}$$

hat neben der offensichtlichen Lösung $u \equiv 0$ auch die Funktion $u(t) = t^2$ als Lösung. Tatsächlich sind alle Funktionen

$$u(t) = (t - t_0)_+^2 = (\max\{0, t - t_0\})^2$$

Lösungen von (6.5.4), wobei der Parameter $t_0 \ge 0$ beliebig gewählt werden kann.



Wir müssen also einerseits im allgemeinen damit rechnen, dass Lösungen von (6.5.1), (6.5.2) in endlicher Zeit "explodieren"; andererseits gilt es, geeignete zusätzliche Voraussetzungen an f zu finden, welche die Eindeutigkeit der Lösungen garantieren.

Satz 6.5.1. (Picard-Lindelöf) *Sei $f = f(t,u)\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ stetig und bzgl. $u \in \mathbb{R}^n$ lokal Lipschitz stetig, lokal gleichmässig in $t \in \mathbb{R}$; das heisst, für alle $u_0 \in \mathbb{R}^n$, alle $t_0 > 0$ gibt es $r_0 > 0$, $T_0 > 0$ und eine Konstante L, so dass*

$$|f(t,u) - f(t,v)| \le L|u - v| \text{ für alle } t \in B_{T_0}(t_0),\ u, v \in B_{r_0}(u_0). \qquad (6.5.5)$$

Dann gilt

i) *Zu jedem $u_0 \in \mathbb{R}^n$ existiert ein $T = T(u_0) > 0$ und genau eine Lösung $u = u(t;u_0) \in C^1([0,T],\mathbb{R}^n)$ von (6.5.1), (6.5.2).*

ii) *Die Lösung $u = u(t;u_0)$ hängt stetig ab von u_0 im folgenden Sinn: Für jedes $u_1 \in B_{r_0/2}(u_0)$ sind die zum Anfangswert u_1 gehörigen Lösungen $u(t;u_1) \in C^1([0,T])$ der Gleichung (6.5.1) für $0 \le t \le T$ erklärt, und*

$$\|u(t;u_0) - u(t;u_1)\|_{C^1([0,T])} \le C\,|u_1 - u_0|\,.$$

wobei $T > 0$ wie in i) und $r_0 > 0$ wie in (6.5.5) zu u_0 und $t_0 = 0$ gewählt sind.

Beispiel 6.5.4. i) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(u) = u^2$ ist gemäss Beispiel 4.1.5 lokal Lipschitz stetig bzgl. $u \in \mathbb{R}$.

ii) Allgemein gilt dies für "autonome" $f = f(u) \in C^1(\mathbb{R})$. Für jedes $R > 0$ und beliebige $u, v \in B_R(0)$ gibt es nämlich gemäss dem Mittelwertsatz, Satz 5.2.1, ein $w \in B_R(0)$ mit

$$f(u) - f(v) = f'(w)(u - v).$$

Zum Nachweis von (6.5.5) genügt es daher,

$$L = \sup_{w \in B_R(0)} |f'(w)|$$

zu setzen.

iii) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}$ mit $f(u) = \sqrt{|u|}$ ist **nicht** Lipschitz stetig bei $u_0 = 0$, da

$$\frac{|f(u) - f(0)|}{|u|} = \frac{1}{\sqrt{|u|}} \to \infty \quad (u \to 0).$$

iv) Sei $f\colon \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ gegeben durch

$$f(t,u) = A(t)u + b(t)$$

mit stetigen Funktionen A, b. Dann ist die Funktion f lokal in $t \in \mathbb{R}$ bzgl. $u \in \mathbb{R}^n$ sogar gleichmässig Lipschitz stetig. Zu gegebenem $T < \infty$ können wir nämlich für beliebiges $|t| < T$ und beliebige $u, v \in \mathbb{R}^n$ abschätzen

$$|f(t,u) - f(t,v)| \le |A(t)|\,|u-v| \le \|A\|_{C^0([-T,T])}\,|u-v|\,;$$

das heisst, wir erhalten (6.5.5) mit $L = \|A\|_{C^0([-T,T])}$.

Den Beweis der Aussage i) von Satz 6.5.1 führen wir zurück auf ein Fixpunktproblem im Funktionenraum $C^0([0,T];\mathbb{R}^n)$ für geeignetes $T > 0$. Zur Lösung dieses Fixpunktproblems verwenden wir das **Kontraktionsprinzip** von Stefan Banach, welches uns später auch in anderem Kontext gute Dienste leisten wird.

Sei $(X, \|\cdot\|_X)$ ein Banachraum. (Im Beweis von Satz 6.5.1 werden wir $X = C^0([0,T];\mathbb{R}^n)$ wählen.) Eine Teilmene $M \subset X$ heisst **abgeschlossen**, falls sie (folgen-) abgeschlossen ist im Sinne von Satz 4.3.5.ii), d.h., falls gilt:

$$\forall (x_k) \subset M:\ x_k \to x\ (k\to\infty) \Rightarrow x \in M.$$

Weiter heisst eine Abbildung $\Phi\colon M \to M$ **kontrahierend**, falls gilt

$$\exists q < 1\ \forall x, y \in M:\ \|\Phi(x) - \Phi(y)\|_X \le q\,\|x-y\|_X\,;$$

d.h., falls Φ Lipschitz stetig ist mit Lipschitz Konstante $q < 1$.

Satz 6.5.2. (Banachscher Fixpunktsatz, Kontraktionsprinzip) *Sei $(X, \|\cdot\|_X)$ ein Banachraum, $M \subset X$ abgeschlossen, und sei $\Phi\colon M \to M$ kontrahierend. Dann gibt es genau ein $\overline{x} \in M$ mit $\Phi(\overline{x}) = \overline{x}$. Zudem gilt für jedes $x_0 \in M$ und die Folge*

$$x_1 = \Phi(x_0),\ \ldots,\ x_k = \Phi(x_{k-1}) = \cdots = \Phi^k(x_0),\ k \in \mathbb{N}, \tag{6.5.6}$$

die Abschätzung

$$\|x_k - \overline{x}\|_X \le q^k\,\|x_0 - \overline{x}\|_X\,.$$

Beweis. **i)** Wähle ein $x_0 \in M$ und definiere $(x_k)_{k\in\mathbb{N}}$ wie in (6.5.6). Schätze ab

$$\begin{aligned}\|x_k - x_{k+1}\|_X &= \|\Phi(x_{k-1}) - \Phi(x_k)\|_X \le q\,\|x_{k-1} - x_k\|_X \\ &\le \cdots \le q^k\,\|x_0 - x_1\|_X\,,\quad k \in \mathbb{N}.\end{aligned}$$

Für $l \ge k \in \mathbb{N}$ folgt so

$$\begin{aligned}\|x_k - x_l\|_X &\le \sum_{j=k}^{l-1} \|x_j - x_{j+1}\|_X \le \sum_{j=k}^{l-1} q^j\,\|x_0 - x_1\|_X \\ &\le \frac{q^k}{1-q}\,\|x_0 - x_1\|_X \to 0\ \ (l \ge k \to \infty);\end{aligned}$$

d.h. $(x_k)_{k\in\mathbb{N}}$ ist Cauchy-Folge in X.

Da X nach Annahme vollständig ist, existiert $\overline{x} = \lim\limits_{k\to\infty} x_k \in X$; da M abgeschlossen, gilt $\overline{x} \in M$. Da Φ insbesondere stetig ist, ergibt (6.5.6) die gewünschte Beziehung

$$\overline{x} = \lim_{k\to\infty} x_k = \lim_{k\to\infty} \Phi(x_{k-1}) = \Phi(\overline{x}).$$

Die behauptete Fehlerabschätzung erhält man nun mittels

$$\|x_k - \overline{x}\|_X = \|\Phi(x_{k-1}) - \Phi(\overline{x})\|_X \le q\,\|x_{k-1} - \overline{x}\|_X \le \cdots \le q^k\,\|x_0 - \overline{x}\|_X\,.$$

ii) Zum Beweis der Eindeutigkeit von $\overline{x}$ seien $x, y \in M$ Fixpunkte von Φ. Mit

$$\|x - y\|_X = \|\Phi(x) - \Phi(y)\|_X \le q\,\|x - y\|_X$$

folgt $\|x - y\|_X = 0$, also $x = y$.

□

Die Annahme der Abgeschlossenheit von M kann man nicht weiter abschwächen, wie das folgende Beispiel 6.5.5.i) zeigt.

Beispiel 6.5.5. i) Die Abbildung

$$f\colon]0,1] \ni x \mapsto \frac{x}{2} \in]0,1]$$

ist kontrahierend mit der Konstanten $q = 1/2$, besitzt aber keinen Fixpunkt in $]0,1]$. (Die auf das abgeschlossene Intervall $[0,1]$ stetig fortgesetzte Abbildung f hingegen hat $\overline{x} = 0$ als Fixpunkt.)

ii) Sei $1 \le a \le 2$, und sei $f\colon [1,\infty[\to \mathbb{R}$ mit

$$f(x) = \frac{1}{2}\big(x + \frac{a}{x}\big).$$

Beachte

$$f'(x) = \frac{1}{2}\big(1 - \frac{a}{x^2}\big) \in [-\frac{1}{2}, \frac{1}{2}],\ x \ge 1,\ f''(x) = \frac{a}{x^3} > 0\,;$$

insbesondere hat f genau eine Minimalstelle bei $x = \sqrt{a}$, und wir erhalten

$$f(x) \ge f(\sqrt{a}) = \sqrt{a} \ge 1 \text{ für } x \ge 1\,.$$

Weiter gibt es zu $1 \le x < y$ gemäss Mittelwertsatz, Satz 5.2.1, ein $\xi \in]x, y[$ mit

$$|f(y) - f(x)| = |f'(\xi)|\,|y - x| \le \frac{1}{2}\,|y - x|\,,$$

und f ist kontrahierend.

Zudem gilt offenbar $1 \le f(x) \le a$ für $1 \le x \le a$; also $f\big|_{[1,a]}\colon [1,a] \to [1,a]$.

Gemäss Satz 6.5.2 besitzt f einen eindeutig bestimmten Fixpunkt $\overline{x} \in [1,a]$. Aus der Gleichung

$$\overline{x} = f(\overline{x}) = \frac{1}{2}\big(\overline{x} + \frac{a}{\overline{x}}\big)$$

folgt

$$\overline{x} = \frac{a}{\overline{x}}, \quad \text{also } \overline{x} = \sqrt{a}.$$

Satz 6.5.2 liefert also ein Verfahren zur näherungsweisen Berechnung von $\sqrt{a}$ für jedes $a \in [1,2]$; zudem liefert der Satz die Fehlerabschätzung

$$\left|x_k - \sqrt{a}\right| \leq \left(\frac{1}{2}\right)^k \left|x_0 - \sqrt{a}\right| \leq \left(\frac{1}{2}\right)^k |a-1|$$

für die gemäss (6.5.6) bestimmte Folge von Näherungen x_k, $k \in \mathbb{N}$.

iii) Die Funktion $f\colon \{0,1\} \to \{0,1\}$ mit $f(0) = 1$, $f(1) = 0$ ist Lipschitz-stetig mit Lipschitz-Konstante $L = 1$, hat aber keinen Fizpunkt.

Beweis von Satz 6.5.1. Zu gegebenem $u_0 \in \mathbb{R}^n$ existieren nach Voraussetzung $r_0 > 0$, $T_0 > 0$, $L \in \mathbb{R}$ so, dass gilt

$$|f(t,v) - f(t,w)| \leq L\,|v - w| \tag{6.5.7}$$

für alle $t \in [0,T_0]$ und alle $v, w \in B_{r_0}(u_0)$. Setze

$$C_0 := Lr_0 + \sup_{0 \leq t \leq T_0} |f(t,u_0)| < \infty.$$

Wähle

$$T = \min\{T_0, \frac{r_0}{2C_0}, \frac{1}{2L}\} > 0$$

und setze

$$M = \{u \in C^0([0,T];\mathbb{R}^n);\ \sup_{0 \leq t \leq T} |u(t) - u_0| \leq r_0\}.$$

Dann ist M abgeschlossen im Banachraum $X = C^0([0,T];\mathbb{R}^n)$.

Für $u_1 \in B_{r_0/2}(u_0)$ definiere die Abbildung $\Phi_{u_1}\colon M \to X$ wie folgt: Zu vorgegebenem $v = v(t) \in M$ sei $\Phi_{u_1}(v) \in C^0([0,T];\mathbb{R}^n)$ die Funktion mit

$$(\Phi_{u_1}(v))(t) := u_1 + \int_0^t f(s,v(s))\ ds, \quad 0 \leq t \leq T.$$

Unser Ziel ist es nun zu zeigen, dass die Abbildung Φ_{u_1} einen Fixpunkt besitzt. Anschliessend benutzen wir Satz 6.3.4 zum Nachweis, dass dieser Fixpunkt die gesuchte Lösung $u = u(t;u_1) \in C^1([0,T],\mathbb{R}^n)$ von (6.5.1) mit $u(0) = u_1$ ergibt. Dazu verifizieren wir zunächst die Voraussetzungen des Satzes 6.5.2.

Behauptung 1 Für alle $u_1 \in B_{r_0/2}(u_0)$, $v \in M$ gilt $\Phi_{u_1}(v) \in M$.

Beweis. Schätze ab mit Korollar 6.3.1

$$\begin{aligned}\left|\big(\Phi_{u_1}(v)\big)(t) - u_0\right| &\leq |u_1 - u_0| + \left|\int_0^t f(s,v(s))\ ds\right| \\ &\leq r_0/2 + T \sup_{0 \leq s \leq T} |f(s,v(s))|, \quad 0 \leq t \leq T.\end{aligned}$$

Da wegen (6.5.7) für $0 \leq s \leq T$ weiter gilt

$$\begin{aligned}|f(s,v(s))| &\leq |f(s,v(s)) - f(s,u_0)| + |f(s,u_0)| \\ &\leq L \sup_{0 \leq s \leq T} \underbrace{|v(s) - u_0|}_{\leq r_0} + \sup_{0 \leq s \leq T} |f(s,u_0)| \leq C_0,\end{aligned}$$

folgt mit unserer Wahl von $T \leq \frac{r_0}{2C_0}$ die Abschätzung

$$\left|\big(\Phi_{u_1}(v)\big)(t) - u_0\right| \leq r_0, \quad 0 \leq t \leq T,$$

und $\Phi_{u_1}(v) \in M$ wie gewünscht. □

Als Vorbereitung zum Nachweis der Kontraktionsbedingung schätzen wir für alle $u_1, u_2 \in B_{r_0/2}(u_0)$ und alle $v, w \in M$ mit (6.5.7) ab

$$\begin{aligned} &\left|\big(\Phi_{u_1}(v)\big)(t) - \big(\Phi_{u_2}(w)\big)(t)\right| \\ &\quad = \left|u_1 - u_2 + \int_0^t \big(f(s, v(s)) - f(s, w(s))\big)\, ds\right| \\ &\quad \leq |u_1 - u_2| + L \int_0^t |v(s) - w(s)|\, ds \\ &\quad \leq |u_1 - u_2| + LT \left\|v - w\right\|_{C^0([0,T])}, \end{aligned} \tag{6.5.8}$$

gleichmässig in $0 \leq t \leq T$.

Insbesondere erhalten wir nun die gewünschte Kontraktionseigenschaft.

Behauptung 2 Für alle $u_1 \in B_{r_0/2}(u_0)$ ist $\Phi_{u_1} \colon M \to M$ kontrahierend.

Beweis. Bei Wahl von $u_1 = u_2 \in B_{r_0/2}(u_0)$ ergibt sich wegen $LT \leq 1/2$ aus (6.5.8) sofort

$$\left\|\Phi_{u_1}(v) - \Phi_{u_1}(w)\right\|_{C^0} \leq \frac{1}{2} \left\|v - w\right\|_{C^0}, \forall v, w \in M;$$

□

Wir können nun den Beweis von Satz 6.5.1 vollenden.

i) Da $\Phi_{u_1} \colon M \to M$ kontrahierend, hat Φ_{u_1} gemäss Satz 6.5.2 für alle $u_1 \in B_{r_0/2}(u_0)$ genau einen Fixpunkt $u = u(t) \in M$ mit

$$u(t) = u_1 + \int_0^t f(s, u(s))\, ds, \quad 0 \leq t \leq T. \tag{6.5.9}$$

Nach Satz 6.3.4 ist $u \in C^1([0,T]; \mathbb{R}^n)$, und u erfüllt

$$\dot{u}(t) = \frac{du}{dt}(t) = f(t, u(t)), \quad 0 < t < T.$$

Weiter gilt offenbar

$$u(0) = u_1;$$

also löst $u = u(t; u_1) \in C^1([0,T], \mathbb{R}^n)$ das Anfangswertproblem (6.5.1) mit $u(0) = u_1$.

Da umgekehrt jede Lösung dieses Anfangswertproblems auch (6.5.9) erfüllt, folgt mit Satz 6.5.2 auch die Eindeutigkeit dieser Lösung, und Aussage i) von Satz 6.5.2 ist bewiesen.

ii) Seien $v = \Phi_{u_1}(v)$, $w = \Phi_{u_2}(w)$ die Lösungen des Anfangswertproblems (6.5.1) mit $v(0) = u_1$, bzw. $w(0) = u_2$. Mit (6.5.8) folgt zunächst

$$\left\|v - w\right\|_{C^0} = \left\|\Phi_{u_1}(v) - \Phi_{u_2}(w)\right\|_{C^0} \leq |u_1 - u_2| + \frac{1}{2} \left\|v - w\right\|_{C^0};$$

d.h.

$$\|v - w\|_{C^0} \leq 2\,|u_1 - u_2|\,.$$

Satz 6.3.4 zusammen mit (6.5.7) ergibt weiter für $0 < t < T$ die Abschätzung

$$\begin{aligned}|\dot{v}(t) - \dot{w}(t)| &= |f(t, v(t)) - f(t, w(t))| \\ &\leq L\,|v(t) - w(t)| \leq L\,\|v - w\|_{C^0}\,.\end{aligned}$$

Wir erhalten also

$$\begin{aligned}\|v - w\|_{C^1} &\leq \|v - w_0\|_{C^0} + L\,\|v - w\|_{C^0} \\ &= (1 + L)\,\|v - w_0\|_{C^0} \leq 2(1 + L)\,|u_1 - u_2|\,,\end{aligned}$$

und somit Aussage ii).

□

Was kann man über den Verlauf der Lösungen "im Grossen" sagen?

Satz 6.5.3. *Sei $f\colon \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ wie in Satz 6.5.1, und zu $u_0 \in \mathbb{R}^n$ sei $u = u(t; u_0) \in C^1([0, T]; \mathbb{R}^n)$ die eindeutig bestimmte Lösung des Anfangswertproblems (6.5.1), (6.5.2) gemäss Satz 6.5.1. Dann gibt es ein maximales $T_{max} > T$, so dass u fortgesetzt werden kann zu einer Lösung $u_{max} \in C^1([0, T_{max}[; \mathbb{R}^n)$ von (6.5.1),(6.5.2), und entweder gilt*

$$T_{max} = \infty,$$

oder

$$|u_{max}(t)| \to \infty \quad (t \to T_{max}).$$

Beweis. Setze

$$T_{max} = \sup\{T;\ \exists u \in C^1([0, T]; \mathbb{R}^n) \text{ mit } (6.5.1), (6.5.2)\}.$$

Wegen der Eindeutigkeitsaussage in Satz 6.5.1 stimmen je zwei Lösungen $u^{(i)} \in C^1([0, T_i]; \mathbb{R}^n)$, $i = 1, 2$, auf ihrem gemeinsamen Definitionsbereich $[0, T_1] \cap [0, T_2]$ überein. Somit ist für $0 \leq t < T_{max}$ die Funktion $u_{max}(t) := u(t)$ wohldefiniert, wobei $u \in C^1([0, T]; \mathbb{R}^n)$ eine beliebige Lösung von (6.5.1),(6.5.2) ist auf einem Intervall $[0, T]$ mit $T \geq t$, und u_{max} löst (6.5.1),(6.5.2) auf $[0, T_{max}[$.

Nimm an, $T_{max} < \infty$. Falls wir widerspruchsweise annehmen, dass

$$\liminf_{t \uparrow T_{max}} |u_{max}(t)| < \infty,$$

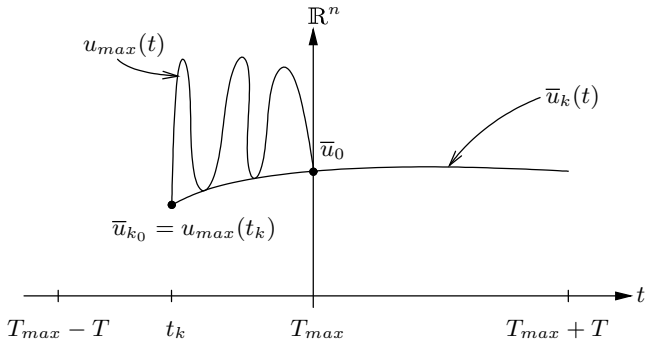
so gibt es $(t_k)_{k \in \mathbb{N}}$ mit

$$\overline{u}_{k_0} = u_{max}(t_k) \to \overline{u}_0 \quad (k \to \infty)$$

für ein $\overline{u}_0 \in \mathbb{R}^n$. Wählen wir im Beweis von Satz 6.5.1 die Konstanten $r_0 > 0$, $T_0 > 0$, $L \in \mathbb{R}$ so, dass (6.5.7) gilt für alle $v, w \in B_{r_0}(\overline{u}_0)$, $t \in \mathbb{R}$ mit $|t - T_{max}| \leq T_0$, so liefert der Beweis ein von k unabhängiges $T > 0$ und Lösungen $\overline{u}_k \in C^1([t_k, T_{max} + T])$ des Anfangswertproblems (6.5.1) mit $\overline{u}_k(t_k) = \overline{u}_{k_0}$, $k \geq k_0$, sofern $k_0 \in \mathbb{N}$ so gewählt ist, dass

$$t_k \geq T_{max} - T,\ \overline{u}_{k_0} \in B_{r_0/2}(\overline{u}_0) \text{ für alle } k \geq k_0.$$

Wegen der Eindeutigkeit der Lösung u des Anfangswertproblems (6.5.1) mit Anfangswert $u(t_k) = \overline{u}_{k_0}$ stimmt $\overline{u}_k$ für $k \geq k_0$ auf $[t_k, T_{max}[$ überein mit u_{max}. Wir können daher u_{max} durch $\overline{u}_{k_0}$ auf das Intervall $[0, T_{max} + T]$ fortsetzen, im Widerspruch zur angenommenen Maximalität von T_{max}. □



Bemerkung 6.5.1. In jedem Fall ist gemäss Satz 6.5.3 das maximale Existenzintervall $[0, T_{max}[$ der Lösung u von (6.5.1),(6.5.2) rechtsseitig offen.

Beispiel 6.5.6. i) Sei $f(t,u) = A(t)u + b(t)$ mit stetigen Koeffizientenfunktionen $A \in C^0(\mathbb{R}; \mathbb{R}^{n\times n})$, $b \in C^0(\mathbb{R}; \mathbb{R}^n)$. Dann besitzt das Anfangswertproblem (6.5.1),(6.5.2) für jedes $u_0 \in \mathbb{R}^n$ eine eindeutig bestimmte "globale" Lösung $u = u(t; u_0) \in C^1(\mathbb{R}; \mathbb{R}^n)$.

Beweis. Nimm an, das maximale Existenzintervall $[0, T_{max}[$ für u wäre endlich. Fixiere ein $T \geq T_{max}$. Schätze ab

$$v \cdot f(t,v) \leq \sup_{0\leq t\leq T} |A(t)|\,|v|^2 + \sup_{0\leq t\leq T} |b(t)|\,|v| \leq C(1 + |v|^2)$$

für alle $0 \leq t \leq T$, $v \in \mathbb{R}^n$. Es folgt

$$\frac{1}{2}\frac{d}{dt}(1 + |u(t)|^2) = (u \cdot \frac{d}{dt}u)(t) = u(t) \cdot f(t, u(t)) \leq C(1 + |u(t)|^2);$$

d.h.

$$\frac{d}{dt}\big(\log(1 + |u(t)|^2) - 2Ct\big) \leq 0, \quad 0 \leq t < T_{max}.$$

Korollar 5.2.1 ergibt nun für alle $0 \leq t < T_{max}$ die Abschätzung

$$\log\big(1 + |u(t)|^2\big) \leq 2Ct + \log(1 + |u_0|^2)\ ;$$

insbesondere erhalten wir die gleichmässige Abschätzung

$$1 + |u(t)|^2 \leq (1 + |u_0|^2)e^{2CT_{max}}$$

im Widerspruch zur erwarteten Divergenz

$$|u(t)| \to \infty \quad (t \uparrow T_{max})$$

gemäss Satz 6.5.3. □

ii) Lorenz-Attraktor. Sei $f = f(u)\colon \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch

$$f(u) = \begin{pmatrix} a(u_2 - u_1) \\ bu_1 - u_2 - u_1 u_3 \\ u_1 u_2 - cu_3 \end{pmatrix}$$

für $u = (u_1, u_2, u_3) \in \mathbb{R}^3$, mit Konstanten $a, b, c \in \mathbb{R}$. Dann gilt

$$u \cdot f(u) = (a+b)u_1 u_2 - (au_1^2 + u_2^2 + cu_3^2) \leq C\,|u|^2\,, \quad \forall u \in \mathbb{R}^3.$$

Also sind wie in Beispiel i) die Lösungen auf ganz $\mathbb{R}$ fortsetzbar.

Speziell für die Wahl

$$a = 10,\ b = 28,\ c = 8/3$$

streben alle Lösungen für $t \to \infty$ hin zu einem kompakten "Attraktor" K, wobei das Langzeitverhalten der Bahnen $u(t; u_0)$ sehr empfindlich auf kleinste Variationen des Startwerts u_0 reagiert. Da nach Satz 6.5.1 die Lösungen auf jedem kompakten Zeitintervall stetig vom Anfangswert abhängen, spricht man von "deterministischem Chaos".

Dieses System enstand als einfaches Modell des globalen Klimageschehens; die Sensitivität bzgl. der Daten führte zur Bezeichnung "Schmetterlingseffekt".

Unter der Adresse http://www.scu.org/∼bm733/attractor.html findet man eine Java-Animation im web.

Kapitel 7

Differentialrechnung im $\mathbb{R}^n$

7.1 Partielle Ableitungen und Differential

Wie kann man die Konzepte der Differentialrechnung in einer reellen Variablen auf Funktionen $f\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}$ erweitern?

Beispiel 7.1.1. Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ gegeben durch

$$f(x,y) = x\ e^y, \quad x,y \in \mathbb{R},$$

und sei $(x_0,y_0) \in \mathbb{R}^2$. Fassen wir $y \in \mathbb{R}$ als Parameter einer Schar von Funktionen

$$f(\cdot,y)\colon \mathbb{R} \to \mathbb{R}$$

auf, so können wir f für festes $y = y_0 \in \mathbb{R}$ "partiell" nach x differentieren und erhalten so die "partielle Ableitung"

$$\begin{aligned} f_x(x_0,y_0) = \frac{\partial f}{\partial x}(x_0,y_0) &= \lim_{x\to x_0\ x\neq x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0} \\ &= \lim_{x\to x_0\ x\neq x_0} \frac{x\ e^{y_0} - x_0\ e^{y_0}}{x-x_0} = e^{y_0}; \end{aligned}$$

ebenso

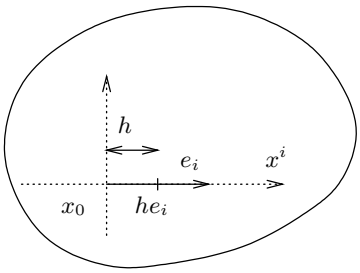
$$f_y(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = x_0\ e^{y_0}.$$

Sei $\Omega \subset \mathbb{R}^n$ offen, $x_0 = (x_0^1,\dots,x_0^n) \in \Omega$. Allgemein definieren wir:

Definition 7.1.1. *Die Funktion $f\colon \Omega \to \mathbb{R}$ heisst an der Stelle x_0 in Richtung $e_i = (0,\dots,0,\underbrace{1}_{i\text{-te Stelle}},0,\dots,0)$ (bzw. nach x^i)* **partiell differenzierbar**, *falls der Limes*

$$\frac{\partial f}{\partial x^i}(x_0) = f_{x^i}(x_0) = \lim_{h\to 0,\ h\neq 0} \frac{f(x_0 + he_i) - f(x_0)}{h}$$
$$= \lim_{h\to 0,\ h\neq 0} \frac{f(x_0^1, \ldots, x_0^i + h, \ldots, x_0^n) - f(x_0^1, \ldots, x_0^i, \ldots, x_0^n)}{h}$$

existiert.



Notation: Von nun ab ist es zweckmässig, die Komponenten des Ortsvektors $x = (x^i)_{1\leq i\leq n}$ mit hochgestelltem Index zu schreiben. Wir werden bald erkennen, welche Vorteile dies bietet.

In einer Raumdimension ($n = 1$) hat die Differenzierbarkeit der Funktion f an einer Stelle x_0 zur Folge, dass f für x nahe x_0 gut durch die affin-lineare Funktion

$$x \mapsto T_1 f(x; x_0) = f(x_0) + f'(x_0)(x - x_0)$$

angenähert wird: Aus

$$\lim_{x\to x_0,\ x\neq x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

erhalten wir sofort

$$\lim_{x\to x_0,\ x\neq x_0} \frac{|f(x) - (f(x_0) + f'(x_0)(x - x_0))|}{x - x_0} = 0.$$

Gilt eine vergleichbare Approximationseigenschaft auch für $n > 1$?

Beispiel 7.1.2. i) Sei $f(x, y) = x\, e^y$ wie in Beispiel 7.1.1 und sei $(x_0, y_0) \in \mathbb{R}^2$ gegeben.

Für $(x, y) \in \mathbb{R}^2$ erhalten wir mit Korollar 5.2.1

$$\begin{aligned} f(x,y) - f(x_0,y_0) &= f(x,y) - f(x_0,y) + f(x_0,y) - f(x_0,y_0) \\ &= \frac{\partial f}{\partial x}(\xi(y), y)(x - x_0) + \frac{\partial f}{\partial y}(x_0, \eta)(y - y_0) \\ &= \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0) + R(x,y) \end{aligned}$$

mit geeigneten Zwischenstellen $\xi = \xi(y)$, η und Restterm

$$R(x,y) = \Big(\frac{\partial f}{\partial x}(\xi(y), y) - \frac{\partial f}{\partial x}(x_0,y_0)\Big)(x - x_0) + \Big(\frac{\partial f}{\partial y}(x_0,\eta) - \frac{\partial f}{\partial y}(x_0,y_0)\Big)(y - y_0).$$

Wegen der Stetigkeit der Funktionen

$$\frac{\partial f}{\partial x}(x,y) = e^y, \quad \frac{\partial f}{\partial y}(x,y) = x\ e^y$$

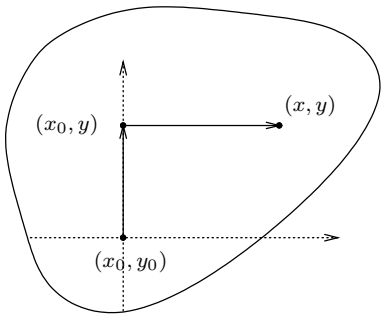
können wir den "Fehler" $R(x,y)$ leicht abschätzen

$$\frac{|R(x,y)|}{|x-x_0|+|y-y_0|} \le \sup_{\substack{|\xi-x_0|<|x-x_0| \\ |\eta-y_0|<|y-y_0|}} \Big(|e^y - e^{y_0}| + |x_0|\,|e^\eta - e^{y_0}| \Big) \to 0$$

für $(x,y) \to (x_0,y_0)$, $(x,y) \neq (x_0,y_0)$; d.h., es gilt

$$\frac{f(x,y) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) - \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)}{|x-x_0|+|y-y_0|} \to 0 \qquad (7.1.1)$$

für $(x,y) \to (x_0,y_0)$, $(x,y) \neq (x_0,y_0)$.



ii) Sei $f\colon \mathbb{R}^2 \to \mathbb{R}$ die Funktion

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

Offenbar ist f an jeder Stelle $(x_0,y_0) \in \mathbb{R}^2$ partiell nach x und y differenzierbar. Insbesondere gilt

$$\frac{\partial f}{\partial x}(0,0) = 0 = \frac{\partial f}{\partial y}(0,0).$$

Jedoch gilt beispielsweise

$$f(x,x) = 1 \not\to f(0,0) = 0 \quad (x \to 0,\ x \neq 0);$$

die Funktion f ist also bei $(x_0,y_0) = (0,0)$ noch nicht einmal stetig; schon gar nicht kann man die Approximationseigenschaft (7.1.1) erwarten.

Definition 7.1.2. *Die Funktion $f\colon \Omega \to \mathbb{R}$ heisst an der Stelle $x_0 \in \Omega$* **differenzierbar**, *falls eine lineare Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}$ existiert mit*

$$\lim_{x \to x_0,\ x \neq x_0} \frac{f(x) - f(x_0) - A(x-x_0)}{|x-x_0|} = 0.$$

In diesem Fall heisst $df(x_0) := A$ das **Differential** *von f an der Stelle x_0.*

Bemerkung 7.1.1. i) Insbesondere ist f in diesem Fall stetig an der Stelle x_0 und es existieren sämtliche partiellen Ableitungen $\frac{\partial f}{\partial x_i}(x_0) = Ae_i$, $1 \leq i \leq n$; die Umkehrung gilt aber nicht (vgl. Beispiel 7.1.2).

ii) Ist f an der Stelle x_0 differenzierbar, so gilt

$$df(x_0) = \Big(\frac{\partial f}{\partial x^1}(x_0), \ldots, \frac{\partial f}{\partial x^n}(x_0)\Big).$$

Notation: Es zeigt sich nun, dass es auch vorteilhaft ist, Zeilen- und Spaltenvektoren zu unterscheiden. Schreiben wir nämlich $x = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ für einen Vektor $x = (x^i)_{1\leq i\leq n} \in \mathbb{R}^n$ und $A = (A_1, \ldots, A_n)$ für die Darstellung einer linearen Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}$ bezüglich der Standardbasis, so ist

$$A(x - x_0) = \sum_{i=1}^{n} A_i(x^i - x_0^i) = (A_1, \ldots, A_n) \begin{pmatrix} x^1 - x_0^1 \\ \vdots \\ x^n - x_0^n \end{pmatrix}$$

interpretierbar als Matrixmultiplikation des **co-Vektors** $A = (A_1, \ldots, A_n)$ mit dem **Vektor** $x - x_0$.

Diese Schreibweise lädt ein zur **Einsteinschen Summenkonvention:** Über doppelt auftretende obere und untere Indizes wird stillschweigend summiert.

Beispiel 7.1.3. i) Jede affin lineare Funktion $f(x) = Ax + b$, $x \in \mathbb{R}^n$, ist an jeder Stelle $x_0 \in \mathbb{R}^n$ differenzierbar mit $df(x_0) = A$.

Beweis: $f(x) - f(x_0) - A(x - x_0) = 0$, $\forall x, x_0 \in \mathbb{R}^n$.

ii) Insbesondere sind die **Koordinatenfunktionen** $x^i :\ x = (x^k)_{1\leq k\leq n} \mapsto x^i$ an jeder Stelle $x_0 \in \mathbb{R}^n$ differenzierbar mit

$$dx^i\big|_{x=x_0} = (0, \ldots, 0, \overbrace{1}^{i\text{-te Stelle}}, 0, \ldots, 0),\ 1 \leq i \leq n.$$

Die Differentiale $dx^1, \ldots, dx^n$ bilden also an jeder Stelle $x_0 \in \mathbb{R}^n$ eine Basis des Raumes

$$L(\mathbb{R}^n; \mathbb{R}) = \{A\colon \mathbb{R}^n \to \mathbb{R};\ A \text{ linear}\},$$

wobei wir $A \in L(\mathbb{R}^n; \mathbb{R})$ mit der Darstellung $A = (A_1, \ldots, A_n)$ bzgl. der Standardbasis $e_1, \ldots, e_n$ des $\mathbb{R}^n$ identifizieren, und mit $A_i = Ae_i$, $1 \leq i \leq n$.

Da offenbar gilt

$$dx^i e_j = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

ist $(dx^i)_{1\leq i\leq n}$ sogar die zur Standardbasis $(e_i)_{1\leq i\leq n}$ des $\mathbb{R}^n$ **duale Basis** von $L(\mathbb{R}^n; \mathbb{R})$.

iii) Jedes $f \in C^1(\mathbb{R})$ besitzt das Differential

$$df(x_0) = \frac{df}{dx}(x_0)dx = f'(x_0)dx;$$

d.h. $f'(x_0)$ ist die Darstellung von $df(x_0)$ bzgl. der Basis dx von $L(\mathbb{R};\mathbb{R})$.

iv) Die Funktion $f(x,y) = xe^y \colon \mathbb{R}^2 \to \mathbb{R}$ ist an jeder Stelle $(x_0,y_0) \in \mathbb{R}^2$ gemäss Beispiel 7.1.2.i) differenzierbar, und es gilt

$$df(x_0,y_0) = \Big(\frac{\partial f}{\partial x}(x_0,y_0), \frac{\partial f}{\partial y}(x_0,y_0)\Big).$$

(Eigentlich müssten wir auch hier und im folgenden $\begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$ anstelle von (x_0,y_0) schreiben; dies wäre aber doch zu umständlich!)

v) Die Funktion aus Beispiel 7.1.2.ii) ist an der Stelle $(x_0,y_0) = (0,0)$ nicht differenzierbar.

Was macht den Unterschied aus zwischen den Beispielen 7.1.2.i) und ii)?

Definition 7.1.3. *Die Funktion $f\colon \Omega \to \mathbb{R}$ heisst* **von der Klasse C^1**, *$f \in C^1(\Omega)$, falls f an jeder Stelle $x_0 \in \mathbb{R}^n$ in jede Richtung e_i partiell differenzierbar ist, und falls die Funktionen $x \mapsto \frac{\partial f}{\partial x^i}(x)$, $1 \le i \le n$, auf Ω stetig sind.*

Beispiel 7.1.4. Die Funktion $f(x,y) = x\ e^y \colon \mathbb{R}^2 \to \mathbb{R}$ aus Beispiel 7.1.2.i) ist von der Klasse C^1, die Funktion aus Beispiel 7.1.2.ii) nicht.

Satz 7.1.1. *Sei $f \in C^1(\Omega)$. Dann ist f an jeder Stelle $x_0 \in \Omega$ differenzierbar. Insbesondere ist f auch stetig auf Ω.*

Beweis. Der Beweis folgt dem Vorgehen in Beispiel 7.1.2.i). Für $x = (x^i)_{1\le i\le n}$ schätze ab mit Satz 5.2.1

$$\begin{aligned} &f(x) - f(x_0) - df(x_0)(x-x_0) \\ &= \sum_{i=1}^n \Big(f(x^1,\dots,x^i,x_0^{i+1},\dots,x_0^n) - f(x^1,\dots,x_0^i,\dots,x_0^n) - \frac{\partial f}{\partial x^i}(x_0)(x^i - x_0^i)\Big) \\ &= \sum_{i=1}^n \Big(\frac{\partial f}{\partial x^i}(x^1,\dots,\xi^i,x_0^{i+1},\dots,x_0^n) - \frac{\partial f}{\partial x^i}(x_0)\Big)(x^i - x_0^i) \end{aligned}$$

für geeignete Punkte ξ^i zwischen x_0^i und x^i, $1 \le i \le n$. Wir erhalten somit für $x \ne x_0$ die Abschätzung

$$\frac{f(x) - f(x_0) - df(x_0)(x-x_0)}{|x-x_0|} \le n \sup_{|z^i - x_0^i| < |x^i - x_0^i|} \Big| \frac{\partial f}{\partial x^i}(z) - \frac{\partial f}{\partial x^i}(x_0)\Big| ,$$

und die rechte Seite strebt wegen der Stetigkeit der partiellen Ableitungen gegen 0 mit $x \to x_0$. D.h., f ist an der Stelle $x_0 \in \Omega$ differenzierbar. □

Beispiel 7.1.5. i) Polynome auf $\mathbb{R}^n$ sind von der Klasse C^1. Eine handliche Notation erhält man mit **Multi-Indices** $\alpha = (\alpha_1,\dots,\alpha_n) \in \mathbb{N}_0^n$, indem man für $x = (x^i)_{1\le i\le n} \in \mathbb{R}^n$ setzt

$$x^\alpha := \prod_{i=1}^n (x^i)^{\alpha_i}.$$

Somit kann man ein Polynom $p\colon \mathbb{R}^n \to \mathbb{R}$ vom Grad N in der Form schreiben

$$p(x) = \sum_{|\alpha| \le N} a_\alpha x^\alpha, \quad x \in \mathbb{R}^n,$$

wobei $|\alpha| = \alpha_1 + \cdots + \alpha_n$.

ii) Rationale Funktionen $r = p/q$ sind von der Klasse C^1 auf ihrem natürlichen Definitionsbereich $\Omega = \{x;\ q(x) \neq 0\}$.

Schliesslich definieren wir für beschränktes, offenes $\Omega \subset \mathbb{R}^n$ analog zu Abschnitt 5 den Raum

$$C^1(\overline{\Omega}) = \{f \in C^1(\Omega);\ f \text{ und } \frac{\partial f}{\partial x^i} \text{ sind stetig auf } \overline{\Omega} \text{ ergänzbar}, 1 \le i \le n\}$$

mit der Norm

$$\|f\|_{C^1} = \|f\|_{C^0} + \sum_{i=1}^n \left\| \frac{\partial f}{\partial x^i} \right\|_{C^0}.$$

Völlig analog zu Satz 5.4.3 zeigt man, dass $C^1(\overline{\Omega})$ metrisch vollständig ist bzgl. der Norm $\|\cdot\|_{C^1}$; der Raum $C^1(\overline{\Omega})$ ist also ein Banachraum.

7.2 Differentiationsregeln

Sei $\Omega \subset \mathbb{R}^n$ offen.

Satz 7.2.1. *Seien $f, g\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar. Dann sind auch die Funktionen $f+g$ und $f \cdot g$ an der Stelle x_0 differenzierbar, und es gilt*

i) $d(f+g)(x_0) = df(x_0) + dg(x_0)$,

ii) $d(f \cdot g)(x_0) = g(x_0)df(x_0) + f(x_0)dg(x_0)$,

sowie -falls $g(x_0) \neq 0$- auch f/g mit

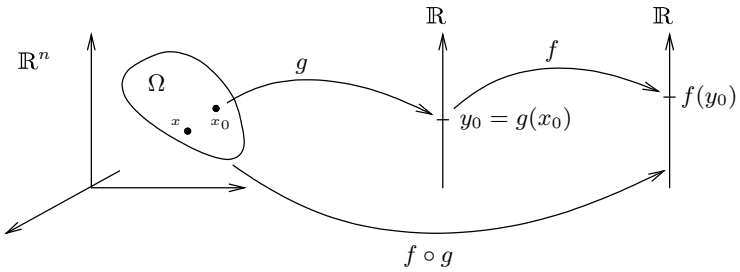
iii)

$$d\Big(\frac{f}{g}\Big)(x_0) = \frac{g(x_0)df(x_0) - f(x_0)dg(x_0)}{(g(x_0))^2}.$$

Beweis. Analog zu Satz 5.1.2. □

Satz 7.2.2. (Kettenregel, 1. Version) *Sei $g\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar, und sei $f\colon \mathbb{R} \to \mathbb{R}$ differenzierbar an der Stelle $g(x_0)$. Dann ist die Funktion $f \circ g\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar, und es gilt*

$$d(f \circ g)(x_0) = f'(g(x_0))dg(x_0).$$



Beweis. Für $x \to x_0$, $x \in \Omega$, gilt $g(x) \to g(x_0)$. Da f bei $g(x_0)$ differenzierbar, folgt

$$f(g(x)) - f(g(x_0)) - f'(g(x_0))\big(g(x) - g(x_0)\big) = R_f(x, x_0)\big(g(x) - g(x_0)\big)$$

mit

$$R_f(x, x_0) \to 0 \quad (x \to x_0).$$

Ebenso gilt

$$g(x) - g(x_0) - dg(x_0)(x - x_0) = R_g(x, x_0)(x - x_0)$$

mit

$$R_g(x, x_0) \to 0 \quad (x \to x_0).$$

Insbesondere erhalten wir $C \in \mathbb{R}$ mit

$$|g(x) - g(x_0)| \le |dg(x_0)(x - x_0)| + |R_g(x, x_0)(x - x_0)| \le C\,|x - x_0|$$

für $x \in \Omega$ nahe x_0. Schätze ab

$$\begin{aligned}
&|(f \circ g)(x) - (f \circ g)(x_0) - f'(g(x_0))dg(x_0)(x - x_0)| \\
&\qquad = |f(g(x)) - f(g(x_0)) - f'(g(x_0))(g(x) - g(x_0)) \\
&\qquad\qquad + f'(g(x_0))\big(g(x) - g(x_0) - dg(x_0)(x - x_0)\big)| \\
&\qquad \le |R_f(x, x_0)(g(x) - g(x_0))| + |f'(g(x_0))|\,|R_g(x, x_0)(x - x_0)| \\
&\qquad \le C\big(|R_f(x, x_0)| + |R_g(x, x_0)|\big)\,|x - x_0|\,.
\end{aligned}$$

Die Behauptung folgt, da $R_f(x, x_0), R_g(x, x_0) \to 0$ für $x \to x_0$. □

Beispiel 7.2.1. Sei $h\colon \mathbb{R}^2 \to \mathbb{R}$ mit

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

Schreibe $h = f \circ g$ mit $f = \exp$ und $g(x, y) = xy$. Mittels direkter Rechnung erhalten wir

$$dh(x, y) = (ye^{xy}, xe^{xy})\,.$$

Dasselbe Resultat erhalten wir durch Anwendung von Satz 7.2.2 in der Form

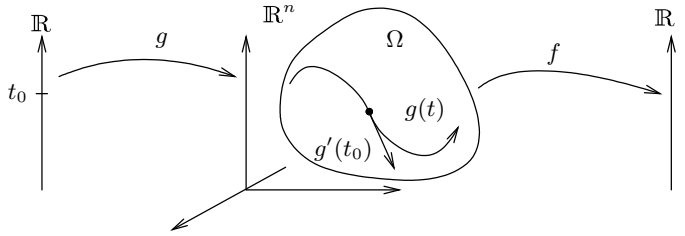
$$dh(x, y) = f'(g(x, y)) \cdot dg(x, y) = e^{xy} \cdot (y, x)\,.$$

Satz 7.2.3. (Kettenregel, 2. Teil) *Sei $\Omega \subset \mathbb{R}^n$ offen. Sei $g\colon I \subset \mathbb{R} \to \Omega$ an der Stelle $t_0 \in I$ differenzierbar, und sei $f\colon \Omega \to \mathbb{R}$ an der Stelle $x_0 = g(t_0)$ differenzierbar. Dann ist die Funktion $f \circ g\colon I \to \mathbb{R}$ an der Stelle t_0 differenzierbar, und es gilt*

$$\frac{d}{dt}(f \circ g)(t_0) = df(g(t_0))\ \frac{dg}{dt}(t_0),$$

oder -dazu äquivalent-

$$d(f \circ g)(t_0) = df(g(t_0))\ dg(t_0).$$



Bemerkung 7.2.1. Im ersten Fall deuten wir $\frac{dg}{dt}(t_0) \in \mathbb{R}^n$ als "Geschwindigkeitsvektor" der Kurve $t \mapsto g(t)$ zur Zeit t_0, auf den die lineare Abbildung $df(g(t_0))\colon \mathbb{R}^n \to \mathbb{R}$ wirkt. Im zweiten Fall deuten wir $dg(t_0)$ als Differential der vektorwertigen Funktion g, d.h. als lineare Abbildung $dg(t_0)\colon \mathbb{R} \to \mathbb{R}^n$, die wir mit der linearen Abbildung $df(g(t_0))\colon \mathbb{R}^n \to \mathbb{R}$ verknüpfen.

Beweis. Der Beweis ist vollkommen analog zum Beweis von Satz 7.2.2. □

Beispiel 7.2.2. i) Sei $f\colon \Omega \to \mathbb{R}$ differenzierbar an der Stelle $x_0 \in \Omega$, und sei $e \in \mathbb{R}^n \backslash \{0\}$. Betrachte die Gerade $g(t) = x_0 + te$, $t \in \mathbb{R}$, durch x_0 mit Richtungsvektor $\frac{dg}{dt}(t_0) = e$, $\forall t_0 \in \mathbb{R}$.

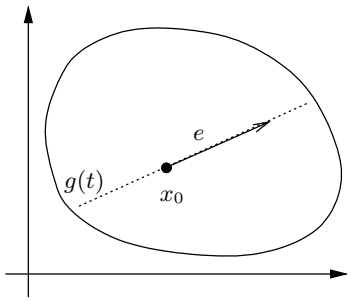
Dann ist die Funktion $f \circ g$ in einer Umgebung von $t_0 = 0$ definiert, und nach Satz 7.2.3 ist $f \circ g$ an der Stelle $t_0 = 0$ differenzierbar mit

$$\frac{d}{dt}(f \circ g)(0) = df(g(0))\frac{dg}{dt}(0) = df(x_0)e.$$

Wir deuten den Ausdruck $df(x_0)e$ als **Richtungsableitung** von f in Richtung e. Für $e = e_i$ ergibt sich insbesondere wieder

$$\frac{\partial f}{\partial x^i}(x_0) = df(x_0)e_i,$$

in Übereinstimmung mit Bemerkung 7.1.1.

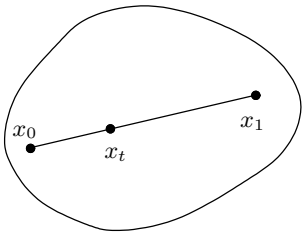


ii) Sei $f\colon B_r(0) \subset \mathbb{R}^n \to \mathbb{R}$ differenzierbar, und seien $x_0, x_1 \in B_r(0)$. Dann existiert $0 < \vartheta < 1$ mit

$$f(x_1) - f(x_0) = df(x_\vartheta)(x_1 - x_0),$$

wobei

$$x_t = (1-t)x_0 + tx_1, \quad 0 \le t \le 1.$$



Beweis: Setze $g(t) = x_t$, $0 \le t \le 1$. Dann ist nach Satz 4.2.1 und Satz 7.2.3 die Funktion $f \circ g\colon [0,1] \to \mathbb{R}$ stetig und in $]0,1[$ differenzierbar. Gemäss Satz 5.2.1 gibt es $\vartheta \in]0,1[$ mit

$$\begin{aligned} f(x_1) - f(x_0) = f(g(1)) - f(g(0)) &= \frac{d}{dt}(f \circ g)(\vartheta) \\ &= df(g(\vartheta))\frac{dg}{dt}(\vartheta) = df(x_\vartheta)(x_1 - x_0). \end{aligned}$$

iii) Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega)$. Dann ist f lokal Lipschitz stetig.

Beweis: Sei $x_0 \in \Omega$, dazu $r > 0$ mit $B_r(x_0) \subset \Omega$. Nach ii) gilt mit $L = \sup_{x \in B_r(x_0)} |df(x)|$

$$|f(y) - f(x)| \le L\,|x - y|\,, \quad \forall x, y \in B_r(x_0).$$

iv) Sei $f(x,y) = x^2 + y^2$, $(x,y) \in \mathbb{R}^2$, und sei $g\colon \mathbb{R} \to \mathbb{R}^2$ die Kurve

$$g(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \ t \in \mathbb{R}.$$

Dann gilt

$$(f \circ g)(t) = \cos^2(t) + \sin^2(t) = 1, \quad \forall t \in \mathbb{R},$$

also

$$\frac{d}{dt}(f \circ g)(t) = 0, \quad \forall t \in \mathbb{R}.$$

Dasselbe Ergebnis erhalten wir auch mit Satz 7.2.3, denn

$$\begin{aligned}\frac{d}{dt}(f \circ g)(t) &= df(g(t))\frac{dg}{dt}(t) \\ &= (2x, 2y)\Big|_{(x,y)=g(t)} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \\ &= 2(\cos t, \sin t) \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = 0, \quad \forall t \in \mathbb{R}.\end{aligned}$$

Integrale mit Parametern. Sei $h = h(s,t)\colon \mathbb{R}^2 \to \mathbb{R}$ stetig und bzgl. t partiell differenzierbar mit $\frac{\partial h}{\partial t} \in C^0(\mathbb{R}^2)$. Setze

$$u(t) = \int_0^t h(s,t)\ ds, \quad t \in \mathbb{R}.$$

Fragen i) Ist $u \in C^1$?

ii) Wie erhält man in diesem Fall $\dot{u}$?

Deute $u = f \circ g$ mit

$$\begin{aligned}f(x,y) &= \int_0^x h(s,y)\ ds\colon \mathbb{R}^2 \to \mathbb{R}, \\ g(t) &= \begin{pmatrix} t \\ t \end{pmatrix} : \mathbb{R} \to \mathbb{R}^2.\end{aligned}$$

Offenbar ist $g \in C^1(\mathbb{R}, \mathbb{R}^2)$, und nach Satz 6.3.4 ist f partiell nach x differenzierbar mit

$$\frac{\partial f}{\partial x}(x,y) = h(x,y) \in C^0(\mathbb{R}^2).$$

Behauptung f ist partiell nach y differenzierbar mit

$$\frac{\partial f}{\partial y}(x,y) = \int_0^x \frac{\partial h}{\partial y}(s,y)\ ds \in C^0(\mathbb{R}^2).$$

Beweis. Für festes $x \in \mathbb{R}$ und $y_0, y \in \mathbb{R}$ gilt

$$\begin{aligned}f(x,y) - f(x,y_0) &= \int_0^x (h(s,y) - h(s,y_0))\ ds \\ &= \int_0^x \frac{\partial h}{\partial y}(s, y(s))(y - y_0)\ ds\end{aligned}$$

mit Zwischenstellen $y(s)$ zwischen y_0 und y gemäss Satz 5.2.1. Mit Korollar 6.3.1 folgt

$$\begin{aligned}\left|\frac{f(x,y) - f(x,y_0)}{y - y_0} - \int_0^x \frac{\partial h}{\partial y}(x,y_0)\ ds\right| &\le \int_0^x \left|\frac{\partial h}{\partial y}(x, y(s)) - \frac{\partial h}{\partial y}(s,y_0)\right|\ ds \\ &\le x \sup_{0 \le s \le x, |\eta - y_0| < |y - y_0|} \left|\frac{\partial h}{\partial y}(s,\eta) - \frac{\partial h}{\partial y}(s,y_0)\right| \to 0 \quad (y \to y_0),\end{aligned}$$

da $\frac{\partial h}{\partial y}$ auf dem kompakten Streifen $\{(s,y);\ 0 \le s \le x;\ |y-y_0| \le 1\}$ gemäss Satz 4.7.3 gleichmässig stetig ist.

Ebenso erhält man die Stetigkeit von $\frac{\partial f}{\partial y}$ aus der Abschätzung

$$\begin{aligned}
&\left|\frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(x_0,y_0)\right| \le \left|\frac{\partial f}{\partial y}(x,y) - \frac{\partial f}{\partial y}(x_0,y)\right| + \left|\frac{\partial f}{\partial y}(x_0,y) - \frac{\partial f}{\partial y}(x_0,y_0)\right| \\
&= \left|\int_{x_0}^{x} \frac{\partial h}{\partial y}(s,y)\ ds\right| + \left|\int_0^{x_0} \Big(\frac{\partial h}{\partial y}(s,y) - \frac{\partial h}{\partial y}(s,y_0)\Big)\ ds\right| \\
&\le \sup_{|s-x_0|\le 1,\ |y-y_0|\le 1} \left|\frac{\partial h}{\partial y}(s,y)\right| \cdot |x-x_0| + x_0 \sup_{0\le s\le x_0} \left|\frac{\partial h}{\partial y}(s,y) - \frac{\partial h}{\partial y}(s,y_0)\right| \\
&\to 0 \quad (x \to x_0,\ y \to y_0)\ .
\end{aligned}$$

□

Satz 7.2.3 ergibt somit $u \in C^1$ mit

$$\begin{aligned}
\dot u(t) &= \frac{d}{dt}(f\circ g)(t) = \Big(\frac{\partial f}{\partial x}(g(t)), \frac{\partial f}{\partial y}(g(t))\Big)\frac{dg}{dt}(t) \\
&= \Big(h(t,t), \int_0^t \frac{\partial h}{\partial t}(s,t)\ ds\Big)\begin{pmatrix}1\\1\end{pmatrix} = h(t,t) + \int_0^t \frac{\partial h}{\partial t}(s,t)\ ds.
\end{aligned} \tag{7.2.1}$$

Beispiel 7.2.3. i) Sei $b \in C^0(\mathbb{R})$, und setze an

$$u(t) = \int_0^t e^{s-t} b(s)\ ds.$$

Die Funktion

$$h(s,t) = e^{s-t}b(s)$$

ist stetig, nach $t \in \mathbb{R}$ partiell differenzierbar mit stetiger Ableitungsfunktion

$$\frac{\partial h}{\partial t}(s,t) = -e^{s-t}b(s) = -h(s,t).$$

Es folgt, $u \in C^1$ mit

$$\dot u(t) = h(t,t) + \int_0^t \frac{\partial h}{\partial t}(s,t)\ ds = b(t) - u;$$

vgl. Beispiel 6.1.5.vi) mit $a = -1$. Beachte, dass

$$\Phi(s,t) = e^{s-t}$$

die Lösung ist des Anfangswertproblems

$$\dot u = -u, \quad u(s) = 1.$$

ii) Analog erhält man die allgemeine **Variation-der-Konstanten Formel** für eine partikuläre Lösung $u = u(t) \in C^1(\mathbb{R};\mathbb{R}^n)$ der inhomogenen linearen DGl

$$\dot u(t) = A(t)u(t) + b(t) \tag{7.2.2}$$

mit $A \in C^0(\mathbb{R}; \mathbb{R}^{n\times n})$, $b \in C^0(\mathbb{R}; \mathbb{R}^n)$. Nach Beispiel 6.5.6.i) besitzt das Anfangswertproblem

$$\frac{d\Phi}{dt}(t;s) = A(t)\Phi(t;s), \quad \Phi(s;s) = id \tag{7.2.3}$$

für alle $s \in \mathbb{R}$ eine eindeutige Lösung $\Phi(t) = \Phi(t;s) \in C^1(\mathbb{R}; \mathbb{R}^{n\times n})$. (Für $n = 1$, $A(t) = -1$ erhalten wir $\Phi(t;s) = e^{s-t}$ für alle $s, t \in \mathbb{R}$.)

Als **Ansatz** für eine Lösung von (7.2.2) wähle nun

$$u(t) = \int_0^t \Phi(t;s)b(s)\ ds, \quad t \in \mathbb{R}. \tag{7.2.4}$$

Analog zu i) erhalten wir $u \in C^1(\mathbb{R}; \mathbb{R}^n)$ mit

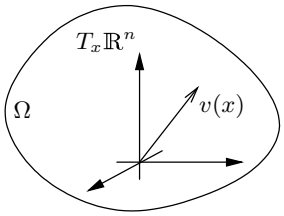
$$\begin{aligned} \dot{u}(t) &= \Phi(t;t)b(t) + \int_0^t \frac{d\Phi}{dt}(t;s)b(s)\ ds \\ &\overset{(7.2.3)}{=} b(t) + A(t)\int_0^t \Phi(t;s)b(s)\ ds = A(t)u(t) + b(t)\ , \end{aligned}$$

wie gewünscht.

Ebenso wie die Darstellung (6.1.5) einer partikulären Lösung im Falle $n = 1$ erhält man auch in Dimensionen $n > 1$ die Lösungsformel (7.2.4) aus dem Ansatz $u(t) = \Phi(t;0)c(t)$ mit variablem $c = c(t)$.

7.3 Differentialformen und Vektorfelder

In Abschnitt 6.5 haben wir bereits Funktionen $v\colon \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ als **Vektorfelder** gedeutet, wobei wir den Vektor $v(x) \in \mathbb{R}^n$ für jedes $x \in \Omega$ als einen von diesem Punkt ausgehenden Richtungsvektor auffassen, also als ein Element des **Tangentialraums** $T_x\mathbb{R}^n$ des $\mathbb{R}^n$ am Punkt x.



Analog können wir auch Abbildungen $\lambda\colon \Omega \subset \mathbb{R}^n \to L(\mathbb{R}^n; \mathbb{R})$ betrachten, welche jedem $x \in \Omega$ eine lineare Abbildung $\lambda(x)\colon T_x\mathbb{R}^n \cong \mathbb{R}^n \to \mathbb{R}$ zuordnen. Bzgl. der Basis $dx^1, \ldots, dx^n$ von $L(\mathbb{R}^n; \mathbb{R})$ schreiben wir

$$\lambda(x) = \sum_{i=1}^n \lambda_i(x)dx^i$$

und können jedes derartige λ so mit einer linearen Abbildung

$$\lambda = (\lambda_1, \ldots, \lambda_n)\colon \Omega \to \mathbb{R}^n$$

identifizieren.

Definition 7.3.1. *Eine Abbildung* $\lambda\colon \Omega \to L(\mathbb{R}^n;\mathbb{R})$ *heisst eine* **Differentialform** *vom Grad 1 (kurz* **1-Form** *oder Pfaffsche Form).*

Beispiel 7.3.1. i) Für jedes $f \in C^1(\Omega)$ ist das Differential df eine 1-Form (von der Klasse C^0).

ii) Der Ausdruck $\lambda(x,y,z) = 3dx + 2zdy + xydz$ definiert eine 1-Form auf $\mathbb{R}^3$.

Bemerkung 7.3.1. i) Mit Hilfe des Skalarprodukts $\langle\cdot,\cdot\rangle_{\mathbb{R}^n}$ kann man ein Vektorfeld $v = (v^i)_{1\le i\le n}\colon \Omega \to \mathbb{R}^n$ in eine 1-Form λ verwandeln. Setze dazu

$$\lambda(x)w := \langle v(x), w\rangle_{\mathbb{R}^n}\,, \quad \forall w \in T_x\mathbb{R}^n, \tag{7.3.1}$$

für jedes $x \in \Omega$. Bzgl. $dx^1,\dots,dx^n$ gilt $\lambda = (\lambda_i)_{1\le i\le n}$ mit

$$\lambda_i(x) = \lambda(x)e_i = \langle v(x), e_i\rangle_{\mathbb{R}^n} = v^i(x), \quad \forall x \in \Omega. \tag{7.3.2}$$

ii) Umgekehrt kann man via (7.3.1) auch 1-Formen λ auf Ω in Vektorfelder $v\colon \Omega \to \mathbb{R}^n$ umwandeln.

Speziell für $\lambda = df$ ergibt Bemerkung 7.3.1.ii) die folgende Definition

Definition 7.3.2. *Sei* $f \in C^1(\Omega)$*. Das durch die Gleichung*

$$\langle \nabla f(x), w\rangle_{\mathbb{R}^n} = df(x)w, \quad \forall w \in \mathbb{R}^n$$

definierte Vektorfeld $\nabla f\colon \Omega \to \mathbb{R}^n$ *heisst* **Gradientenfeld von** f **bzgl.** $\langle\cdot,\cdot\rangle_{\mathbb{R}^n}$.

Bzgl. der Standardbasis $e_1,\dots,e_n$ des $\mathbb{R}^n$ folgt mit (7.3.2) die Darstellung

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x^1}(x) \\ \vdots \\ \frac{\partial f}{\partial x^n}(x) \end{pmatrix}, \quad \forall x \in \Omega.$$

Bemerkung 7.3.2. Sei $f \in C^1(\Omega)$, und sei $x_0 \in \Omega$. Dann gibt $\nabla f(x_0)$ die Richtung und den Betrag des "steilsten Anstiegs" des Graphen $\mathcal{G}(f)$ an der Stelle x_0 an in dem Sinne, dass

$$df(x_0)\frac{\nabla f(x_0)}{|\nabla f(x_0)|} = |\nabla f(x_0)| = \max_{e\in T_{x_0}\mathbb{R}^n,\ |e|=1} df(x_0)e.$$

Beweis. Für jedes $e \in T_{x_0}\mathbb{R}^n \cong \mathbb{R}^n$ mit $|e| = 1$ gilt

$$\begin{aligned} df(x_0)e &\overset{\text{(Def)}}{=} \langle \nabla f x_0, e\rangle_{\mathbb{R}^n} \overset{\text{(Cauchy-Schwarz)}}{\le} |\nabla f(x_0)| \\ &= \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{|\nabla f(x_0)|}\right\rangle = df(x_0)\frac{\nabla f(x_0)}{|\nabla f(x_0)|}, \end{aligned}$$

mit Gleichheit für $e = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$. □

Beispiel 7.3.2. i) Sei $f(x,y) = \frac{x^2-y^2}{2}$, $(x,y) \in \mathbb{R}^2$, mit

$$df(x,y) = (x,-y), \quad \nabla f(x,y) = \begin{pmatrix} x \\ -y \end{pmatrix},$$

und sei $(x_0, y_0) = (1,-1)$ mit

$$|\nabla f(1,-1)| = \sqrt{2}, \ \Big(\frac{\nabla f}{|\nabla f|}\Big)(1,-1) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

ii) Sei $f \in C^1(\mathbb{R}^2)$ mit

$$df(0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = df(0)e_2 = \frac{\partial f}{\partial x^2}(0) = 3,$$
$$df(0) \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{\partial f}{\partial x^1}(0) - \frac{\partial f}{\partial x^2}(0) = 1.$$

Es folgt $\frac{\partial f}{\partial x^1}(0) = 4$; d.h.

$$\nabla f(0) = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

7.4 Wegintegrale

Sei $\Omega \subset \mathbb{R}^n$ offen, $\gamma\colon [0,1] \to \Omega$ ein "Weg" in Ω von der Klasse C^1, $\gamma \in C^1([0,1];\Omega)$, mit Geschwindigkeitsvektorfeld

$$\dot\gamma(t) = \frac{d\gamma}{dt}(t) = \begin{pmatrix} \frac{d\gamma^1}{dt}(t) \\ \vdots \\ \frac{d\gamma^n}{dt}(t) \end{pmatrix}, \quad 0 \le t \le 1.$$

Sei weiter $\lambda = \sum_{i=1}^n \lambda_i(x)dx^i$ mit

$$\lambda = (\lambda_1, \dots, \lambda_n) \in C^0(\Omega; \mathbb{R}^n)$$

eine 1-Form auf Ω. Dann wird durch

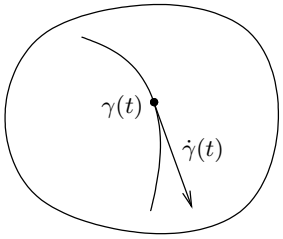
$$t \mapsto \lambda(\gamma(t))\dot\gamma(t) = \sum_{i=1}^n \lambda_i(\gamma(t)) \frac{d\gamma^i}{dt}(t)$$

eine stetige Funktion auf $[0,1]$ definiert.

Definition 7.4.1. *Der Ausdruck*

$$\int_\gamma \lambda := \int_0^1 \lambda(\gamma(t))\dot\gamma(t)\ dt$$

heisst **Wegintegral** *von λ längs γ.*



Bemerkung 7.4.1. i) Das Wegintegral $\int_\gamma \lambda$ ist unabhängig von orientierungserhaltenden Umparametrisierungen von γ.

ii) Wege $\gamma_1, \gamma_2 \in C^1([0,1];\Omega)$ mit $\gamma_1(1) = \gamma_2(0)$ kann man aneinanderhängen zu einem stückweise C^1-Weg $\gamma = \gamma_1 + \gamma_2 \colon [0,2] \to \Omega$ mit

$$t \mapsto \begin{cases} \gamma_1(t), & 0 \le t \le 1 \\ \gamma_2(t-1), & 1 \le t \le 2. \end{cases}$$

Offenbar kann man das Wegintegral einer 1-Form λ auch für derartige $\gamma = \gamma_1 + \gamma_2 \in C^1_{pw}([0,2];\Omega)$ erklären (mit Index "pw" für Engl. "piece-wise"), und es gilt

$$\int_{\gamma_1+\gamma_2} \lambda = \int_{\gamma_1} \lambda + \int_{\gamma_2} \lambda.$$

Beispiel 7.4.1. i) Sei $\gamma \in C^1([0,2\pi];\mathbb{R}^2)$ mit

$$\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad 0 \le t \le 2\pi,$$

eine Parametrisierung des Einheitskreises, $\lambda = \lambda(x,y)$ die 1-Form mit

$$\lambda(x,y) = -y\ dx + x\ dy, \quad (x,y) \in \mathbb{R}^2.$$

Dann gilt

$$\begin{aligned} \int_\gamma \lambda &= \int_0^{2\pi} (-\sin t, \cos t) \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}\ dt \\ &= \int_0^{2\pi} (\sin^2(t) + \cos^2(t))\ dt = 2\pi. \end{aligned}$$

ii) Sei $\Omega \in \mathbb{R}^n$ offen, $\gamma \in C^1([0,1];\Omega)$, $f \in C^1(\Omega)$. Betrachte $\lambda = df$. Gemäss Satz 7.2.3 gilt

$$df(\gamma(t))\dot\gamma(t) = \frac{d}{dt}(f \circ \gamma)(t);$$

d.h.

$$\int_\gamma df = \int_0^1 \frac{d}{dt}(f \circ \gamma)(t)\ dt = f(\gamma(1)) - f(\gamma(0))$$

hängt nur von Anfangs- und Endpunkt des Weges γ ab.

Beispiel 7.4.1.ii) liefert ein Analogon zu Korollar 5.2.1.i).

Satz 7.4.1. *Sei $\Omega \subset \mathbb{R}^n$ offen und (C^1-) wegzusammenhängend im Sinne von Definition 4.6.2, und sei $f \in C^1(\Omega)$ mit $df = 0$. Dann ist f konstant.*

Beweis. Zu je zwei Punkten $x_0, x_1 \in \Omega$ gibt es ein $\gamma \in C^1([0,1];\Omega)$ mit $x_0 = \gamma(0)$, $x_1 = \gamma(1)$. Mit Beispiel 7.4.1.ii) erhalten wir

$$f(x_1) - f(x_0) = f(\gamma(1)) - f(\gamma(0)) = \int_\gamma df = 0.$$

□

Wie kann man entscheiden, ob eine Differentialform λ von der Form $\lambda = df$ ist für ein $f \in C^1(\Omega)$?

Satz 7.4.2. *Sei $\lambda \in C^0(\Omega;\mathbb{R}^n)$. Es sind äquivalent:*

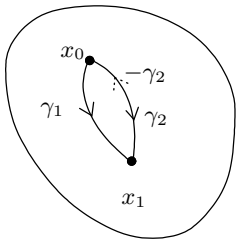
i) *$\exists f \in C^1(\Omega)$: $\lambda = df$ ("Potential").*

ii) *Für je zwei Wege $\gamma_{1,2} \in C^1_{pw}([0,1];\Omega)$ mit $\gamma_1(0) = \gamma_2(0)$, $\gamma_1(1) = \gamma_2(1)$ gilt*

$$\int_{\gamma_1} \lambda = \int_{\gamma_2} \lambda.$$

iii) *Für jeden "geschlossenen" Weg $\gamma \in C^1_{pw}([0,1];\Omega)$ mit $\gamma(0) = \gamma(1)$ gilt*

$$\int_\gamma \lambda = 0.$$



Beweis. *i)* $\Rightarrow$ *ii)*: Beispiel 7.4.1.ii).

ii) $\Rightarrow$ *i)*: Fixiere $p_0 \in \Omega$. Setze $f(p_0) = 0$. Für $x \in \Omega$ sei $\gamma \in C^1_{pw}([0,1];\Omega)$ ein Weg mit $\gamma(0) = p_0$, $\gamma(1) = x$. Nach Annahme ii) ist die durch

$$f(x) := \int_\gamma \lambda$$

definierte Funktion f auf Ω wohldefiniert.

Behauptung $f \in C^1(\Omega)$, $df = \lambda$.

Beweis. Sei $x_0 \in \Omega$, $\gamma_0 \in C^1_{pw}([0,1];\Omega)$ ein Weg von $p_0 = \gamma_0(0)$ nach $x_0 = \gamma_0(1)$. Sei $r > 0$ mit $B_r(x_0) \subset \Omega$. Für beliebiges $i \in \{1,\dots,n\}$, $0 < |h| < r$ gilt

$$\gamma(t) = x_0 + the_i \in C^1([0,1];\Omega)$$

und

$$f(x_0 + he_i) - f(x_0) = \int_{\gamma_0+\gamma} \lambda - \int_{\gamma_0} \lambda = \int_{\gamma} \lambda$$
$$= \int_0^1 \lambda(x_0 + the_i) he_i \; dt.$$

Da λ stetig ist, folgt, dass

$$\lim_{h\to 0} \frac{f(x_0 + he_i) - f(x_0)}{h} = \lim_{h\to 0} \int_0^1 \lambda(x_0 + the_i) e_i \; dt = \lambda(x_0) e_i$$

existiert; d.h. f ist auf Ω partiell in Richtung e_i differenzierbar mit

$$\frac{\partial f}{\partial x^i}(x_0) = \lambda(x_0) e_i \in C^0(\Omega), \quad 1 \le i \le n,$$

und $f \in C^1(\Omega)$ mit $df = \lambda$. □

ii) $\Rightarrow$ *iii)*: Sei $\gamma \in C^1_{pw}([0,1];\Omega)$ geschlossen mit $\gamma(0) = \gamma(1) = x_0$. Wähle $\gamma_1(t) = x_0$, $0 \le t \le 1$. Mit ii) folgt

$$\int_{\gamma} \lambda = \int_{\gamma_1} \lambda = 0.$$

iii) $\Rightarrow$ *ii)*: Seien $\gamma_{1,2} \in C^1_{pw}([0,1];\Omega)$ mit $\gamma_1(0) = \gamma_2(0)$, $\gamma_1(1) = \gamma_2(1)$. Definiere den Weg $-\gamma_2(t) := \gamma_2(1-t) \in C^1_{pw}([0,1];\Omega)$ mit

$$\int_{-\gamma_2} \lambda = -\int_0^1 \lambda(\gamma_2(1-t))\dot{\gamma}_2(1-t) \; dt = -\int_{\gamma_2} \lambda.$$

Der Weg $\gamma = \gamma_1 - \gamma_2 = \gamma_1 + (-\gamma_2) \in C^1_{pw}([0,1];\Omega)$ ist geschlossen, also

$$0 = \int_{\gamma} \lambda = \int_{\gamma_1} \lambda + \int_{-\gamma_2} \lambda = \int_{\gamma_1} \lambda - \int_{\gamma_2} \lambda.$$

□

Bemerkung 7.4.2. Der Beweis von Satz 7.4.2 liefert offenbar ein Verfahren zur Berechnung des Potentials f der 1-Form $\lambda = df$.

Beispiel 7.4.2. i) Sei

$$\lambda(x,y) = 2xy^2 dx + 2x^2 y dy, \quad (x,y) \in \mathbb{R}^2.$$

Wir setzen an $f(0,0) = 0$ und bestimmen zunächst einen **Ansatz** für $f(x,0)$ durch

$$f(x,0) = \int_{\gamma} \lambda, \quad \text{wo } \gamma(t) = (tx, 0), \quad 0 \le t \le 1.$$

Da $y = 0$ längs γ, folgt

$$f(x,0) = \int_0^1 \lambda(\gamma(t))\dot{\gamma}(t) \; dt = \int_0^1 (2tx \cdot 0, 2t^2x^2 \cdot 0) \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} \; dt = 0.$$

Anschliessend machen wir für beliebiges $(x, y) \in \mathbb{R}^2$ den **Ansatz**

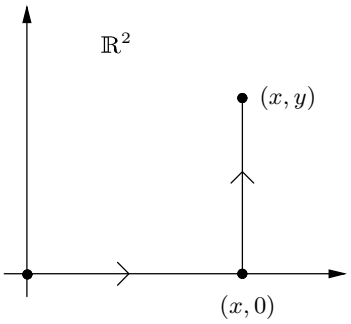
$$f(x, y) = f(x, 0) + \int_\gamma \lambda,$$

wobei wir nun als Weg $\gamma(t) = (x, ty), \quad 0 \leq t \leq 1$ wählen. Dies ergibt

$$\begin{aligned} f(x, y) &= 0 + \int_0^1 (2x(ty)^2, 2x^2(ty)) \begin{pmatrix} 0 \\ y \end{pmatrix} \, dt \\ &= \int_0^1 2x^2 ty^2 \, dt = x^2 y^2. \end{aligned}$$

Zum Schluss verifizieren wir

$$df(x, y) = (2xy^2, 2x^2 y) = \lambda, \quad (x, y) \in \mathbb{R}^2.$$



ii) Das analoge Vorgehen im Fall

$$\lambda(x, y) = 2xy^2 dx + 2y dy, \quad (x, y) \in \mathbb{R}^2$$

ergibt $f(x, 0) = 0$, $x \in \mathbb{R}$, und

$$f(x, y) = \int_0^1 2ty^2 \, dt = y^2, \quad \forall (x, y) \in \mathbb{R}^2.$$

Die Probe **versagt** jedoch, da

$$df(x, y) = (0, 2y) \neq \lambda, \text{ falls } x \neq 0 \neq y.$$

iii) Ebenso besitzt die 1-Form $\lambda(x, y) = -y \, dx + x \, dy$ aus Beispiel 7.4.1 kein Potential auf $\mathbb{R}^2$.

In Bemerkung 7.3.1 haben wir gesehen, dass wir Vektorfelder in 1-Formen verwandeln können mittels dem Skalarprodukt. Somit können wir auch das Wegintegral für Vektorfelder erklären.

Sei $\Omega \subset \mathbb{R}^n$ offen, $v = (v^i)_{1 \leq i \leq n} \in C^0(\Omega; \mathbb{R}^n)$ ein Vektorfeld mit zugehöriger 1-Form λ, wobei

$$\lambda(x)w = \langle v(x), w \rangle_{\mathbb{R}^n}, \ \forall x \in \Omega, \ w \in T_x \mathbb{R}^n,$$

und sei $\gamma \in C^1([0, 1]; \Omega)$.

Definition 7.4.2. *Das Wegintegral von v längs γ ist erklärt als*

$$\int_\gamma v \; d\vec{s} = \int_\gamma \lambda = \int_0^1 \langle v(\gamma(t)), \dot{\gamma}(t)\rangle_{\mathbb{R}^n} \; dt.$$

$d\vec{s}$ $(= \dot{\gamma}(t)\; dt)$ heisst **gerichtetes Längenelement** .

Definition 7.4.3. *Das Vektorfeld $v \in C^0(\Omega; \mathbb{R}^n)$ heisst* **konservativ**, *falls für jeden "geschlossenen" Weg $\gamma \in C^1_{pw}([0,1]; \Omega)$ mit $\gamma(0) = \gamma(1)$ gilt*

$$\int_\gamma v \cdot d\vec{s} = 0.$$

(Mit einem konservativen Kraftfeld v kann man also kein "perpetuum mobile" konstruieren.)

Aus Satz 7.4.2 folgt unmittelbar:

Satz 7.4.3. *Für $v \in C^0(\Omega; \mathbb{R}^n)$ sind äquivalent:*

i) *v ist konservativ;*

ii) *$\exists f \in C^1(\Omega)$: $v = \nabla f$.*

In diesem Fall heisst v **Potentialfeld** *mit dem* **Potential** *f.*

7.5 Höhere Ableitungen

Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega)$.

Definition 7.5.1. *Die Funktion f heisst* **von der Klasse** C^2, *$f \in C^2(\Omega)$, falls $\frac{\partial f}{\partial x^i} \in C^1(\Omega)$, $1 \le i \le n$.*

Satz 7.5.1. *Sei $f \in C^2(\Omega)$. Dann gilt*

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^i}\Big(\frac{\partial f}{\partial x^j}\Big) = \frac{\partial^2 f}{\partial x^j \partial x^i}, \quad 1 \le i, j \le n.$$

Bemerkung 7.5.1. Die Voraussetzung $f \in C^2(\Omega)$ ist wichtig, wie das Beispiel der Funktion

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

zeigt.

Beweis von Satz 7.5.1. Sei $x_0 \in \Omega$, $i \neq j$. Für $h, k > 0$ betrachte den Ausdruck

$$I := (f(x_0 + he_i + ke_j) - f(x_0 + he_i)) - (f(x_0 + ke_j) - f(x_0)) \; .$$

Indem wir schreiben

$$f(x_0 + ke_j) - f(x_0) = \int_\gamma df = k \int_0^1 \frac{\partial f}{\partial x^j}(x_0 + tke_j)\ dt\ ,$$

wobei

$$\gamma(t) = x_0 + tke_j,\ 0 \le t \le 1,$$

und ebenso für den ersten Term, erhalten wir zunächst

$$I = k \int_0^1 \Big(\frac{\partial f}{\partial x^j}(x_0 + he_i + tke_j) - \frac{\partial f}{\partial x^j}(x_0 + tke_j) \Big)\ dt\ .$$

Nach einer analogen Umformung des Integranden ergibt sich schliesslich

$$I = hk \int_0^1 \Big(\int_0^1 \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j)\ ds \Big)\ dt\ .$$

Vertauschen der Summationsfolge lässt den Ausdruck

$$I := (f(x_0 + he_i + ke_j) - f(x_0 + ke_j)) - (f(x_0 + he_i) - f(x_0))\ .$$

unverändert; jedoch werden dabei die Rollen von i und j vertauscht, und wir erhalten

$$I = hk \int_0^1 \Big(\int_0^1 \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0 + she_i + tke_j)\ dt \Big)\ ds\ .$$

Da $\frac{\partial^2 f}{\partial x^i \partial x^j}$ nach Voraussetzung stetig, ergibt Korollar 6.3.1 für den Term

$$R_{ij} := \int_0^1 \Big(\int_0^1 \Big(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j) - \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) \Big)\ ds \Big)\ dt$$

die Abschätzung

$$|R_{ij}| \le \sup_{|x - x_0| < h+k} \left| \frac{\partial^2 f}{\partial x^i \partial x^j}(x) - \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) \right| \to 0 \quad (h, k \to 0);$$

analog $R_{ji} \to 0\ \ (h, k \to 0)$. Subtraktion der obigen beiden Ausdrücke für I und Division durch hk ergibt somit

$$\begin{aligned} 0 &= \int_0^1 \Big(\int_0^1 \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j)\ ds \Big)\ dt \\ &\quad - \int_0^1 \Big(\int_0^1 \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0 + she_i + tke_j)\ dt \Big)\ ds \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) + R_{ij} - \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0) - R_{ji}\ . \end{aligned}$$

Nach Grenzübergang $h, k \to 0$ folgt

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) = \frac{\partial^2 f}{\partial x^j \partial x^i},$$

wie gewünscht. □

Mit Satz 7.5.1 erhalten wir eine einfach zu handhabende notwendige Bedingung für ein konservatives Vektorfeld.

Korollar 7.5.1. *Sei $v = (v^i)_{1\le i\le n} \in C^1(\Omega;\mathbb{R}^n)$ konservativ. Dann gilt*

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial v^j}{\partial x^i}, \quad 1 \le i,j \le n.$$

Beweis. Nach Voraussetzung ist $v = \nabla f$ für ein $f \in C^2(\Omega)$. Mit Satz 7.5.1 folgt

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i} = \frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial v^j}{\partial x^i}, \ 1 \le i,j \le n.$$

□

Beispiel 7.5.1. Sei $v(x,y) = \begin{pmatrix} 2xy^2 \\ 2y \end{pmatrix}$, $(x,y) \in \mathbb{R}^2$. Es gilt

$$\frac{\partial v^1}{\partial y}(x,y) = 4xy, \quad \frac{\partial v^2}{\partial x}(x,y) = 0;$$

also ist v nicht konservativ; vgl. Beispiel 7.4.2.ii).

Für beliebiges $m \in \mathbb{N}$ definieren wir induktiv analog zu Definition 7.5.1:

Definition 7.5.2. *Die Funktion $f \in C^1(\Omega)$ heisst* **von der Klasse C^m**, *$f \in C^m(\Omega)$, falls $\frac{\partial f}{\partial x^i} \in C^{m-1}(\Omega)$, $1 \le i \le n$.*

Bemerkung 7.5.2. Gemäss Satz 7.5.1 sind für $f \in C^m(\Omega)$ partielle Ableitungen der Ordnung $\le m$ beliebig vertauschbar.

Sei $f \in C^m(\Omega)$, und seien $x_0, x_1 \in \Omega$ mit

$$x_t = (1-t)x_0 + tx_1 \in \Omega, \quad 0 \le t \le 1. \tag{7.5.1}$$

Gemäss Satz 7.2.3 ist die Funktion

$$\phi(t) = f(x_t) \in C^m([0,1])$$

mit

$$\begin{aligned}
\frac{d\phi}{dt}(t) &= df(x_t)(x_1 - x_0) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(x_t)(x_1^i - x_0^i) \\
\frac{d^2\phi}{dt}(t) &= \sum_{i_1,i_2=1}^n \frac{\partial^2 f}{\partial x^{i_1}\partial x^{i_2}}(x_t)(x_1^{i_1} - x_0^{i_1})(x_1^{i_2} - x_0^{i_2}) \\
&\vdots \quad \vdots \\
\frac{d^m\phi}{dt}(t) &= \sum_{i_1,\dots,i_m=1}^n \frac{\partial^m f}{\partial x^{i_1}\dots\partial x^{i_m}}(x_t) \prod_{j=1}^m (x_1^{i_j} - x_0^{i_j}).
\end{aligned}$$

Satz 5.5.1, angewandt auf ϕ, ergibt

Satz 7.5.2. (Taylorformel) *Sei $f \in C^m(\Omega)$, und seien $x_0, x_1 \in \Omega$ mit (7.5.1). Dann gibt es eine Zahl $0 < \vartheta < 1$ mit*

$$f(x_1) = f(x_0) + df(x_0)(x_1 - x_0) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x_1^i - x_0^i)(x_1^j - x_0^j)$$
$$+ \cdots + \frac{1}{m!}\sum_{i_1,\ldots,i_m=1}^{n} \frac{\partial^m f}{\partial x^{i_1} \ldots \partial x^{i_m}}(x_\vartheta) \prod_{j=1}^{m}(x_1^{i_j} - x_0^{i_j}).$$

Bemerkung 7.5.3. Mit der Multi-Index Schreibweise

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{(\partial x^1)^{\alpha_1} \ldots (\partial x^n)^{\alpha_n}}$$

für $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$ analog zu Beispiel 7.1.5.i) kann man den Ausdruck

$$\sum_{i_1,\ldots,i_m=1}^{n} \frac{\partial^m f}{\partial x^{i_1} \ldots \partial x^{i_m}}(x_0) \prod_{j=1}^{m}(x_1^{i_j} - x_0^{i_j}) = \sum_{|\alpha|=m} \partial^\alpha f(x_0)(x_1 - x_0)^\alpha$$

kompakt schreiben.

Wie in Abschnitt 5.5 definieren wir das **Taylor-Polynom m-ter Ordnung**

$$\begin{aligned} T_n f(x, x_0) &= f(x_0) + df(x_0)(x - x_0) + \ldots \\ &\quad + \frac{1}{m!}\sum_{i_1,\ldots,i_m=1}^{n} \frac{\partial^m f}{\partial x^{i_1} \ldots \partial x^{i_m}}(x_0) \prod_{j=1}^{m}(x_1^{i_j} - x_0^{i_j}) \\ &= f(x_0) + df(x_0)(x - x_0) + \cdots + \frac{1}{m!}\sum_{|\alpha|=m} \partial^\alpha f(x_0)(x_1 - x_0)^\alpha. \end{aligned}$$

Bemerkung 7.5.4. i) Gemäss Satz 7.5.2 gilt für $f \in C^m(\Omega)$

$$f(x_1) = T_m f(x; x_0) + r_m f(x, x_0)$$

mit

$$|r_m f(x_1; x_0)| \leq \frac{n^m}{m!} \sup_{0<\vartheta<1,\ |\alpha|=m} |\partial^\alpha f(x_\vartheta) - \partial^\alpha f(x_0)|\, |x_1 - x_0|^m .$$

ii) Falls $f \in C^{m+1}(\Omega)$, so liefert Satz 7.5.2 alternativ die Abschätzung

$$|r_m f(x_1; x_0)| \leq \frac{n^{m+1}}{(m+1)!} \sup_{0<\vartheta<1,\ |\alpha|=m+1} |\partial^\alpha f(x_\vartheta)|\, |x_1 - x_0|^{m+1} .$$

iii) Insbesondere für $m = 2$ erhalten wir für f die **quadratische** Näherung

$$\begin{aligned} f(x_1) &= f(x_0) + df(x_0)(x_1 - x_0) + \\ &\quad \frac{1}{2}\sum_{i,j=1}^{2} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x_1^i - x_0^i)(x_1^j - x_0^j) + r_2(f(x_1, x_0)) \end{aligned}$$

mit Fehler

$$\frac{r_2 f(x_1, x_0)}{|x_1 - x_0|^2} \to 0 \quad (x_1 \to x_0).$$

Analog zu Korollar 5.5.1 für $n = 1$ gilt nun:

Satz 7.5.3. *Sei $f \in C^2(\Omega)$, $x_0 \in \Omega$.*

i) *Falls $x_0 \in \Omega$ lokale Minimalstelle von f ist, so gilt $df(x_0) = 0$.*

ii) *Falls $df(x_0) = 0$ und falls die symmetrische quadratische "***Hesse-Matrix***"*

$$Hess_f(x_0) = \Big(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)\Big)_{1 \le i,j \le n}$$

positiv definit ist im Sinne von

$$Hess_f(x_0)(\xi, \xi) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)\xi^i \xi^j > 0$$

für alle $\xi \in \mathbb{R}^n \backslash \{0\}$, so ist x_0 eine strikte lokale Minimalstelle von f.

Beweis. **i)** Sei x_0 eine lokale Minimalstelle von f, und nimm widerspruchsweise an, $df(x_0) \neq 0$. Setze $e = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$. Dann hat die Funktion

$$\phi(t) = f(x_0 + te), \quad |t| << 1,$$

bei $t = 0$ ein lokales Minimum; jedoch gilt

$$\frac{d\phi}{dt}(0) = df(x_0)e = |\nabla f(x_0)| > 0$$

im Widerspruch zu Korollar 5.5.1.i).

ii) Da $S^{n-1} = \partial B_1(0; \mathbb{R}^n)$ kompakt, gibt es $\lambda > 0$ mit

$$Hess_f(x_0)(\xi, \xi) \ge \lambda |\xi|^2, \quad \forall \xi \in S^{n-1}.$$

Mit Bemerkung 7.5.3.iii) folgt für $x \neq x_0$ genügend nahe bei x_0 die Ungleichung

$$f(x) - f(x_0) \ge \lambda |x - x_0|^2 + r_2 f(x; x_0) \ge \frac{\lambda}{2} |x - x_0|^2 > 0 \ ,$$

wie gewünscht.

□

Bemerkung 7.5.5. Analog zu Satz 7.5.3.ii) ist ein kritischer Punkt x_0 von f, wo die Hesse-Matrix negativ definit ist mit

$$Hess_f(x_0)(\xi, \xi) < 0, \quad \forall \xi \in \mathbb{R}^n \backslash \{0\},$$

eine strikte lokale Maximalstelle von f.

Definition 7.5.3. i) *Ein Punkt* $x_0 \in \Omega$ *mit* $df(x_0) = 0$ *heisst* **kritischer Punkt** *von* f.

ii) *Die quadratische Form*

$$\xi \mapsto Hess_f(x_0)(\xi, \xi)$$

heisst **Hessesche Form** *von* f.

Beispiel 7.5.2. Sei $f(x,y) = \frac{1}{2}(x^2 + \alpha y^2) + \beta xy$, $(x,y) \in \mathbb{R}^2$, mit Parametern $\alpha, \beta \in \mathbb{R}$. Es gilt

$$df(x,y) = (x + \beta y, \alpha y + \beta x)$$

sowie

$$Hess_f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{(\partial x)^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{(\partial y)^2} \end{pmatrix}(x,y) = \begin{pmatrix} 1 & \beta \\ \beta & \alpha \end{pmatrix} =: H.$$

Offenbar ist $(x_0, y_0) = (0,0)$ kritischer Punkt von f. Die Eigenwerte von H entscheiden, um welchen Typ es sich handelt.

Das charakteristische Polynom

$$\begin{aligned} p(\lambda) &= det(H - \lambda \cdot id) = (1-\lambda)(\alpha - \lambda) - \beta^2 \\ &= \lambda^2 - (1+\alpha)\lambda + \alpha - \beta^2 \end{aligned}$$

der Matrix H hat die Nullstellen

$$\lambda_{1,2} = \frac{1+\alpha}{2} \pm \sqrt{\frac{(1+\alpha)^2}{4} - \alpha + \beta^2}.$$

(Beachte: $(1+\alpha)^2 - 4\alpha = (1-\alpha)^2 \geq 0$.) Um das Verhalten von f in der Nähe von $(x_0, y_0) = (0,0)$ zu verstehen, unterscheiden wir die folgenden Fälle:

i) $\alpha > \beta^2$. In diesem Fall gilt $\lambda_{1,2} > 0$; also ist H positiv definit, und der Punkt $(x_0, y_0) = (0,0)$ ist eine strikte lokale (sogar die globale) Minimalstelle.

ii) $\alpha = \beta^2$. Es gilt $\lambda_1 > 0 = \lambda_2$, und H ist positiv semi-definit. Da f quadratisch ist, folgt

$$f(x,y) \geq 0, \ \forall (x,y) \in \mathbb{R}^2 \ .$$

Andererseits gilt offenbar

$$f(-\beta y, y) = 0, \ \forall y \in \mathbb{R} \ ;$$

der Punkt $(x_0, y_0) = (0,0)$ ist also ein (entartetes) lokales Minimum.

iii) $\alpha < \beta^2$. Dann gilt $\lambda_1 > 0 > \lambda_2$; die Matrix H ist **indefinit**. Der Punkt $(x_0, y_0) = (0,0)$ ist also weder ein lokales Minimum noch ein lokales Maximum von f sondern ein **Sattelpunkt**: Jede Umgebung U von (0,0) enthält Punkte $p, q \in U$ mit

$$f(p) > 0 > f(q).$$

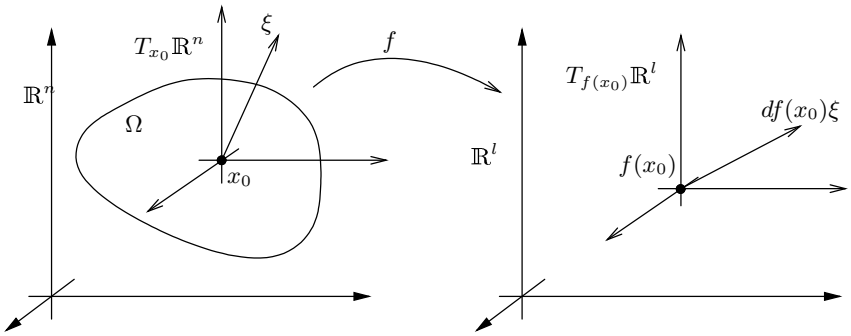
7.6 Vektorwertige Funktionen

Sei $\Omega \subset \mathbb{R}^n$ offen, $f = (f^i)_{1\le i\le l}\colon \Omega \to \mathbb{R}^l$.

Definition 7.6.1. i) *Die Funktion f heisst an der Stelle $x_0 \in \Omega$ differenzierbar, falls jede Komponente f^i, $1 \le i \le l$, an der Stelle x_0 differenzierbar ist. Das Differential $df(x_0)$ hat die Gestalt*

$$df(x_0) = \begin{pmatrix} df^1(x_0) \\ \vdots \\ df^l(x_0) \end{pmatrix} : \mathbb{R}^n \cong T_{x_0}\mathbb{R}^n \to T_{f(x_0)}\mathbb{R}^l \cong \mathbb{R}^l.$$

ii) *f heisst auf Ω differenzierbar, bzw. von der Klasse C^m, $m \ge 1$, falls jedes f^i differenzierbar, bzw. $f^i \in C^m(\Omega)$, $1 \le i \le l$.*



Notation: $C^m(\Omega;\mathbb{R}^l) = \{f = (f^i)_{1\le i\le l};\ f^i \in C^m(\Omega),\ 1 \le i \le l\}$.

Bemerkung 7.6.1. i) Bzgl. der Standardbasis dx^j, $1 \le j \le n$, erhalten wir mit

$$df^i(x_0) = \sum_{j=1}^n \frac{\partial f}{\partial x^j}(x_0)dx^j = \Big(\frac{\partial f^i}{\partial x^1}(x_0), \dots, \frac{\partial f^i}{\partial x^n}(x_0)\Big)$$

die Darstellung

$$df(x_0) = \sum_{j=1}^n \frac{\partial f}{\partial x^j}(x_0)dx^j = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(x_0) & \dots & \frac{\partial f^1}{\partial x^n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f^l}{\partial x^1}(x_0) & \dots & \frac{\partial f^l}{\partial x^n}(x_0) \end{pmatrix}.$$

Die Matrix

$$df(x_0) = \Big(\frac{\partial f^i}{\partial x_j}(x_0)\Big)_{1\le i\le l,\ 1\le j\le n}$$

heisst **Jacobi-** oder **Funktionalmatrix** von f an der Stelle x_0.

ii) Auch im vektorwertigen Fall ist die Funktion f genau dann differenzierbar an der Stelle x_0, wenn eine lineare Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^l$ existiert mit

$$\lim_{x\to x_0,\ 0\ne x\in\Omega} \frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0,$$

analog zu Definition 7.1.2.

Beispiel 7.6.1. i) Sei $f = f_1 \colon \mathbb{R}^2 \to \mathbb{R}^2$ gegeben mit

$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}, \quad (x,y) \in \mathbb{R}^2.$$

Offenbar gilt $f \in C^\infty(\mathbb{R}^2; \mathbb{R}^2)$ mit

$$df(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}, \ (x,y) \in \mathbb{R}^2.$$

ii) Sei $f = f_2 \colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben mit

$$f(x,y,z) = \begin{pmatrix} x^2 + y^2 + z^2 \\ xyz \end{pmatrix}, \quad (x,y,z) \in \mathbb{R}^3.$$

Dann ist $f \in C^\infty(\mathbb{R}^3; \mathbb{R}^2)$ mit

$$df(x,y,z) = \begin{pmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{pmatrix}, \quad (x,y,z) \in \mathbb{R}^3.$$

Es gelten die üblichen Differentiationsregeln:

Satz 7.6.1. *Seien $f, g \colon \Omega \to \mathbb{R}^l$ an der Stelle $x_0 \in \Omega$ differenzierbar. Dann sind die Summe sowie das Skalarprodukt von f und g an der Stelle x_0 differenzierbar, und*

i) $d(f+g)(x_0) = df(x_0) + dg(x_0)$

ii) $d(f \cdot g)(x_0) = f(x_0) \cdot dg(x_0) + g(x_0) \cdot df(x_0)$,

wobei $f(x_0) \cdot dg(x_0) = \sum_{i=1}^{l} (f^i(x_0) dg^i(x_0)$, *etc.*

Beweis. Der Beweis von Satz 7.2.1 lässt sich unmittelbar übertragen. □

Beispiel 7.6.2. Seien $f(x) = g(x) = x$ mit $df(x) = id$, $\forall x$. Satz 7.6.1 ergibt

$$d(|x|^2)\xi = 2x \cdot \xi, \quad \forall x \in \mathbb{R}^n, \ \xi \in \mathbb{R}^n.$$

Satz 7.6.2. (Kettenregel, 3. Teil) *Seien $g \colon \Omega \to \mathbb{R}^l$ an der Stelle $x_0 \in \Omega$ und $f \colon \mathbb{R}^l \to \mathbb{R}^m$ an der Stelle $y_0 = g(x_0) \in \mathbb{R}^l$ differenzierbar. Dann ist die Funktion $f \circ g \colon \Omega \to \mathbb{R}^m$ an der Stelle x_0 differenzierbar, und*

$$d(f \circ g)(x_0) = df(g(x_0)) \cdot dg(x_0).$$

Proof. Der Beweis ist derselbe wie von Satz 7.2.2. □

Bemerkung 7.6.2. i) Falls $g \colon \mathbb{R}^n \to \mathbb{R}^l$, $f \colon \mathbb{R}^l \to \mathbb{R}^m$ linear mit

$$f(y) = Ay, \ y \in \mathbb{R}^l, \quad g(x) = Bx, \ x \in \mathbb{R}^n,$$

so ist $f \circ g \colon \mathbb{R}^n \to \mathbb{R}^m$ linear mit

$$(f \circ g)(x) = ABx,\ x \in \mathbb{R}^n.$$

ii) Im folgenden Abschnitt sind die Rollen von f und g meist vertauscht. D.h., wir betrachten differenzierbare Funktionen $f \colon \Omega \to \mathbb{R}^l$, $g \colon \mathbb{R}^l \to \mathbb{R}^m$ und wenden Satz 7.6.2 an auf die Funktion $g \circ f$ mit

$$d(g \circ f)(x_0) = dg(f(x_0)) \cdot df(x_0). \tag{7.6.1}$$

In Koordinaten und mit $y_0 = f(x_0)$ können wir die Formel (7.6.1) schreiben

$$\frac{\partial (g \circ f)^i}{\partial x^k}(x_0) = \sum_{j=1}^{l} \frac{\partial g^i(y_0)}{\partial y^j} \cdot \frac{\partial f^j(x_0)}{\partial x^k},\ 1 \le i \le m,\ 1 \le k \le n.$$

Beispiel 7.6.3. Betrachte die Funktionen $g = f_1 \colon \mathbb{R}^2 \to \mathbb{R}^2$, $f = f_2 \colon \mathbb{R}^3 \to \mathbb{R}^2$ aus Beispiel 7.6.1. Die Funktion

$$(g \circ f)(x,y,z) = \begin{pmatrix} (x^2+y^2+z^2)^2 - x^2y^2z^2 \\ 2(x^2+y^2+z^2)xyz \end{pmatrix} : \mathbb{R}^3 \to \mathbb{R}^2$$

ist differenzierbar mit

$$\begin{aligned} d(g \circ f)(x,y,z) &= dg(f(x,y,z)) \cdot df(x,y,z) \\ &= \begin{pmatrix} 2(x^2+y^2+z^2) & -2xyz \\ 2xyz & 2(x^2+y^2+z^2) \end{pmatrix} \begin{pmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{pmatrix} \\ &= \begin{pmatrix} 4x(x^2+y^2+z^2) - 2xy^2z^2 & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}. \end{aligned}$$

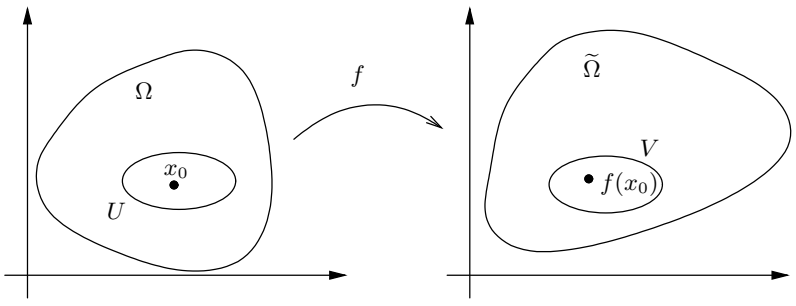
Probe: Differenziere direkt.

7.7 Der Umkehrsatz

Sei $\Omega \subset \mathbb{R}^n$, $f \in C^1(\Omega, \mathbb{R}^n)$ injektiv, $\widetilde{\Omega} = f(\Omega)$ der Wertebereich von f.

Fragen i) Unter welchen Bedingungen ist die Umkehrabbildung $f^{-1} \colon \widetilde{\Omega} \to \Omega$ wieder von der Klasse C^1?

ii) Gibt es Bedingungen an $df(x_0)$, die gewährleisten, dass f in einer Umgebung von x_0 injektiv ist?



Offenbar liefert Satz 7.6.2 eine notwendige Bedingung: Falls $f \in C^1(\Omega;\mathbb{R}^n)$, $g \in C^1(\widetilde{\Omega};\mathbb{R}^n)$ zu f invers, so folgt für jedes $x_0 \in \Omega$:

$$id = d(g \circ f)(x_0) = dg(f(x_0))df(x_0);$$

die lineare Abbildung $df(x_0)$ muss also invertierbar sein.

Gemäss dem folgenden Satz ist diese Bedingung auch hinreichend für die lokale Invertierbarkeit von f.

Satz 7.7.1. (Umkehrsatz) *Sei $f \in C^1(\Omega;\mathbb{R}^n)$ und sei $df(x_0)\colon \mathbb{R}^n \to \mathbb{R}^n$ invertierbar an einer Stelle $x_0 \in \Omega$. Dann existieren Umgebungen U von x_0, V von $f(x_0) = y_0$ und eine Funktion $g \in C^1(V;\mathbb{R}^n)$ mit $g = (f\big|_U)^{-1}$; d.h.*

$$g(f(x)) = x,\ \forall x \in U,\ \ f(g(y)) = y,\ \forall y \in V.$$

Weiter gilt für alle $x \in U$ gemäss (7.6.1) *die Beziehung*

$$dg(f(x)) = \big(df(x)\big)^{-1}\ .$$

Bevor wir diesen Satz beweisen, diskutieren wir die Aussage durch Vergleich mit dem Fall $n = 1$ und anhand von Beispielen.

Beispiel 7.7.1. i) Falls $n = 1$, $f \in C^1(]a,b[)$ mit $f'(x_0) > 0$ für ein $x_0 \in]a,b[$, so folgt aus der Stetigkeit von f' die Bedingung $f'(x) > 0$, $\forall x \in]x_0 - r, x_0 + r[$ für ein $r > 0$. Nach Satz 5.2.2 ist $f\colon]x_0 - r, x_0 + r[\to]c,d[$ invertierbar mit $g = f^{-1} \in C^1(]c,d[)$, und

$$\frac{d\big(f^{-1}\big)}{dy}\Big|_{y=f(x)} = \Big(\frac{df}{dx}(x)\Big)^{-1}.$$

D.h. das Differential $d\big(f^{-1}\big)(y)$ wird bzgl. der Standardbasis dy an der Stelle $y = f(x)$ dargestellt durch $\frac{1}{f'(x)}$.

ii) Betrachte die Funktion $f \in C^\infty(\mathbb{R}^2;\mathbb{R}^2)$ aus Beispiel 7.6.1.i) mit

$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}, \quad df(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Da für $(x,y) \neq (0,0)$ stets gilt

$$det\big(df(x,y)\big) = 4(x^2 + y^2) > 0,$$

ist f nach Satz 7.7.1 lokal um jeden Punkt $(x_0,y_0) \in \mathbb{R}^2\backslash\{(0,0)\}$ invertierbar.

Ist f auch "global" invertierbar? - Deute dazu $(x,y) = z = x + iy \in \mathbb{C}$,

$$f(x+iy) = x^2 - y^2 + 2ixy = (x+iy)^2 = z^2.$$

Wegen $f(-z) = f(z)$, $z \in \mathbb{C}$ ist f nicht "global" invertierbar auf $\mathbb{R}^2\backslash\{(0,0)\}$. Satz 7.7.1 zeigt jedoch, dass dies lokal auf $\mathbb{R}^2\backslash\{(0,0)\}$ möglich ist. Entsprechend kann man **lokal** auf $\mathbb{C}\backslash\{0\}$ eine Quadratwurzelfunktion definieren.

iii) Polarkoordinaten in $\mathbb{R}^2$. Die Abbildung $f\colon]0,\infty[\times\mathbb{R} \to \mathbb{R}^2$ mit

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

erfüllt

$$df(r,\varphi) = \begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix},$$

also

$$det(df(r,\varphi)) = r(\cos^2\varphi + \sin^2\varphi) = r > 0.$$

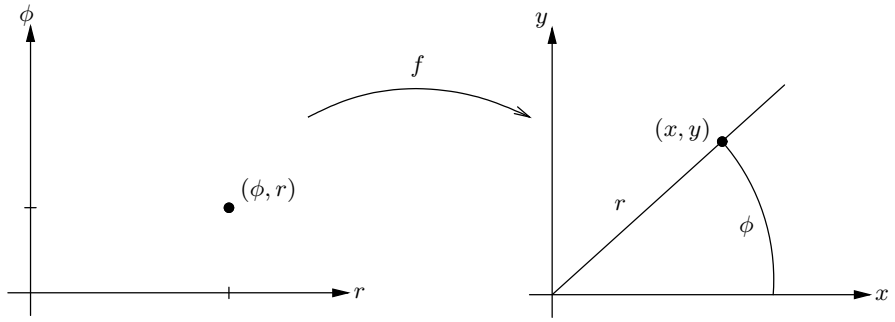
Gemäss Satz 7.7.1 kann man mittels f lokal Polarkoordinaten auf $\mathbb{R}^2$ einführen.

f ist injektiv zum Beispiel auf $]0,\infty[\times]-\pi,\pi[=:U$ mit

$$\left(f|_U\right)^{-1} = g\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} r = \sqrt{x^2+y^2} \\ \varphi = \arctan(y/x) \end{pmatrix} \quad \text{(lokal)}.$$

Die **Koordinatenlinien** $\alpha(r) = f(r,\varphi)$ (φ fest) und $\beta(\varphi) = f(r,\varphi)$ (r fest) schneiden sich senkrecht, da in jedem Punkt (r,φ) gilt

$$\frac{d\alpha}{dr}\cdot\frac{d\beta}{d\varphi} = \frac{\partial f}{\partial r}\cdot\frac{\partial f}{\partial \varphi} = 0.$$



Beweis von Satz 7.7.1. OBdA seien $x_0 = 0$, $y_0 = f(x_0) = 0$. (Betrachte sonst die Funktion $\tilde{f}(x) = f(x+x_0) - f(x_0)$.) Wähle $r_0 > 0$ so, dass $df(x)$ invertierbar für alle $x \in B_{r_0}(0)$. Wir zeigen zunächst:

Behauptung 1 Für genügend kleines $0 < r < r_0$ und $\delta = \delta(r) > 0$ gilt:

$$\forall y \in \overline{B_\delta(0)}\ \exists! x \in \overline{B_r(0)}:\ y = f(x).$$

Zum Beweis dieser Behauptung versuchen wir, für $y \in \mathbb{R}^n$ das gesuchte Urbild als Fixpunkt der Abbildung

$$\Phi_y\colon x \mapsto x + df(0)^{-1}(y - f(x))$$

zu erhalten. Gemäss Satz 6.5.2 genügt es dafür, die folgenden Aussagen zu beweisen.

Behauptung 2 $\exists 0 < r < r_0\ \forall y \in \mathbb{R}^n\ \forall x, \tilde{x} \in \overline{B_r(0)}$:

$$|\Phi_y(x) - \Phi_y(\tilde{x})| \leq \frac{1}{2}\,|x - \tilde{x}|\,.$$

Fixieren wir $0 < r < r_0$ gemäss Behauptung 2, so zeigen wir weiter:

Behauptung 3 $\exists \delta = \delta(r) > 0\ \forall y \in \overline{B_\delta(0)}$:

$$\Phi_y \colon \overline{B_r(0)} \to \overline{B_r(0)}\ .$$

Beweis von Behauptung 1. Gemäss den Behauptungen 2 und 3 ist für jedes $y \in \overline{B_\delta(0)}$ die Abbildung $\Phi_y \colon \overline{B_r(0)} \to \overline{B_r(0)}$ kontrahierend. Mit dem Kontraktionsprinzip, Satz 6.5.2, folgt Behauptung 1 nun unmittelbar. □

Beweis von Behauptung 3. Zu gegebenem $r > 0$ wähle $0 < \delta < \frac{r}{2\|df(0)^{-1}\|}$. Für $|y| \leq \delta$ schätze ab

$$|\Phi_y(0)| = \big|df(0)^{-1}y\big| \leq \big\|df(0)^{-1}\big\| \cdot \delta < \frac{r}{2}.$$

Mit Behauptung 2 folgt für $x \in \overline{B_r(0)}$:

$$|\Phi_y(x)| \leq |\Phi_y(x) - \Phi_y(0)| + |\Phi_y(0)| \leq \frac{1}{2}\,|x| + |\Phi_y(0)| < \frac{r}{2} + \frac{r}{2} = r.$$

□

Beweis von Behauptung 2. Schreibe für $x, \tilde{x} \in \overline{B_r(0)}$

$$\begin{aligned}\Phi_y(x) - \Phi_y(\tilde{x}) &= (x - \tilde{x}) + df(0)^{-1}\big(f(\tilde{x}) - f(x)\big)\\ &= df(0)^{-1}\big(f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)\big).\end{aligned}$$

Mit der Darstellung

$$\begin{aligned}f(\tilde{x}) - f(x) &= \int_0^1 \frac{d}{dt} f(x + t(\tilde{x} - x))\ dt\\ &= \int_0^1 df(x + t(\tilde{x} - x))(\tilde{x} - x)\ dt\end{aligned}$$

erhalten wir

$$\frac{f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)}{|\tilde{x} - x|} = \int_0^1 \big(df(x + t(\tilde{x} - x)) - df(0)\big)\frac{\tilde{x} - x}{|\tilde{x} - x|}\ dt$$

und können daher abschätzen

$$\begin{aligned}\frac{|f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)|}{|\tilde{x} - x|} &\leq \int_0^1 |df(x + t(\tilde{x} - x)) - df(0)|\ dt\\ &\leq \sup_t \|df(x + t(\tilde{x} - x)) - df(0)\|\\ &\leq \sup_{z \in \overline{B_r(0)}} \|df(z) - df(0)\| \to 0\ \ (r \to 0).\end{aligned}$$

Die Behauptung folgt. □

Bemerkung: Das in Satz 6.5.2 eingeführte Iterationsverfahren liefert im obigen Kontext ein sehr effizientes Verfahren zur numerischen Bestimmung des Urbildes von y unter f, ausgehend von einer "Startnäherung" $x_0 \in B_r(0)$, mittels der Vorschrift

$$x_{n+1} = \Phi_y(x_n), \quad n \in \mathbb{N}_0 \ .$$

Insbesondere im Fall $n = 1$, $y = 0$ wird dieses **Newton-Verfahren** gern herangezogen zur näherungsweisen Berechnung der Nullstellen einer Funktion $f \in C^1(\mathbb{R})$.

Setze nun $V := B_\delta(0) \subset \mathbb{R}^n$ und definiere

$$U := f^{-1}(V) \cap B_r(0).$$

Dann sind U und V offen, und gemäss Behauptung 1 ist $f\big|_U : U \to V$ bijektiv. Die Funktion

$$g := \big(f\big|_U\big)^{-1} : V \to U$$

ist also wohldefiniert. Beachte weiter, dass $df(x)$ nach Wahl von $r < r_0$ invertierbar ist an jeder Stelle $x \in U$.

Behauptung 4 $g \in C^1(V; \mathbb{R}^n)$, und es gilt

$$dg(y) = \big(df(g(y))\big)^{-1}, \quad \forall y \in V.$$

Beweis. **i)** Für $y = f(x)$, $\tilde{y} = f(\tilde{x}) \in V$ mit $x = g(y)$, $\tilde{x} = g(\tilde{y}) \in U \subset B_r(0)$ erhalten wir mit Behauptung 2 bei Wahl von $y = 0$ die Abschätzung

$$\begin{aligned} |\tilde{x} - x| &= \big|\Phi_0(\tilde{x}) - \Phi_0(x) + df(0)^{-1}(f(\tilde{x}) - f(x))\big| \\ &\le |\Phi_0(\tilde{x}) - \Phi_0(x)| + \big\|df(0)^{-1}\big\|\, |\tilde{y} - y| \\ &\le \frac{1}{2}\,|\tilde{x} - x| + C\,|\tilde{y} - y|\,; \end{aligned}$$

also

$$|\tilde{x} - x| \le 2C\,|\tilde{y} - y|\,.$$

ii) Fixiere $y = f(x) \in V$ mit $x \in U$. Falls $V \ni \tilde{y} = f(\tilde{x}) \to y$, $\tilde{y} \neq y$, so folgt mit i) auch $\tilde{x} = g(\tilde{y}) \to x = g(y)$, $\tilde{x} \neq x$. Schreibe

$$\begin{aligned} R_g := g(\tilde{y}) - g(y) - \big(df(x)\big)^{-1}(\tilde{y} - y) &= \tilde{x} - x - \big(df(x)\big)^{-1}(f(\tilde{x}) - f(x)) \\ &= -\big(df(x)\big)^{-1}\big(f(\tilde{x}) - f(x) - df(x)(\tilde{x} - x)\big)\ . \end{aligned}$$

Mit i) erhalten wir

$$\frac{|R_g|}{|\tilde{y} - y|} \le 2C\,\big\|df(x)^{-1}\big\|\,\frac{|f(\tilde{x}) - f(x) - df(x)(\tilde{x} - x)|}{|\tilde{x} - x|} \to 0 \quad (\tilde{y} \to y).$$

D.h. g ist an der Stelle $y \in V$ differenzierbar mit

$$dg(y) = \big(df(x)\big)^{-1} = \big(df(g(y))\big)^{-1},$$

und $g \in C^1(V; \mathbb{R}^n)$.

□

Mit Behauptung 4 ist nun auch Satz 7.7.1 vollständig bewiesen. □

7.8 Implizite Funktionen

Wir beginnen mit einfachen Beispielen.

Beispiel 7.8.1. i) Der Einheitskreis

$$S^1 = \{(x, y) \in \mathbb{R}^2;\ x^2 + y^2 = 1\}$$

mit der **impliziten** Darstellung $S^1 = f^{-1}(\{0\})$, wo

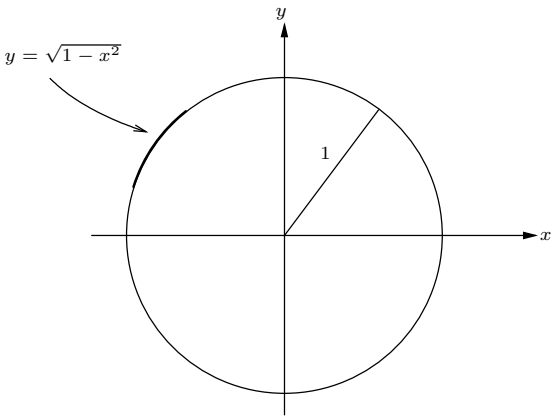
$$f(x, y) = x^2 + y^2 - 1$$

lässt sich lokal darstellen als Graph der Funktion

$$y = h(x) = \pm\sqrt{1 - x^2}, \quad -1 < x < 1$$

bzw. als Graph der Funktion

$$x = l(y) = \pm\sqrt{1 - y^2}, \quad -1 < y < 1.$$



ii) Sei K_b der Doppelkegel

$$K_b = \{(x, y, z) \in \mathbb{R}^3;\ x^2 + y^2 = b^2 z^2\}$$

mit Öffnungsverhältnis $b > 0$. Indem wir K_b schneiden mit der Ebene

$$E_\mu = \{(x, y, z) \in \mathbb{R}^3;\ z = 1 + \mu x\},$$

erhalten wir die Schnittkurven

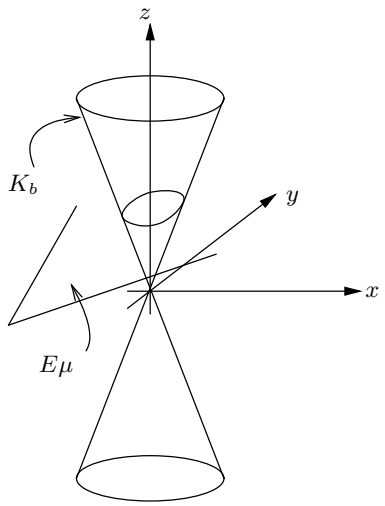
$$\begin{aligned} \Gamma: \quad x^2 + y^2 &= b^2(1 + \mu x)^2 = b^2 + 2\mu b^2 x + b^2\mu^2 x^2, \\ z &= 1 + \mu x. \end{aligned}$$

Für $b^2\mu^2 < 1$ handelt es sich dabei um eine Ellipse, für $b^2\mu^2 = 1$ um eine Parabel, und für $b^2\mu^2 > 1$ um eine Hyperbel.

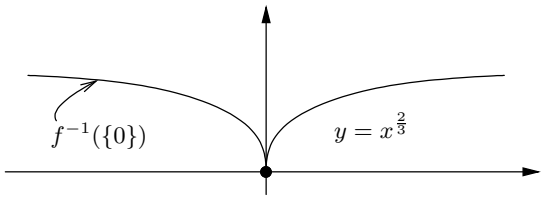
Implizit lassen sich alle diese Kegelschnitte wiederum bequem darstellen in der Form $\Gamma = f^{-1}(\{(0,0)\})$, wobei $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ gegeben ist durch

$$f(x,y,z) = \begin{pmatrix} x^2 + y^2 - b^2 z^2 \\ z - (1 + \mu x) \end{pmatrix}.$$

Wie in i) lassen sich alle diese Schnittkurven offenbar ebenfalls lokal als Graph von Funktionen $(y,z) = h(x)$, bzw. $(x,z) = l(y)$ bzgl. x oder y schreiben. Beachte, dass h vektorwertig ist und ebensoviele Komponenten besitzt wie f.



iii) Sei $f(x,y) = x^2 - y^3$, $(x,y) \in \mathbb{R}^2$. Die implizit durch $\Gamma = f^{-1}(\{0\})$ gegebene Kurve hat eine Spitze bei $x = y = 0$.



Wodurch unterscheiden sich diese Beispiele? Gibt es eine allgemeine Theorie?

Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega; \mathbb{R}^l)$, $p_0 \in \Omega$.

Definition 7.8.1. *Der* **Rang** *von $df(p_0)$ ist die Dimension des Bildraumes*

$$df(p_0)(\mathbb{R}^n) = \{df(p_0)\xi;\ \xi \in \mathbb{R}^n\} \subset \mathbb{R}^l.$$

Bemerkung 7.8.1. Offenbar gilt stets

$$Rang(df(p_0)) \leq \min\{n, l\},$$

und Gleichheit gilt in folgenden Fällen:

$$\begin{aligned} n \le l: &\text{ falls } df(p_0) \text{ injektiv,} \\ n \ge l: &\text{ falls } df(p_0) \text{ surjektiv,} \\ n = l: &\text{ falls } df(p_0) \text{ bijektiv.} \end{aligned}$$

Definition 7.8.2. *Der Punkt p_0 heisst* **regulärer Punkt** *von f, falls*

$$Rang(df(p_0)) = \min\{n, l\} \ ,$$

d.h., falls der Rang von $df(p_0)$ maximal ist.

Falls $n = l$, so ist f in der Nähe eines regulären Punktes invertierbar nach Satz 7.7.1. Im folgenden interessiert uns jedoch der Fall $n > l$. Wir betrachten erneut die Beispiele 7.8.1.i)-iii).

Beispiel 7.8.2. i) Für $f(x, y) = x^2 + y^2 - 1$ gilt

$$df(x, y) = (2x, 2y) \neq (0, 0), \quad \forall (x, y) \in f^{-1}(\{0\});$$

d.h. jedes $p_0 = (x_0, y_0)$ mit $f(p_0) = 0$ ist regulär.

ii) Für die Funktion

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 - 1 \\ z - (1 + \mu x) \end{pmatrix} : \mathbb{R}^3 \to \mathbb{R}^2$$

hat

$$df(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ -\mu & 0 & 1 \end{pmatrix}$$

den Rang 2 für alle $p_0 = (x_0, y_0, z_0) \in f^{-1}(\{(0, 0)\})$; diese Punkte sind also allesamt regulär.

iii) Für $f(x, y) = x^2 - y^3$ mit

$$df(x, y) = (2x, -3y^2)$$

ist der Punkt $p_0 = (0, 0)$ mit $f(p_0) = 0$ **nicht** regulär.

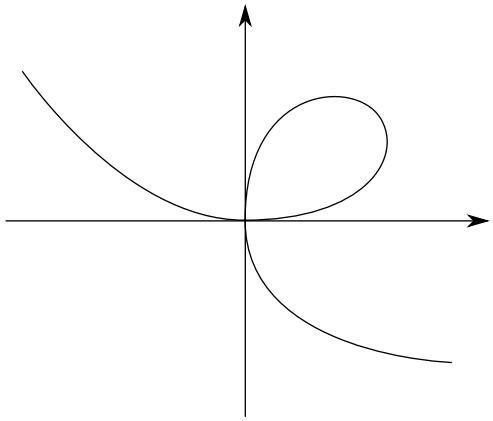
iv) Sei $f(x, y) = x^3 + y^3 - 3xy$ mit

$$df(x, y) = 3(x^2 - y, y^2 - x), \quad (x, y) \in \mathbb{R}^3.$$

Beachte:

$$df(x, y) = (0, 0) \Leftrightarrow x^2 = y \text{ und } y^2 = x \ ;$$

in einem nicht regulären Punkt (x, y) gilt also die Gleichung $x = x^4$. Somit sind $(x_0, y_0) = (0, 0)$ sowie $(x, y) = (1, 1)$ die einzigen nicht regulären Punkte von f. Die Kurve $\Gamma = f^{-1}(\{0\})$ heisst **Descartesches Blatt**. Offenbar ist der Punkt $p_0 = (0, 0)$ der einzige Punkt in Γ, wo Γ nicht lokal als Graph beschrieben werden kann.



Der folgende Satz liefert die Erklärung für den in den obigen Beispielen zutage tretenden Zusammenhang zwischen regulären Punkten und der Existenz einer lokalen Darstellung der Niveau-Menge $f^{-1}(\{0\})$ als Graph.

Satz 7.8.1. *Sei $f \in C^1(\Omega;\mathbb{R}^l)$, $l < n$, und sei $p_0 \in \Omega$ regulär mit $f(p_0) = 0$. Wähle Koordinaten $(x,y) \in \mathbb{R}^k \times \mathbb{R}^l$, $k+l=n$, auf $\Omega \subset \mathbb{R}^n \cong \mathbb{R}^k \times \mathbb{R}^l$, so dass*

$$\partial_y f(p_0) = \Big(\frac{\partial f^i}{\partial y^j}(p_0)\Big)_{1\le i,j\le l} : \mathbb{R}^l \to \mathbb{R}^l$$

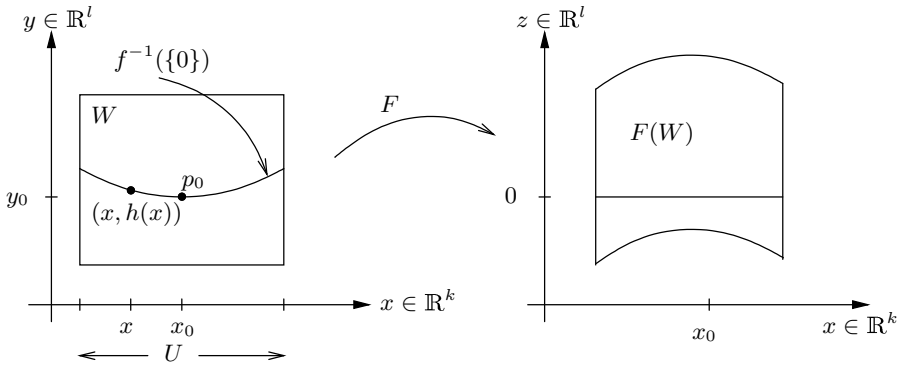
invertierbar. Sei $p_0 = (x_0, y_0)$ in diesen Koordinaten.

Dann gibt es Umgebungen U von x_0 in $\mathbb{R}^k$, W von p_0 in $\mathbb{R}^n$ und eine Funktion $h \in C^1(U;\mathbb{R}^l)$ mit $h(x_0) = y_0$, so dass gilt:

$$f(x,h(x)) = 0, \quad \forall x \in U, \tag{7.8.1}$$

und

$$f^{-1}(\{0\}) \cap W = \mathcal{G}(h) = \{(x,h(x));\ x \in U\}. \tag{7.8.2}$$



Beweis. OBdA sei $p_0 = 0 \in \mathbb{R}^n$. Betrachte die Abbildung $F \in C^1(\Omega; \mathbb{R}^n)$ mit

$$F\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ f(x,y) \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^l \cong \mathbb{R}^n.$$

Beachte

$$dF(p) = \begin{pmatrix} id_{\mathbb{R}^k} & 0 \\ \partial_x f(p) & \partial_y f(p) \end{pmatrix}, \quad p \in \Omega.$$

Es folgt

$$det(dF(p_0)) = det(\partial_y f(p_0)) \neq 0\,;$$

d.h., $dF(p_0)$ ist invertierbar. Nach dem Umkehrsatz Satz 7.7.1 gibt es Umgebungen $\tilde{U}$ von $p_0 = 0$ in Ω, $\tilde{V}$ von $F(p_0) = 0$ in $\mathbb{R}^n$ und eine lokale Inverse

$$G = (g_1, g_2) = (F\big|_{\tilde{U}})^{-1} \in C^1(\tilde{V}; \mathbb{R}^k \times \mathbb{R}^l)$$

von F. Mit der Darstellung von F folgt

$$(x,z) = F(G(x,z)) = \big(g_1(x,z), f(g_1(x,z), g_2(x,z))\big)$$

für alle $(x,z) \in \tilde{V}$.

Insbesondere erhalten wir für $z = 0$ die Identität

$$g_1(x,0) = x, \quad \forall x \in U := \{x;\ (x,0) \in \tilde{V}\} \tag{7.8.3}$$

und somit auch

$$f(x, g_2(x,0)) = 0, \quad \forall x \in U.$$

Für $h(x) := g_2(x,0) \in C^1(U; \mathbb{R}^l)$ folgt somit (7.8.1), wie gewünscht.

Mit

$$F(f^{-1}(\{0\}) \cap \tilde{U}) = \{(x,0) \in \tilde{V}\}$$

und (7.8.3) folgt nun

$$\begin{aligned} f^{-1}(\{0\}) \cap \tilde{U} &= G(\{(x,0) \in \tilde{V}) = \{G(x,0);\ (x,0) \in \tilde{V}\} \\ &= \{(x, h(x));\ x \in U\} = \mathcal{G}(h)\,. \end{aligned}$$

Bei Wahl von $W := \tilde{U}$ erhalten wir dann auch (7.8.2). □

Bemerkung 7.8.2. Mittels Kettenregel kann man aus (7.8.1) eine Gleichung für das Differential dh der "impliziten Funktion" h herleiten. Sei dazu $\Phi \in C^1(U; \mathbb{R}^n)$ die Funktion

$$\Phi(x) = \begin{pmatrix} x \\ h(x) \end{pmatrix}, \quad x \in U.$$

Dann folgt mit (7.8.1):

$$\begin{aligned} 0 = d(f \circ \Phi)(x) = df(x, h(x)) d\Phi(x) &= \Big(\partial_x f(x, h(x)), \partial_y f(x, h(x))\Big) \begin{pmatrix} id_{\mathbb{R}^k} \\ dh(x) \end{pmatrix} \\ &= \partial_x f(x, h(x)) + \partial_y f(x, h(x)) dh(x), \end{aligned}$$

also

$$dh(x) = -\big(\partial_y f(x, h(x))\big)^{-1} \partial_x f(x, h(x))\,. \tag{7.8.4}$$

Beispiel 7.8.3. Sei $f(x,y) = x + y + x^2 y$, $(x,y) \in \mathbb{R}^2$, mit

$$df(x,y) = (1 + 2xy, 1 + x^2).$$

Da $\frac{\partial f}{\partial y} \geq 1$, ist jeder Punkt (x,y) regulär. Die Gleichung $f(x,y) = 0$ definiert also lokal um $x = 0$ implizit eine Funktion $h = h(x)$ mit $h(0) = 0$. Gemäss (7.8.4) gilt

$$h'(x) = \frac{\partial h}{\partial x}(x) = -\frac{\partial_x f(x,h(x))}{\partial_y f(x,h(x))} = -\frac{1 + 2xh(x)}{1 + x^2}, \quad h(0) = 0\,;$$

d.h.

$$\big((1 + x^2)h(x)\big)' = (1 + x^2)h'(x) + 2xh(x) = -1\,.$$

Als Lösung dieses Anfangswertproblems erhalten wir

$$h(x) = -\frac{x}{1 + x^2}, \quad x \in \mathbb{R}.$$

Wir verifizieren leicht

$$f(x,h(x)) = x - \frac{x}{1 + x^2}(1 + x^2) = 0\,.$$

7.9 Extrema mit Nebenbedingungen

Auch diesen Abschnitt beginnen wir mit einem einfach zu durchschauenden Beispiel.

Beispiel 7.9.1. Sei $f(x,y) = x(1+y)$, $(x,y) \in \mathbb{R}^2$, und sei $g(x,y) = x^2 + y^2 - 1$,

$$S = \{(x,y) \in \mathbb{R}^2;\ g(x,y) = 0\} = S^1.$$

Parametrisiere S^1 via $\gamma(t) = (\cos(t), \sin(t))$, $t \in \mathbb{R}$. Notwendig für eine Extremalstelle von f an der Stelle $p_0 = \gamma(t_0) = (x_0, y_0) \in S^1$ ist die Bedingung

$$\begin{aligned} 0 = \frac{d}{dt} f(\gamma(t))\Big|_{t=t_0} &= \frac{d}{dt}\big(\cos(t)(1 + \sin(t))\big)\Big|_{t=0} \\ &= \cos^2(t_0) - \sin(t_0)(1 + \sin(t_0)) = 1 - 2y_0^2 - y_0\,; \end{aligned}$$

d.h., $y_0 = \frac{1}{2}$, $x_0 = \pm\frac{\sqrt{3}}{2}$, $f(p_0) = \pm\frac{3\sqrt{3}}{4}$ oder $y_0 = -1$, $x_0 = 0$ mit $f(p_0) = 0$. Offenbar ist $p_0 = \big(\frac{\sqrt{3}}{2}, \frac{1}{2}\big)$ die gesuchte Maximalstelle.

Allgemein sei $\Omega \subset \mathbb{R}^n$ und seien $f \in C^1(\Omega)$, $g \in C^1(\Omega; \mathbb{R}^l)$, $l < n$. Wir möchten f unter der Nebenbedingung $g(p) = 0$ maximieren; d.h. wir suchen

$$\max\{f(p);\ p \in \Omega,\ g(p) = 0\}.$$

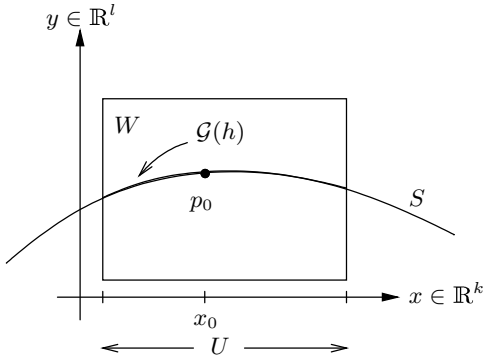
Kann man die gewünschten Extremalstellen auch ohne eine explizite Parametrisierung der "zulässigen Menge"

$$S = \{p \in \Omega;\ g(p) = 0\}$$

finden? Satz 7.8.1 liefert dafür notwendige Bedingungen.

Sei dazu $p_0 \in S$ eine lokale Maximalstelle von f in S. Nimm an, p_0 ist regulär für g. Wähle Koordinaten $(x, y) \in \mathbb{R}^k \times \mathbb{R}^l \cong \mathbb{R}^n$ um $p_0 = (x_0, y_0)$ wie in Satz 7.8.1, dazu Umgebungen $x_0 \in U \subset \mathbb{R}^k$, $p_0 \in W \subset \mathbb{R}^n$ und eine Funktion $h \in C^1(U; \mathbb{R}^l)$ mit

$$S \cap W = \mathcal{G}(h) = \{(x, h(x));\ x \in U\}.$$



Die Abbildung

$$\Phi(x) = (x, h(x)) \in C^1(U; \mathbb{R}^n)$$

liefert dann eine Parameterdarstellung für S nahe p_0. Falls p_0 ein lokales Maximum von f auf S liefert, so ist x_0 eine lokale Maximalstelle der Funktion $f \circ \Phi$ auf U. Mit Satz 7.5.3.i) und der Kettenregel folgt

$$0 = d\big(f \circ \Phi\big)(x_0) = df(p_0)d\Phi(x_0) = \partial_x f(p_0) + \partial_y f(p_0)dh(x_0).$$

Gemäss (7.8.4) in Bemerkung 7.8.2 gilt andererseits

$$dh(x_0) = -\partial_y g(p_0)^{-1}\partial_x g(p_0)\,.$$

Wir erhalten somit die Gleichung

$$0 = \partial_x f(p_0) + \lambda \partial_x g(p_0), \tag{7.9.1}$$

wobei λ die lineare Abbildung

$$\lambda = -\partial_y f(p_0)(\partial_y g(p_0))^{-1} \colon \mathbb{R}^l \to \mathbb{R} \tag{7.9.2}$$

bezeichnet, dargestellt durch $\lambda = (\lambda_1, \ldots, \lambda_l)$. Beachte, dass mit (7.9.2) automatisch auch gilt

$$0 = \partial_y f(p_0) + \lambda \partial_y g(p_0); \tag{7.9.3}$$

d.h., es gilt

$$0 = df(p_0) + \lambda dg(p_0)\,.$$

Wir haben gezeigt:

Satz 7.9.1. (Lagrange-Multiplikatorenregel) *Sei $p_0 \in S$ lokales Maximum oder Minimum von f unter der Nebenbedingung $g(p_0) = 0$, und sei p_0 regulärer Punkt von g. Dann existiert $\lambda = (\lambda_1, \ldots, \lambda_l) \in \mathbb{R}^l$, so dass für $L = f + \lambda g$ gilt*

$$dL(p_0) = df(p_0) + \lambda dg(p_0) = 0.$$

Definition 7.9.1. i) *Der Vektor $\lambda = (\lambda_1, \ldots, \lambda_l) = \lambda(p_0) \in \mathbb{R}^l$ mit (7.9.2) heisst* **Lagrange-Multiplikator**.

ii) *Die Funktion $L = f + \lambda p$ mit $\lambda = \lambda(p_0)$ heisst* **Lagrangefunktion** *(am Punkt p_0).*

iii) *Der Punkt $p_0 \in S$ heisst* **kritischer Punkt von f auf S**, *falls $dL(p_0) = 0$. Mit (7.9.3) erhalten wir dann auch die Charakterisierung (7.9.2) von $\lambda = \lambda(p_0)$.*

Beispiel 7.9.2. i) Sei $f(x, y) = x(1+y)$ wie in Beispiel 7.9.1. Satz 7.9.1 ergibt als **notwendige Bedingung** für das Vorliegen einer Maximalstelle unter der Nebenbedingung

$$g(x, y) = x^2 + y^2 - 1 = 0 \tag{7.9.4}$$

in (x_0, y_0) die Gleichung

$$0 = d(f + \lambda g)(x_0, y_0) = (1 + y_0, x_0) + 2\lambda(x_0, y_0). \tag{7.9.5}$$

Aus den 3 Gleichungen (7.9.4), (7.9.5) können wir $(x_0, y_0) \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ wie folgt bestimmen:

$$(7.9.5) \Rightarrow 1 + y_0 + 2\lambda x_0 = 0, \;\; x_0 + 2\lambda y_0 = 0.$$

Falls $y_0 = 0$, so folgt $x_0 = -2\lambda y_0 = 0$, also $0 = g(x_0, y_0) = -1$, und wir erhalten einen Widerspruch. Also gilt $y_0 \neq 0$, $\lambda = -\frac{x_0}{2y_0}$ und

$$1 + y_0 - \frac{x_0^2}{y_0} = 0;$$

d.h.

$$0 = y_0^2 + y_0 - x_0^2 \overset{(7.9.4)}{=} 2y_0^2 + y_0 - 1$$

wie in Beispiel 7.9.1, und

$$y_0 = -\frac{1}{4} \pm \sqrt{\frac{1}{16} + \frac{1}{2}} = \frac{-1 \pm 3}{4} \in \{\frac{1}{2}, -1\}.$$

Schliesslich ergibt (7.9.4):

$$\begin{aligned} y_0 = \frac{1}{2} \Rightarrow x_0 = \pm\frac{\sqrt{3}}{2} = -\lambda, \quad f(x_0, y_0) = \pm\frac{3\sqrt{3}}{2}, \\ y_0 = -1 \Rightarrow x_0 = 0 = \lambda, \quad f(x_0, y_0) = 0. \end{aligned}$$

Die Funktion f nimmt daher auf S^1 ihr Maximum an im Punkt $\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$, ihr Minimum im Punkt $\left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)$.

ii) Sei $f(x,y) = y$, $g(x,y) = x^2 - y^3$, $(x,y) \in \mathbb{R}^2$. Offenbar gilt

$$f(0,0) = 0 \leq f(x,y), \quad \forall (x,y) \in S = g^{-1}(\{0\}),$$

jedoch gilt für $L = f + \lambda g$ stets

$$dL(x,y) = (2xy, 1 - 2\lambda y^2) \neq 0,$$

falls $x^2 = y^3$; Satz 7.9.1 versagt also. Der **Grund** ist natürlich, dass der Punkt $p_0 = (0,0)$ nicht regulär ist für g.

Bemerkung 7.9.1. i) Wir können Satz 7.9.1 auch geometrisch deuten. Seien dazu $p_0 = (x_0, y_0) \in S = g^{-1}(\{0\})$, U, W, $h \in C^1(U, \mathbb{R}^l)$ wie in Satz 7.8.1 mit

$$S \cap W = \mathcal{G}(h) = \{(x, h(x));\ x \in U\}.$$

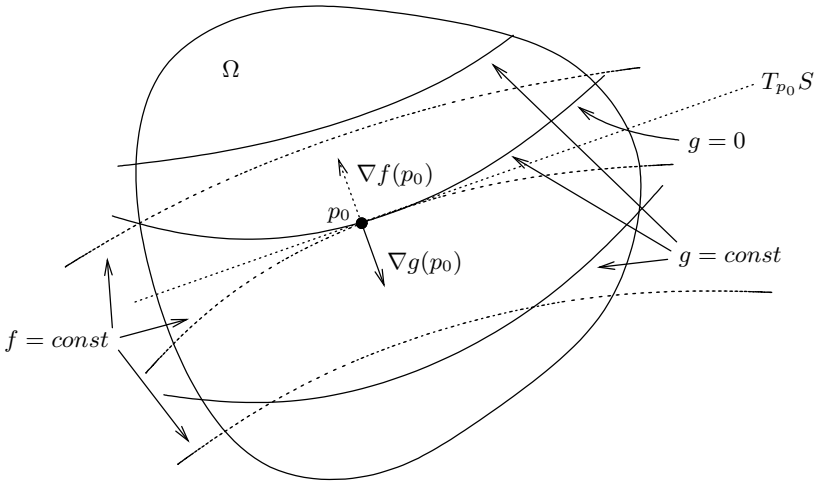
Mittels

$$\Phi(x) = (x, h(x)) \colon U \to S \subset \mathbb{R}^n$$

erhalten wir die Darstellung

$$T_{p_0}S = \{d\Phi(x_0)\xi = (\xi, dh(x_0)\xi);\ \xi \in \mathbb{R}^k\} \cong \mathbb{R}^k \tag{7.9.6}$$

des **Tangentialraums** an S im Punkt $p_0 = \Phi(x_0)$.



Da $g \circ \Phi \equiv 0$, folgt mit

$$0 = d(g \circ \Phi)(x_0) = dg(p_0)d\Phi(x_0)$$

sofort, dass $T_{p_0}S = im(d\Phi(x_0)) \subset ker(dg(p_0))$; d.h.

$$dg(p_0)\eta = 0, \quad \forall \eta \in T_{p_0}S. \tag{7.9.7}$$

Da wegen der Rangformel zudem gilt

$$dim(ker(dg(p_0))) = n - l = k = dim(T_{p_0}S)\,,$$

folgt sogar die Gleichheit

$$T_{p_0}S = ker(dg(p_0))\,.$$

Andererseits gilt an einer Maximalstelle $p_0 = \Phi(x_0)$ von f auf S die Bedingung $d(f \circ \Phi)(x_0) = 0$; d.h.

$$df(p_0)\eta = 0, \quad \forall \eta \in T_{p_0}S; \tag{7.9.8}$$

falls nämlich $df(p_0)\eta > 0$ für ein $\eta = (\xi, dh(x_0)\xi)$, so wäre

$$\frac{d}{d\epsilon}\Big|_{\epsilon=0} f\big(\Phi(x_0 + \epsilon\xi)\big) > 0$$

im Widerspruch zu Korollar 5.5.1.i).

ii) Falls $l = 1$, so sind (7.9.7) und (7.9.8) äquivalent zu den Bedingungen

$$\langle \nabla g(p_0), \eta \rangle_{\mathbb{R}^n} = 0, \quad \forall \eta \in T_{p_0}S,$$

bzw.

$$\langle \nabla f(p_0), \eta \rangle_{\mathbb{R}^n} = 0, \quad \forall \eta \in T_{p_0}S,$$

wobei $\nabla f(p_0)$, $\nabla g(p_0)$ die Gradienten von f, bzw. g am Punkt p_0 bezeichnen. Da $T_{p_0}S \cong \mathbb{R}^{n-1}$, sind $\nabla f(p_0)$, $\nabla g(p_0)$ notwendig proportional zum Normalvektor auf $T_{p_0}S$, und es existiert $\lambda \in \mathbb{R}$ mit

$$\nabla f(p_0) + \lambda \nabla g(p_0) = 0.$$

Beispiel 7.9.3. Sei $S^n = \{x \in \mathbb{R}^{n+1};\ |x| = 1\}$. Es gilt $S^n = g^{-1}(\{0\})$, wobei $g(x) = |x|^2 - 1$ mit

$$dg(x) = 2x^t \neq 0, \quad \forall x \in S^n.$$

Es folgt

$$T_x S^n = \{\xi \in \mathbb{R}^{n+1};\ x \cdot \xi = 0\}.$$

Bemerkung 7.9.2. Analog zu Satz 7.5.3.ii) erhalten wir auch eine hinreichende Bedingung für das Vorliegen eines Maximums in einem kritischen Punkt von f in $S = g^{-1}(\{0\})$.

Satz 7.9.2. *Sei $p_0 \in S$ regulär und sei p_0 kritischer Punkt von f mit Lagrange-Multiplikator $\lambda = \lambda(p_0) \in \mathbb{R}^l$, $L = f + \lambda g$ die zugehörige Lagrange-Funktion. Falls*

$$Hess_L(p_0)(\eta, \eta) > 0$$

für alle $\eta \in T_{p_0}S \backslash \{0\}$, so ist p_0 ein striktes relatives Minimum von f auf S.

Kapitel 8

Integration im $\mathbb{R}^n$

8.1 Riemannsches Integral über einem Quader

Zur Definition des R-Integrals über einem n-dimensionalen Quader gehen wir vollkommen analog vor wie im Fall $n = 1$; vgl. Abschnitt 6.2.

Definition 8.1.1. i) *Ein n-dimensionaler* **Quader** *ist ein Produkt*

$$Q = \prod_{i=1}^{n} I_i = \{x = (x^i)_{1\le i\le n};\ x^i \in I_i,\ 1 \le i \le n\}$$

von (offenen, abgeschlossenen, oder halb-offenen) Intervallen $I_1, \dots, I_n$. Solch ein Q hat den **Elementarinhalt**

$$\mu(Q) = \prod_{i=1}^{n} |I_i| .$$

ii) *Eine* **Zerlegung** $P = \{Q_k;\ 1 \le k \le K\}$ *eines Quaders* $Q = \bigcup_{k=1}^{K} Q_k$ *in disjunkte Teilquader $Q_k \subset Q$, $1 \le k \le K$, hat die* **Feinheit**

$$\delta_P = \max_{1\le k\le K} diam\ Q_k,$$

wobei

$$diam\ Q_k = \sup_{x,y\in Q_k} |x - y| , \quad 1 \le k \le K,$$

den **Durchmesser** *von Q_k bezeichnet.*

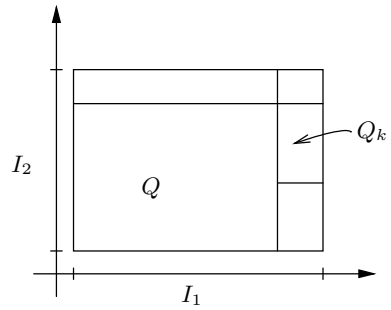
iii) *Eine Funktion $f\colon Q \to \mathbb{R}$ auf einem Quader Q heisst* **Treppenfunktion**, *falls f eine Darstellung der Form*

$$f = \sum_{k=1}^{K} c_k \chi_{Q_k}$$

besitzt mit einer Zerlegung $P = \{Q_k;\ 1 \le k \le K\}$ *von* Q *und Konstanten* $c_k \in \mathbb{R}$, $1 \le k \le K$.

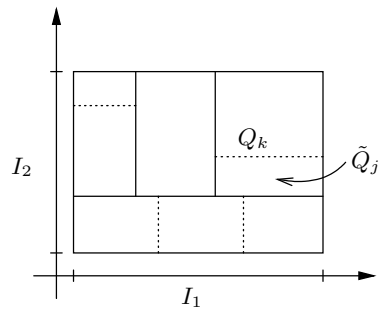
iv) *Das* **Riemann-Integral** *(R-Integral) einer Treppenfunktion* $f = \sum_{k=1}^{K} c_k \chi_{Q_k}$ *wie in iii) ist dann wie folgt definiert*

$$\int_Q f \, d\mu = \int_Q \Big(\sum_{k=1}^{K} c_k \ \chi_{Q_k} \Big) \, d\mu = \sum_{k=1}^{K} c_k \ \mu(Q_k). \tag{8.1.1}$$



Bemerkung 8.1.1. Analog zu Bemerkung 6.2.1 ist die Definition des R-Integrals einer Treppenfunktion $f\colon Q \to \mathbb{R}$ unabhängig von der gewählten Darstellung $f = \sum_{k=1}^{K} c_k \ \chi_{Q_k}$; insbesondere ändert sich der Wert der Summe (8.1.1) nicht bei "Verfeinerungen" der Zerlegung, die wir wie folgt definieren.

Definition 8.1.2. *Eine Zerlegung* $\tilde{P} = \{\tilde{Q}_j;\ 1 \le j \le J\}$ *ist eine* **Verfeinerung** *der Zerlegung* $P = \{Q_k;\ 1 \le k \le K\}$ *des Quaders* Q*, falls jedes* $\tilde{Q}_j$ *in einem Quader* Q_k *enthalten ist.*



Beispiel 8.1.1. Seien $P = \{Q_k;\ 1 \le k \le K\}$, $R = \{S_l;\ 1 \le l \le L\}$ Zerlegungen von Q. Dann ist

$$\tilde{P} = \{Q_k \cap S_l;\ 1 \le k \le K,\ 1 \le l \le L\}$$

Zerlegung von Q, welche sowohl P als auch R verfeinert. (Vgl. den Beweis von Lemma 6.2.1.)

Sei $Q \subset \mathbb{R}^n$ ein Quader, $f\colon Q \to \mathbb{R}$ beschränkt. Analog zu Definition 6.2.2 können wir nun das R-Integral von f definieren.

Definition 8.1.3. i) *Das* **untere**, *bzw.* **obere** *R-Integral von f sind erklärt durch*

$$\underline{\int_Q} f \, d\mu = \sup\{\int_Q fe \, d\mu;\ e \text{ Treppenfunktion};\ e \le f\},$$

bzw.

$$\overline{\int_Q} f \, d\mu = \inf\{\int_Q g \, d\mu;\ g \text{ Treppenfunktion};\ f \le g\}.$$

ii) *Die Funktion f heisst* **R-integrabel** *über Q, falls*

$$\underline{\int_Q} f \, d\mu = \overline{\int_Q} f \, d\mu =: \int_Q f \, d\mu.$$

Bemerkung 8.1.2. i) Analog zu Bemerkung 6.2.2 gilt (unter Verwendung von Beispiel 8.1.1) für jedes beschränkte f die Ungleichung

$$\underline{\int_Q} f \, d\mu \le \overline{\int_Q} f \, d\mu.$$

ii) Weiter ist f R-integrabel genau dann, wenn für jedes $\epsilon > 0$ Treppenfunktionen $e, g\colon Q \to \mathbb{R}$ existieren mit $e \le f \le g$ und

$$\int_Q g \, d\mu - \int_Q e \, d\mu < \epsilon.$$

Vollkommen analog zu den Sätzen 6.2.2 und 6.2.3 erhalten wir sodann die folgende Aussage

Satz 8.1.1. *Sei $f \in C^0(\overline{Q})$. Dann ist f über Q R-integrabel, und für jede Folge $(P^{(l)})_{l\in\mathbb{N}}$ von Zerlegungen $P^{(l)} = \{Q_k^{(l)};\ 1 \le k \le K^{(l)}\}$ von Q mit Feinheit $\delta_{P^{(l)}} \to 0 \ \ (l \to \infty)$ gilt für eine beliebige Auswahl von Punkten $x_k^{(l)} \in Q_k^{(l)}$, $1 \le k \le K^{(l)}$, $l \in \mathbb{N}$, stets*

$$\int_Q \Big(\sum_{k=1}^{K^{(l)}} f(x_k^{(l)})\ \chi_{Q_k^{(l)}}\Big)\, d\mu = \sum_{k=1}^{K^{(l)}} f(x_k^{(l)})\ \mu(Q_k^{(l)}) \to \int_Q f \, d\mu \ \ (l \to \infty).$$

Weiter gelten Linearität und Monotonie des R-Integrals analog zu den Sätzen 6.3.1 und 6.3.2, und wir erhalten die den Korollaren 6.3.1-6.3.3 entsprechenden Aussagen.

Satz 8.1.2. *Seien $f, f_1, f_2 \colon Q \to \mathbb{R}$ beschränkt und über Q R-integrabel, und sei $\alpha \in \mathbb{R}$. Dann sind die Funktionen αf, $f_1 + f_2 \colon Q \to \mathbb{R}$ über Q R-integrabel, und*

$$\int_Q (\alpha f)\ d\mu = \alpha \int_Q f\ d\mu,$$

bzw.

$$\int_Q (f_1 + f_2)\ d\mu = \int_Q f_1\ d\mu + \int_Q f_2\ d\mu.$$

Satz 8.1.3. *Seien $f, g \colon Q \to \mathbb{R}$ beschränkt und R-integrabel, und sei $f \leq g$. Dann gilt*

$$\int_Q f\ d\mu \leq \int_Q g\ d\mu.$$

Insbesondere gilt für $f \in C^0(\overline{Q})$ die Abschätzung

$$\left| \int_Q f\ d\mu \right| \leq \int_Q |f|\ d\mu \leq \sup_Q |f| \cdot \mu(Q).$$

(Beachte, dass gemäss Satz 8.1.1 für $f \in C^0(\overline{Q})$ stets sowohl f als auch $|f|$ über Q R-integrabel sind.)

Kombination von Satz 8.1.2 mit Satz 8.1.3 ergibt

Korollar 8.1.1. *Seien $f, f_k \in C^0(\overline{Q})$ mit $f_k \overset{glm.}{\to} f$ $(k \to \infty)$. Dann gilt*

$$\left| \int_Q f_k\ d\mu - \int_Q f\ d\mu \right| \leq \int_Q |f_k - f|\ d\mu \leq \|f_k - f\|_{C^0} \cdot \mu(Q) \overset{(k\to\infty)}{\to} 0.$$

Schliesslich gilt auch Satz 6.3.3 analog.

Satz 8.1.4. (Gebietsadditivität) *Sei $f \colon Q \to \mathbb{R}$ beschränkt und über Q R-integrabel, und sei $P = \{Q_k;\ 1 \leq k \leq K\}$ eine Zerlegung von Q in disjunkte Quader Q_k, $1 \leq k \leq K$. Dann gilt*

$$\int_Q f\ d\mu = \sum_{k=1}^{K} \int_{Q_k} f\ d\mu.$$

8.2 Der Satz von Fubini

Soweit die Theorie; wie kann man jedoch das R-integral konkret berechnen? Ausser im Falle von Treppenfunktionen gelingt dies mit den Mitteln aus Abschnitt 8.1 allenfalls approximativ; vgl. Satz 8.1.1.

Der folgende Satz hilft uns weiter.

Satz 8.2.1. (Fubini) *Sei $Q = [a,b] \times [c,d] \subset \mathbb{R}^2$, und sei $f \in C^0(Q)$. Dann gilt*

$$\int_Q f \; d\mu = \int_a^b \Big(\int_c^d f(x,y) \; dy \Big) \; dx = \int_c^d \Big(\int_a^b f(x,y) \; dx \Big) \; dy.$$

D.h. das Integral von f über Q kann iterativ durch 1-dimensionale Integration bestimmt werden!

Beispiel 8.2.1. Sei $Q = [0,1] \times [0,2\pi]$, $f(x,y) = \sin(y-x) \in C^0(Q)$. Nach Satz 8.2.1 gilt

$$\int_Q f \; d\mu = \int_0^1 \Big(\underbrace{\int_0^{2\pi} \sin(y-x) \; dy}_{=0} \Big) \; dx = 0.$$

Bemerkung 8.2.1. Die Voraussetzung $f \in C^0(Q)$ in Satz 8.2.1 ist wichtig. Insbesondere kann man für allgemeine (beschränkte) Funktionen $f \colon Q \to \mathbb{R}$ aus der Existenz eines der iterierten Integrale **nicht** auf die Existenz der R-Integrale $\int_Q f \; d\mu$ schliessen. Dies zeigt das folgende

Beispiel 8.2.2. Sei $Q = [0,1] \times [0,2\pi] \subset \mathbb{R}^2$,

$$f(x,y) = \chi_Q(x) \cdot \sin(y-x) \colon Q \to \mathbb{R}.$$

Es gilt

$$\int_0^{2\pi} f(x,y) \; dy = \chi_Q(x) \int_0^{2\pi} \sin(y-x) \; dy = 0$$

für alle $x \in [0,1]$; jedoch ist f über Q nicht R-integrabel, und auch $\int_0^1 f(x,y) \; dx$ existiert für kein $y \in [0,2\pi]$.

Beweis von Satz 8.2.1. Seien

$$P_1 = \{I_{1j};\ 1 \le j \le J\},\ P_2 = \{I_{2k};\ 1 \le k \le K\}$$

Zerlegungen von $I_1 = [a,b]$, bzw. $I_2 = [c,d]$.

Die zugehörige **Produktzerlegung**

$$P = P_1 \times P_2 = \{Q_{jk} = I_{1j} \times I_{2k};\ 1 \le j \le J,\ 1 \le k \le K\}$$

hat offenbar die Feinheit

$$\delta_P \le \sqrt{2} \max\{\delta_{P_1}, \delta_{P_2}\}.$$

Für $x \in I_1$ setze

$$g(x) = \int_c^d f(x,y) \; dy.$$

Behauptung $g \in C^0(I_1)$.

Beweis. Wir benutzen das Folgenkriterium. Für $(x_k)_{k\in\mathbb{N}} \in I_1$ mit $x_k \to x_0$ $(k \to \infty)$ gilt gemäss Korollar 6.3.2

$$\begin{aligned}|g(x_k) - g(x_0)| &= \left| \int_c^d \big(f(x_k, y) - f(x_0, y)\big)\, dy \right| \\ &\le \sup_{c\le y\le d} |f(x_k, y) - f(x_0, y)| \cdot |d - c| \to 0 \quad (k \to \infty),\end{aligned}$$

da f auf Q gemäss Satz 4.7.3 gleichmässig stetig ist. □

Für beliebig gewählte Punkte $x_j \in I_{1j}$, $1 \le j \le J$, $y_k \in I_{2k}$, $1 \le k \le K$, gilt nun gemäss Satz 8.1.1, bzw. Satz 6.2.3

$$\begin{aligned}\int_Q f\, d\mu &\overset{\text{(Satz 8.1.1)}}{=} \lim_{\delta_{P_1},\delta_{P_2}\to 0} \Big(\sum_{j,k} f(x_j, y_k) \underbrace{\mu(Q_{jk})}_{=|I_{1j}|\cdot|I_{2k}|} \Big) \\ &= \lim_{\delta_{P_1}\to 0} \sum_{j=1}^{J} \underbrace{\lim_{\delta_{P_2}\to 0} \Big(\sum_{k=1}^{K} f(x_j, y_k)\, |I_{2k}| \Big)}_{\overset{\text{(Satz 6.2.3)}}{=} \int_c^d f(x_j,y)\ dy = g(x_j)} \cdot |I_{1j}| \\ &= \lim_{\delta_{P_1}\to 0} \sum_{j=1}^{J} g(x_j)\, |I_{1j}| \overset{\text{(Satz 6.2.3)}}{=} \int_a^b g(x)\, dx = \int_a^b \Big(\int_c^d f(x, y)\, dx \Big)\, dx.\end{aligned}$$

Die 2. Identität erhält man analog nach Vertauschen von x und y. □

Beispiel 8.2.3. i) Sei Q der Würfel $Q = [0, 1] \times [0, 1] = [0, 1]^2$. Mit dem Additionstheorem $e^{x+y} = e^x \cdot e^y$ und Satz 8.2.1 erhalten wir

$$\begin{aligned}\int_Q e^{x+y}\, d\mu &= \int_0^1 \Big(\int_0^1 e^{x+y}\, dy \Big)\, dx \\ &= \int_0^1 e^x \cdot \Big(\int_0^1 e^y\, dy \Big)\, dx = \Big(\int_0^1 e^x \Big)^2 = (e - 1)^2.\end{aligned}$$

ii) Ebenso erhalten wir bei geschickter Wahl der Integrationsreihenfolge und Substitution

$$\begin{aligned}\int_{[0,1]^2} y e^{xy}\, d\mu &= \int_0^1 \Big(\int_0^1 y e^{xy}\, dx \Big)\, dy \\ &\overset{(z=xy)}{=} \int_0^1 \Big(\int_0^y e^z\, dz \Big)\, dy = \int_0^1 (e^y - 1)\, dy = e - 2.\end{aligned}$$

Satz 8.2.1 gilt analog auch in höheren Dimensionen

Satz 8.2.2. *Sei $Q = \prod_{i=1}^{n} [a_i, b_i]$, $f \in C^0(Q)$. Dann gilt*

$$\int_Q f\, d\mu = \int_{a_1}^{b_1} \Big(\int_{a_2}^{b_2} \Big(\dots \Big(\int_{a_n}^{b_n} f(x^1, \dots, x^n)\, dx^n \Big) \dots \Big)\, dx^2 \Big)\, dx^1,$$

und die Reihenfolge der Integration darf beliebig vertauscht werden.

Beispiel 8.2.4. i) Das folgende Integral kann elementar berechnet werden

$$\int_{[0,1]^3} xy^2z^3 \, d\mu = \int_0^1 \Big(\int_0^1 \Big(\int_0^1 xy^2z^3 \, dz \Big) \, dy \Big) \, dx$$
$$= \int_0^1 x \, dx \cdot \int_0^1 y^2 \, dy \cdot \int_0^1 z^3 \, dz = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{24}.$$

ii) Im nächsten Beispiel kommt es wieder auf die geschickte Wahl der Integrationsreihenfolge an.

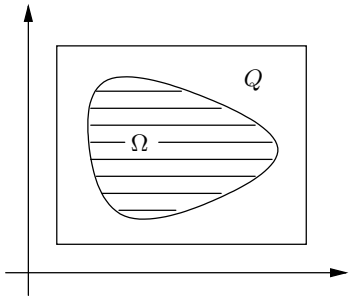
$$\int_{[0,\pi]\times[0,1]^2} x^2 y \cos(xyz) \, d\mu = \int_0^\pi \Big(\int_0^1 \Big(\underbrace{\int_0^1 x^2 y \cos(xyz) \, dz}_{\overset{(t=xyz)}{=} \int_0^{xy} x \cos t \; dt = x \sin(xy)} \Big) \, dy \Big) \, dx$$
$$= \int_0^\pi \Big(\underbrace{\int_0^1 x \sin(xy) \, dy}_{\overset{(s=xy)}{=} \int_0^x \sin s \; ds = 1 - \cos x} \Big) \, dx = \int_0^\pi (1 - \cos x) \, dx = \pi - \sin \pi = \pi.$$

8.3 Jordan-Bereiche

Mit dem in Abschnitt 8.1 eingeführten Integralbegriff können wir nun auch gewisse krummlinig berandete Gebiete $\Omega \subset \mathbb{R}^n$ "ausmessen".

Sei $\Omega \subset \mathbb{R}^n$ beschränkt, $Q \subset \mathbb{R}^n$ ein Quader mit $\Omega \subset Q$, χ_Ω die charakteristische Funktion

$$\chi_\Omega(x) = \begin{cases} 1, & x \in \Omega \\ 0, & \text{sonst} \end{cases}.$$



Definition 8.3.1. *Die Menge Ω heisst* **Jordan-messbar**, *falls χ_Ω über Q R-integrabel ist. In diesem Fall ist*

$$\mu(\Omega) = \int_Q \chi_\Omega \, d\mu$$

das n-dimensionale **Jordansche Mass** *von Ω.*

Bemerkung 8.3.1. Wegen Satz 8.1.4 ist die Definition 8.3.1 unabhängig von der Wahl von Q.

Beispiel 8.3.1. i) Jeder Quader $Q \subset \mathbb{R}^n$ ist Jordan-messbar.

ii) Die Vereinigung $\Omega = \bigcup_{k=1}^{K} Q_k$ disjunkter Quader Q_k, $1 \le k \le K$, ist Jordan-messbar. (Die Funktion $\chi_\Omega = \sum_{i=1}^{K} \chi_{Q_k}$ ist als Treppenfunktion R-integrabel.)

iii) Der Rhombus

$$\Omega = \{(x,y) \in \mathbb{R}^2;\ |x| + |y| \le 1\}$$

ist Jordan-messbar. Gemäss Satz 8.2.1 gilt

$$\int_{[-1,1]^2} \chi_\Omega \, d\mu^{(2)} = \int_{-1}^{1} \Big(\int_{-1}^{1} \chi_\Omega(x,y)\, dx \Big)\, dy = 4 \int_0^1 (1-y)\, dy = 2.$$

Wir können also den Flächeninhalt von Ω berechnen, indem wir die die Längen der Schnittmengen $\Omega_y = \{x;\ (x,y) \in \Omega\}$ bestimmen und bzgl. y aufintegrieren ("Cavalierisches Prinzip").

Definition 8.3.2. *Ein $\Omega = \bigcup_{k=1}^{K} Q_k$ mit disjunkten Quadern Q_k, $1 \le k \le K$, heisst* **Elementarfigur**.

Bemerkung 8.3.2. i) Gemäss Beispiel 8.3.1.ii) sind Elementarfiguren Jordan-messbar.

ii) Ein beschränktes $\Omega \subset \mathbb{R}^n$ ist gemäss Bemerkung 8.1.2.ii) Jordan-messbar genau dann, wenn zu jedem $\epsilon > 0$ Elementarfiguren $E, G \subset \mathbb{R}^n$ existieren mit $E \subset \Omega \subset G$ und

$$\mu(G \backslash E) = \mu(G) - \mu(E) < \epsilon, \tag{8.3.1}$$

also wenn

$$\mu(\partial\Omega) = 0.$$

In diesem Fall gilt

$$\mu(\Omega) = \inf\{\mu(G);\ G \supset \Omega \text{ El.Fig.}\} = \sup\{\mu(E);\ E \subset \Omega \text{ El.Fig.}\} \tag{8.3.2}$$

Beweis: Zum Beweis der Messbarkeit von Ω genügt es offenbar, in Bemerkung 8.1.2.ii) Treppenfunktionen $e \le \chi_\Omega \le g$ mit Werten 0 oder 1 zu betrachten, also $e = \chi_E$, $g = \chi_G$ für Elementarfiguren $E \subset \Omega \subset G$. Weiter gilt in diesem Fall

$$\chi_{G\backslash E} = \chi_G - \chi_E,$$

also

$$\mu(G\backslash E) = \mu(G) - \mu(E),$$

und (8.3.1) folgt. Die Identität (8.3.2) ergibt sich analog aus Definition 8.1.3.

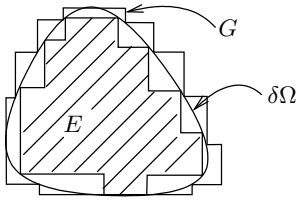
Schliesslich gilt $\mu(\partial\Omega) = 0$ genau dann, wenn zu jedem $\epsilon > 0$ eine Elementarfigur $U \subset \mathbb{R}^n$ existiert mit $\partial\Omega \subset U$ und $\mu(U) < \epsilon$. Dann sind

$$E = \Omega \backslash U, \quad G = \Omega \cup U$$

Elementarfiguren mit $E \subset \Omega \subset G$, und $U = G \setminus E$, also

$$\mu(G \backslash E) = \mu(U) < \epsilon,$$

und umgekehrt.



Beispiel 8.3.2. i) Sei $\psi \in C^0([a,b])$, $\psi \geq 0$. Dann ist die Menge

$$\Omega = \Omega_\psi = \{(x,y) \in \mathbb{R}^2;\ a \leq x \leq b,\ 0 \leq y \leq \psi(x)\}$$

Jordan-messbar, und

$$\mu(\Omega_\psi) = \int_a^b \psi(x)\ dx.$$

(Es genügt anzunehmen, dass $\psi \in C^0_{pw}([a,b])$ stückweise stetig ist.)

Beweis: Die Funktion ψ ist gemäss Satz 6.2.2 R-integrabel; also existieren zu vorgegebenem $\epsilon > 0$ gemäss Bemerkung 6.2.2.ii) Treppenfunktionen $e, g\colon [a,b] \to \mathbb{R}$ mit $0 \leq e \leq \psi \leq g$ und

$$\int_a^b g\ dx - \int_a^b e\ dx < \epsilon.$$

Die Funktionen e und g definieren Elementarfiguren $E = \Omega_e$ und $G = \Omega_g$ mit $E \subset \Omega \subset G$, und

$$\mu(G) = \int_a^b g\ dx, \quad \mu(E) = \int_a^b e\ dx.$$

Also ist $\Omega = \Omega_\psi$ Jordan-messbar gemäss Bemerkung 8.3.1.ii) und

$$\mu(\Omega) = \inf\{\mu(G);\ G \supset \Omega \text{ El.fig}\} = \overline{\int_a^b} \psi\ dx = \int_a^b \psi\ dx.$$

ii) Seien $a, b > 0$, $\psi(x) = b\sqrt{1 - \frac{x^2}{a^2}} \in C^0([-a,a])$. Dann ist

$$\begin{aligned}\Omega_\psi &= \{(x,y) \in \mathbb{R}^2;\ |x| \leq a,\ 0 \leq y \leq \psi(x)\} \\ &= \{(x,y) \in \mathbb{R}^2;\ \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1,\ y \geq 0\}\end{aligned}$$

der obere Teil einer Ellipse mit Halbachsen a und b. Mit i) folgt

$$\mu(\Omega_\psi) = \int_{-a}^{a} b\sqrt{1-\frac{x^2}{a^2}}\; dx \overset{(x=a\sin\varphi)}{=} ab\int_{-\pi/2}^{\pi/2}\cos^2\varphi\; d\varphi = \frac{\pi}{2}ab.$$

Unter Verwendung von Satz 8.2.2 erhalten wir ein analoges Resultat auch in höheren Dimensionen.

iii) Sei $Q' \subset \mathbb{R}^{n-1}$ ein $(n-1)$-Quader, $\psi \in C^0_{pw}(\overline{Q'})$, $\psi \geq 0$. Dann ist die Menge

$$\Omega_\psi = \{(x', x^n) \in \mathbb{R}^n;\ x' \in Q',\ 0 \leq x^n \leq \psi(x')\}$$

Jordan-messbar, und mit Satz 8.2.2 folgt

$$\mu_n(\Omega_\psi) = \int_{Q'} \psi(x')\; d\mu_{n-1}(x').$$

Der Deutlichkeit halber bezeichnet hier μ_n das n-dimensionale Jordansche Mass.

iv) Insbesondere erhalten wir für $Q' = [-1,1]^2$,

$$\psi(x,y) = \sqrt{\max\{0, 1-x^2-y^2\}} \in C^0(Q'),$$

die obere Halbkugel

$$\Omega_\psi = \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 \leq 1,\ z \geq 0\}.$$

Mit iii) sowie Satz 8.2.1 folgt

$$\begin{aligned}\mu_3(\Omega_\psi) = \int_{[-1,1]^2} \psi(x,y)\; d\mu_2(x,y) &= \int_{-1}^{1}\Big(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2}\; dy\Big)\; dx \\ &\overset{(y=\sqrt{1-x^2}\sin\varphi)}{=} \int_{-1}^{1}(1-x^2)\underbrace{\Big(\int_{-\pi/2}^{\pi/2}\cos^2\varphi\; d\varphi\Big)}_{=\pi/2}\; dx = \frac{2\pi}{3}.\end{aligned}$$

Schliesslich definieren wir noch das R-Integral über Jordan-messbare Bereiche.

Sei $\Omega \subset \mathbb{R}^n$ geschränkt und Jordan-messbar, $f\colon \Omega \to \mathbb{R}$ beschränkt, und sei $Q \subset \mathbb{R}^n$ ein Quader mit $\Omega \subset Q$.

Definition 8.3.3. *f heisst* **R-integrabel über** Ω*, falls die durch $\overline{f}(x) = 0$ für $x \in Q\backslash\Omega$ fortgesetzte Funktion f über Q R-integrabel ist, und*

$$\int_\Omega f d\mu := \int_Q \overline{f}\; d\mu.$$

Bemerkung 8.3.3. Wegen Satz 8.1.4 ist die Definition unabhängig von Q.

Mit Satz 8.2.1, bzw. Satz 8.2.2 können wir für Mengen $\Omega = \Omega_\psi$ wie in Beispiel 8.3.2 das Integral aus Definition 8.3.3 auf iterierte 1-dimensionals Integrale zurückführen.

Beispiel 8.3.3. i) Sei $0 \leq \psi \in C^0([a,b])$, $f \in C^0(\Omega_\psi)$, wo

$$\Omega_\psi = \{(x,y) \in \mathbb{R}^2;\ a \leq x \leq b,\ 0 \leq y \leq \psi(x)\}.$$

Dann gilt

$$\int_{\Omega_\psi} f\ d\mu = \int_a^b \Big(\int_0^{\psi(x)} f(x,y)\ dy \Big)\ dx.$$

(Dies gilt für Treppenfunktionen ψ gemäss Satz 8.2.1, also auch für (stückweise) stetige Funktionen ψ.)

ii) Speziell für $\psi(x) = b\sqrt{1 - \frac{x^2}{a^2}} \in C^0([-a,a])$ wie in Beispiel 8.3.3 und die Funktion $f(x,y) = y$ erhalten wir

$$\int_{\Omega_\psi} f\ d\mu = \int_{-a}^a \int_0^{b\sqrt{1-\frac{x^2}{a^2}}} y\ dy = \frac{1}{2} \int_{-a}^a b^2 \big(1 - \frac{x^2}{a^2}\big)\ dx = \frac{2ab^2}{3}.$$

8.4 Der Satz von Green

Falls $\Omega \subset Q \subset \mathbb{R}^2$ glatt berandet, und falls f von der Form ist

$$f = \frac{\partial h}{\partial x} - \frac{\partial g}{\partial y}$$

mit $g, h \in C^1(\overline{\Omega})$, so lässt sich das Integral von f über Ω auf ein Randintegral zurückführen.

Beispiel 8.4.1. i) Sei $Q = [a,b] \times [c,d]$, $g \in C^1(Q)$. Mit Satz 8.2.1 folgt

$$-\int_Q \frac{\partial g}{\partial y}\ d\mu = \int_a^b \Big(-\int_c^d \frac{\partial g}{\partial y}(x,y)\ dy \Big)\ dx = \int_a^b \big(g(x,c) - g(x,d) \big)\ dx.$$

Analog erhalten wir für $h \in C^1(Q)$ die Gleichung

$$\int_Q \frac{\partial h}{\partial x}\ d\mu = \int_c^d \Big(\int_a^b \frac{\partial h}{\partial x}(x,y)\ dx \Big)\ dy = \int_c^d \big(h(b,y) - h(a,y) \big)\ dy;$$

also

$$\int_Q \Big(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \Big)\ d\mu = \int_{\gamma_1+\gamma_2-\gamma_3-\gamma_4} (g dx + h dy) = \int_{\partial Q} (g dx + h dy),$$

wobei

$$\gamma_1(x) = (x,c),\ \gamma_3(x) = (x,d),\ a \leq x \leq b,$$

sowie

$$\gamma_2(y) = (b,y),\ \gamma_4(y) = (a,y),\ c \leq y \leq d$$

den Rand von ∂Q parametrisieren. Die Teilstücke $\gamma_1, \ldots, \gamma_4$ werden dabei so aneinander gehängt, dass der zusammengesetzte Weg $\gamma_1 + \gamma_2 - \gamma_3 - \gamma_4$ eine Parametrisierung von ∂Q ergibt, die so orientiert ist, dass Ω stets zur Linken des Weges liegt.

ii) Insbesondere erhalten wir bei Wahl von $g(x,y) = -y$, $h(x,y) = 0$ den Flächeninhalt

$$\mu(Q) = -\int_Q \frac{\partial g}{\partial y}\, d\mu = \int_a^b \big(g(x,c) - g(x,d)\big)\, dx = (b-a)(d-c)$$

Analog können wir für eine grosse Klasse von Gebieten argumentieren.

Definition 8.4.1. i) $\Omega \subset \mathbb{R}^2$ *heisst* **Normalbereich bzgl.** y *der Klasse* C^1, *falls*

$$\Omega = \{(x,y) \in \mathbb{R}^2;\ a \le x \le b,\ \varphi(x) \le y \le \psi(x)\}$$

für geeignete $-\infty < a < b < \infty$ *und mit Funktionen* $\varphi \le \psi \in C^1([a,b])$. *Analog definieren wir einen Normalbereich bzgl.* x, *bzw. der Klasse* C^1_{pw}, C^k, *etc.*

ii) $\Omega \subset \mathbb{R}^2$ *heisst* **Normalbereich**, *falls* Ω *sowohl bzgl.* x *als auch bzgl.* y *ein Normalbereich ist.*

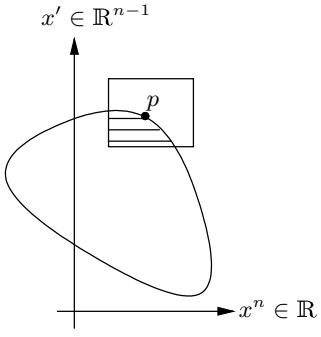
Beispiel 8.4.2. i) Ein bzgl. der Achsen gedrehter Quader ist ein Normalbereich der Klasse C^1_{pw}.

ii) $\overline{B_1(0)} = \{(x,y) \in \mathbb{R}^2;\ |y| \le \sqrt{1-x^2},\ |x| \le 1\} \subset \mathbb{R}^2$ ist ein Normalbereich der Klasse C^0.

iii) Der Kreisring $B_2 \backslash B_1(0)$ ist kein Normalbereich.

Definition 8.4.2. *Ein Gebiet* $\Omega \subset \mathbb{R}^n$ *ist* **von der Klasse** C^1 *(bzw.* C^1_{pw}, C^k*), falls zu jedem Punkt* $p \in \partial\Omega$ *Koordinaten* $(x', x^n) \in \mathbb{R}^{n-1} \times \mathbb{R}$, *ein Quader* $Q' \subset \mathbb{R}^{n-1}$, *eine Umgebung* $W = Q' \times]c,d[$ *von* p *und eine Funktion* $\psi \in C^1(Q')$ *(bzw.* $\psi \in C^1_{pw}(Q')$, $\psi \in C^k(Q')$*), existieren, so dass*

$$\Omega \cap W = \{(x', x^n) \in \mathbb{R}^n;\ x' \in Q',\ c < y < \psi(x)\}.$$



Beispiel 8.4.3. $B_1(0) \subset \mathbb{R}^2$ ist von der Klasse C^k für beliebiges $k \ge 0$.

Die folgende elementare Beobachtung wird später entscheidend benutzt.

Bemerkung 8.4.1. i) Jedes Gebiet $\Omega \subset Q \subset \mathbb{R}^2$ der Klasse C^1_{pw} kann man in

endlich viele disjunkte Gebiete $\Omega_1, \ldots, \Omega_L \in C^1_{pw}$ zerlegen, wobei jedes Ω_l ein Normalbereich ist bzgl. geeignet gewählter Achsen.

ii) Selbst für $\Omega \in C^1$ sind die Gebiete Ω_l in der Regel nur von der Klasse C^1_{pw}.

Beispiel 8.4.4. i) $\Omega = B_1(0)$ und

ii) $\Omega = B_2 \backslash B_1(0)$ sind in endlich viele disjunkte Normalbereiche zerlegbar.

Satz 8.4.1. (Green) *Sei $\Omega \subset Q \subset \mathbb{R}^2$ von der Klasse C^1_{pw}, und seien $g, h \in C^1(\overline{\Omega})$. Dann gilt*

$$\int_\Omega \Big(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y}\Big)\, d\mu = \int_{\partial\Omega} (g dx + h dy),$$

wobei der Rand von Ω so parametrisiert wird, dass Ω zur Linken liegt.

Beweis. **i)** Sei zunächst Ω ein Normalbereich, d.h. insbesondere

$$\Omega = \{(x,y);\ a \le x \le b,\ \varphi(x) \le y \le \psi(x)\}\ ,$$

wobei $a < b$, $\varphi \le \psi \in C^1_{pw}([a,b])$, und sei $h = 0$. Mit Satz 8.2.1 folgt

$$\begin{aligned}\int_\Omega \Big(-\frac{\partial g}{\partial y}\Big)\, d\mu &= \int_a^b \Big(-\int_{\varphi(x)}^{\psi(x)} \frac{\partial g}{\partial y}(x,y)\ dy\Big)\ dx \\ &= \int_a^b \big(g(x,\varphi(x)) - g(x,\psi(x))\big)\ dx.\end{aligned}$$

Parametrisiere $\partial\Omega = \gamma_1 + \gamma_2 - \gamma_3 - \gamma_4$, wobei

$$\begin{aligned}&\gamma_1(x) = (x, \varphi(x)),\ a \le x \le b,\ \gamma_2(x) = (b,y),\ \varphi(b) \le y \le \psi(b),\\ &\gamma_3(x) = (x, \psi(x)),\ a \le x \le b,\ \gamma_4(x) = (a,y),\ \varphi(a) \le y \le \psi(a).\end{aligned}$$

Sei λ die 1-Form $\lambda = (g, 0)$. Dann gilt

$$\begin{aligned}\int_{\partial\Omega} \lambda = \int_{\gamma_1+\gamma_2-\gamma_3-\gamma_4} g\ dx &= \int_{\gamma_1} g\ dx - \int_{\gamma_3} g\ dx \\ &= \int_a^b g(x,\varphi(x))\ dx - \int_a^b g(x,\psi(x))\ dx\ ;\end{aligned}$$

d.h.

$$\int_\Omega \Big(-\frac{\partial g}{\partial y}\Big)\, d\mu = \int_{\partial\Omega} g\ dx\ .$$

ii) Analog zu i) erhalten wir im Falle $g = 0$ für $h \in C^1(\overline{\Omega})$ nach Vertauschen von x und y unter Beachtung der Orientierung von $\partial\Omega$ die Identität

$$\int_\Omega \frac{\partial h}{\partial x}\, d\mu = \int_{\partial\Omega} h\ dy\ ;$$

zusammen also

$$\int_\Omega = \Big(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y}\Big)\, d\mu = \int_{\partial\Omega} (g dx + h dy)\ ,$$

wie gewünscht.

iii) Zerlege $\Omega = \Omega_1 \cup \cdots \cup \Omega_L$ in disjunkte Normalbereiche. Beachte, dass jede innere Randkomponente γ zu genau zwei Gebieten Ω_k, Ω_l gehört und als Teil von $\partial\Omega_k$ mit der entgegengesetzten Orientierung durchlaufen wird wie als Teil von $\partial\Omega_l$; die entsprechenden Wegintegrale heben einander also auf. Es folgt

$$\begin{aligned}\int_\Omega \Big(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y}\Big)\, d\mu &= \sum_{l=1}^{L} \int_{\Omega_l} \Big(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y}\Big)\, d\mu \\ &= \sum_{l=1}^{L} \int_{\partial\Omega_l} (g dx + h dy) = \int_{\partial\Omega} \big(g dx + h dy\big).\end{aligned}$$

□

Beispiel 8.4.5. i) Mit $g(x,y) = -y$, $h(x,y) = 0$ erhalten wir wie in Beispiel 8.4.1.ii) für ein beliebiges beschränktes Gebiet $\Omega \subset \mathbb{R}^2$ von der Klasse C^1_{pw} den Flächeninhalt

$$\mu(\Omega) = \int_\Omega \Big(-\frac{\partial g}{\partial y}\Big)\, d\mu = \int_{\partial\Omega} g\ dx = -\int_{\partial\Omega} y\ dx.$$

Analog ergibt die Wahl $g = 0$, $h(x,y) = x$ die Formel

$$\mu(\Omega) = \int_\Omega \frac{\partial h}{\partial x}\, d\mu = \int_{\partial\Omega} h\ dy = \int_{\partial\Omega} x\ dy,$$

und Kombination dieser Gleichungen liefert

$$\mu(\Omega) = \frac{1}{2} \int_{\partial\Omega} \big(x\ dy - y\ dx\big).$$

ii) Für $\Omega = B_1(0)$ mit der Parametrisierung $\gamma(t) = (\cos t, \sin t)^t$, $0 \le t \le 2\pi$, des Randes ergibt i) den Wert

$$\begin{aligned}\mu\big(B_1(0)\big) &= \frac{1}{2} \int_{\partial\Omega} \Big(x dy - y dx\Big) = \frac{1}{2} \int_\gamma (-y, x) \\ &= \frac{1}{2} \int_0^{2\pi} (-\sin t, \cos t) \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}\, dt = \pi.\end{aligned}$$

Statt als Koffizienten einer 1-Form können wir die Funktionen g und h in Satz 8.4.1 auch als die Komponenten eines Vektorfeldes $v = (g,h)^t \in C^1(\overline{\Omega}; \mathbb{R}^2)$ auffassen. Setzen wir noch in diesem Fall

$$rot\ v := \frac{\partial h}{\partial x} - \frac{\partial g}{\partial y},$$

so nimmt Satz 8.4.1 die folgende Gestalt an.

Satz 8.4.2. *Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C^1_{pw}, und sei $v \in C^1(\overline{\Omega}; \mathbb{R}^2)$. Dann gilt*

$$\int_\Omega rot\ v\ d\mu = \int_{\partial\Omega} v \cdot \vec{ds},$$

wobei $\partial\Omega$ so orientiert durchlaufen wird, dass Ω zur Linken liegt.

Satz 8.4.2 liefert insbesondere ein Kriterium für konservative Kraftfelder.

Definition 8.4.3. *Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C^1_{pw}; weiter sei Ω wegzusammenhängend. Dann heisst Ω* **einfach zusammenhängend** *, falls $\partial\Omega$ nur eine "Komponente" hat.*

Beispiel 8.4.6. i) $B_1(0)$ ist einfach zusammenhängend.

ii) $B_2(0)\backslash B_1(0)$ ist nicht einfach zusammenhängend.

Bemerkung 8.4.2. Eine andere Charakterisierung einfach zusammenhängender Mengen $\Omega \subset \mathbb{R}^2$ ist oft nützlich. Sei $\gamma \in C^1([0,1];\Omega)$ "geschlossen" mit $\gamma(0) = \gamma(1)$ und ohne Selbstschnitte; d.h., $\gamma(s) \neq \gamma(t)$ für alle $0 \leq s < t < 1$. Nach dem "Jordanschen Kurvensatz" berandet γ ein beschränktes Gebiet $\Omega_\gamma \subset \mathbb{R}^2$. Die Relation $\Omega_\gamma \subset \Omega$ gilt nun genau dann für alle derartigen Kurven, wenn Ω einfach zusammenhängend ist. Beispiel 8.4.6 illustriert diesen Zusammenhang auf einfache Weise.

Satz 8.4.3. (Poincaré) *Sei $\Omega \subset \mathbb{R}^2$ in C^1_{pw} beschränkt, zusammenhängend sowie einfach zusammenhängend, und sei $v \in C^1(\overline{\Omega};\mathbb{R}^2)$. Dann sind äquivalent*

i) *v ist konservativ,*

ii) *rot $v = 0$.*

Beweis. *i) $\Rightarrow$ ii)* Nach Satz 7.4.3 besitzt jedes konservative Vektorfeld $v \in C^1(\overline{\Omega};\mathbb{R}^2)$ ein Potential $f \in C^2(\overline{\Omega})$ mit $v = \nabla f$, und

$$rot\ \nabla f = \frac{\partial(\partial f/\partial y)}{\partial x} - \frac{\partial(\partial f/\partial x)}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} = 0$$

nach Satz 7.5.1.

ii) $\Rightarrow$ i) Sei $\gamma \in C^1([0,1];\Omega)$ ein geschlossener Weg in Ω ohne Selbstschnitte, Ω_γ das von γ berandete Gebiet. Nach Bemerkung 8.4.2 gilt $\Omega_\gamma \subset \Omega$, und die Annahme *rot* $v = 0$ zusammen mit Satz 8.4.2 ergibt

$$\int_\gamma v \cdot \vec{ds} = \int_{\Omega_\gamma} rot\ v\ d\mu = 0,$$

wie gewünscht. □

Beispiel 8.4.7. i) Sei $v(x,y) = \begin{pmatrix} y + \sin x \\ x + \cos y \end{pmatrix} \in C^1(\mathbb{R}^2;\mathbb{R}^2)$.

Jede Kugel $B_R(0) \subset \mathbb{R}^2$ ist einfach zusammenhängend, und

$$rot\ v = -\frac{\partial(y + \sin x)}{\partial y} + \frac{\partial(x + \cos y)}{\partial x} = 0.$$

Also ist v konwervativ nach Satz 8.4.3. (Die Funktion

$$f(x,y) = xy + \sin y - \cos x, \quad (x,y) \in \mathbb{R}^2,$$

ist ein Potential für v.)

ii) Sei $v(x,y) = \frac{1}{x^2+y^2} \cdot (-y,x)^t \in C^1(\mathbb{R}^2\backslash\{0\};\mathbb{R}^2)$. Es gilt

$$rot\ v = \frac{2}{x^2+y^2} - \frac{2(x^2+y^2)}{(x^2+y^2)^2} = 0,$$

jedoch ist $B_R(0)\backslash\{0\}$ nicht einfach zusammenhängend für $R > 0$.

In der Tat gilt für $\gamma(t) = (\cos t, \sin t)^t$, $0 \le t \le 2\pi$:

$$\int_\gamma v \cdot \vec{ds} = \int_0^{2\pi} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}\ dt = 2\pi \neq 0;$$

d.h. v ist **nicht** konservativ.

Auf der einfach zusammenhängenden Menge $\mathbb{R}^2\backslash\{(x,0);\ x \le 0\}$ besitzt jedoch v die **Argumentfunktion**

$$f(x,y) = \arctan\Big(\frac{y}{x}\Big)$$

als Potential.

Weiter gilt auf jedem einfach geschlossenen stückweise C^1-Weg $\tilde{\gamma}\colon [0,1] \to \mathbb{R}^2\backslash\{0\}$ mit $0 \notin \Omega_{\tilde{\gamma}}$ nach Satz 8.4.1

$$\int_{\tilde{\gamma}} v \cdot \vec{ds} = \int_{\Omega_{\tilde{\gamma}}} rot\ v\ d\mu = 0.$$

8.5 Substitutionsregel

Gibt es eine zur Substitutionsregel Satz 6.1.5 analoge Regel in $\mathbb{R}^n$? Wird insbesondere eine messbare Menge $\Omega \subset \mathbb{R}^n$ bei "Transformantion" mit einer Abbildung $\Phi \in C^1(\overline{\Omega};\mathbb{R}^n)$ wieder in eine messbare Menge $\Phi(\Omega)$ überführt?

Wir betrachten zunächst lineare Abbildungen.

Beispiel 8.5.1. i) Sei $Q \subset \mathbb{R}^2$ ein Quader, $R\colon \mathbb{R}^2 \to \mathbb{R}^2$ eine Rotation. Dann ist $R(Q)$ ein Normalbereich, und gemäss Satz 8.4.1 gilt

$$\mu(R(Q)) = \frac{1}{2}\int_{R\circ\gamma} x\ dy - y\ dx = \frac{1}{2}\int_{R\circ\gamma} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot \vec{ds}\ ,$$

wobei $\gamma \in C^1_{pw}([0,1];\mathbb{R}^2)$ eine mit der Orientierung verträgliche Parametrisierung von ∂Q ist. Identifizieren wir $(x,y)^t \in \mathbb{R}^2$ mit $x+iy \in \mathbb{C}$, so können wir den Vektor

$$(-y,x)^t = -y + ix = i(x+iy)$$

als den um 90° gedrehten Ortsvektor auffassen. Bezeichnen wir diese Drehung mit i und beachten wir die Vertauschungsrelation $R\circ i = i\circ R$ sowie die Beziehung $\frac{d}{dt}\big(R\circ\gamma)(t) = R\dot{\gamma}(t)$, so erhalten wir

$$\int_{R\circ\gamma} \begin{pmatrix} -y \\ x \end{pmatrix} \cdot \vec{ds} = \int_0^1 \big(i\circ R\circ\gamma\big)(t) \cdot \frac{d}{dt}\big(R\circ\gamma)(t)\big)\ dt = \int_0^1 \big(R\circ i\circ\gamma\big)(t) \cdot \big(R\dot{\gamma}(t)\big)\ dt\ .$$

Da das Skalarprodukt im $\mathbb{R}^2$ unter Drehungen invariant ist, folgt schliesslich

$$\mu(R(Q)) = \int_0^1 \big(R \circ i \circ \gamma\big)(t) \cdot \big(R\dot{\gamma}(t)\big)\ dt = \int_0^1 \big(i \circ \gamma\big)(t) \cdot \dot{\gamma}(t)\ dt = \mu(Q)\ .$$

ii) Analog gilt für einen Quader $Q \subset \mathbb{R}^3$ und eine beliebige Drehung R die Gleichheit $\mu(R(Q)) = \mu(Q)$, da man R als Produkt von Drehungen um eine Koordinatenachse schreiben kann; ebenso im $\mathbb{R}^n$, $n \geq 3$.

iii) Sei $A\colon \mathbb{R}^2 \to \mathbb{R}^2$ eine Scherstreckung mit Matrixdasrtellung $A = \begin{pmatrix} 1 & 0 \\ \lambda & \nu \end{pmatrix}$, und sei Q der Quader $Q = [a,b] \times [c.d]$. Falls $\nu > 0$ ist dann

$$AQ = \{(x,y);\ a \leq x \leq b,\ \lambda x + \nu c \leq y \leq \lambda x + \nu d\}$$

ein Normalbereich, und mit Satz 8.2.1 erhalten wir

$$\mu(AQ) = \int_a^b |\nu|\,|d-c|\ dx = |\nu|\,|b-a|\,|d-c| = |detA|\ \mu(Q).$$

Diese Formel gilt offenbar auch für beliebige $\nu \in \mathbb{R}$ und ebenso in höheren Dimensionen für Transformationen $A = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 0\ 1 & 0 \\ 0 & \dots & 0\ \lambda & \nu \end{pmatrix} : \mathbb{R}^n \to \mathbb{R}^n$, und wir erhalten die Identität

$$\mu(AQ) = |det\ A|\ \mu(Q).$$

für jeden n-Quader Q. Mit ii) gilt diese Formel auch für Drehungen. Weiter können wir jede Jordan-messbare Menge durch Elementarfiguren annähern und erhalten so die obige Beziehung für jede Jordan-messbare Menge Ω anstelle von Q.

iv) Schliesslich lässt sich jede $n \times n$-Matrix A mit der QR-Zerlegung darstellen als Produkt $A = RS_1 \circ \cdots \circ S_L$ einer Drehung R und Scherstreckungen $S_1, \dots, S_L$. Mit ii) und iii) erhalten wir dann für jede lineare Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^n$ und jede Jordan-messbare Menge $\Omega \subset \mathbb{R}^n$ die Beziehung

$$\mu(A\Omega) = \mu\big(RS_1 \circ \cdots \circ S_L(\Omega)\big) = |det\ R| \prod_{l=1}^{L} |det\ S_l|\,\mu(\Omega) = |det\ A|\ \mu(Q).$$

Ähnliches gilt auch für Transformationen mit $\Phi \in C^1(\overline{\Omega};\ \mathbb{R}^n)$. Da wir aber (im Unterschied zum Fall $n = 1$) für $n \geq 2$ im Massbegriff keine Rücksicht auf die Orientierung nehmen, müssen wir die Klasse der zulässigen Abbildungen einschränken.

Definition 8.5.1. *Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^1(U;\mathbb{R}^n)$. Die Abbildung Φ heisst ein* **Diffeomorphismus** *von U auf $\Phi(U) = V$, falls Φ injektiv ist und falls die Umkehrabbildung $\Psi = \Phi^{-1} \in C^1(V;\mathbb{R}^n)$.*

Bemerkung 8.5.1. Mit dem Umkehrsatz folgt, dass eine Abbildung $\Phi \in C^1(U;\mathbb{R}^n)$ genau dann ein Diffeomorphismus ist, wenn Φ injektiv ist mit

$$det(d\Phi(x_0)) \neq 0, \quad \forall x_0 \in U.$$

Beispiel 8.5.2. i) Eine lineare Abbildung $A\colon \mathbb{R}^n \to \mathbb{R}^n$ mit Matrixdarstellung A ist ein Diffeomorphismus genau dann, wenn $det\ A \neq 0$.

ii) Eine Abbildung $g \in C^1(]a,b[)$ mit $g' > 0$ ist gemäss Satz 5.2.2 ein Diffeomorphismus auf $g(]a,b[)$.

iii) Die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ aus Beispiel 7.6.1 mit

$$f(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = z^2,\ \forall (x,y) = z = x + iy \in \mathbb{R}^2.$$

liefert einen Diffeomorphismus $\Phi = f\big|_{\mathbb{R}^2_+} : \mathbb{R}^2_+ \to \mathbb{R}^2\backslash\{(x,0);\ x \leq 0\}$, wobei $\mathbb{R}^2_+ = \{(x,y);\ x > 0\}$.

Satz 8.5.1. (Transformationssatz) *Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^1(U;\mathbb{R}^n)$ ein Diffeomorphismus von U auf $V = \Phi(U) \subset \mathbb{R}^n$, und sei $\Omega \subset \overline{\Omega} \subset U$ beschränkt und Jordan messbar. Dann ist $\Phi(\Omega)$ Jordan messbar, und*

$$\mu(\Phi(\Omega)) = \int_\Omega |det(d\Phi(x))|\ d\mu(x).$$

Beweis. **i)** Wir zeigen zunächst, dass $\Phi(\Omega)$ Jordan messbar ist. Nach Bemerkung 8.3.1.ii) genügt es zu zeigen, dass man zu jedem $\epsilon > 0$ eine Elementarfigur $F_\epsilon \subset \mathbb{R}^n$ finden kann mit $\partial(\Phi(\Omega)) \subset F_\epsilon$ und $\mu(F_\epsilon) < \epsilon$.

Da Φ ein Diffeomorphismus ist, bildet Φ innere Punkte von Ω ab auf innere Punkte von $\Phi(\Omega)$, ebenso innere Punkte von $U\backslash\Omega$ auf innere Punkte des Komplements, und umgekehrt; also folgt

$$\partial(\Phi(\Omega)) = \Phi(\partial\Omega).$$

Da Ω Jordan messbar, gibt es zu jedem $\epsilon > 0$ eine Elementarfigur E_ϵ mit $\partial\Omega \subset E_\epsilon \subset \mathbb{R}^n$ und $\mu(E_\epsilon) < \epsilon$. Indem wir $\mathbb{R}^n$ mittels eines Gitters der Kantenlänge δ in Würfel zerlegen, können wir annehmen, dass E_ϵ die Vereinigung disjunkter derartiger Würfel W_l ist, $1 \leq l \leq L$, mit $\mu(E_\epsilon) = L\delta^n$, und weiter, dass $E_\epsilon \subset E_{\epsilon_0} \subset K = \overline{K} \subset U$ für alle $0 < \epsilon < \epsilon_0$, für ein $\epsilon_0 > 0$ und ein kompaktes $K \subset U$.

Nach Satz 4.2.3 existiert $C > 0$ mit

$$\sup_{x\in K} |d\Phi(x)| \leq C < \infty.$$

Für $1 \leq l \leq L$ folgt

$$\sup_{x,y\in W_l} |\Phi(x) - \Phi(y)| \leq \int_0^1 \Big| d\Phi\big(\underbrace{x + \vartheta(y-x)}_{\in W_l \subset K}\big)\Big| \underbrace{|x-y|}_{\leq \sqrt{n}\delta}\ d\vartheta \leq C\sqrt{n}\delta\ ;$$

d.h., $\Phi(W_l)$ ist enthalten in einem Würfel V_l der Kantenlänge $C_1\delta$ mit einer von ϵ unabhängigen Konstanten C_1. Somit gilt

$$\partial\big(\Phi(\Omega)\big) = \Phi(\partial\Omega) \subset \bigcup_{l=1}^{L} \Phi(W_l) \subset \bigcup_{l=1}^{L} V_l =: F_\epsilon$$

mit

$$\mu(F_\epsilon) \leq \sum_{l=1}^{L} \mu(V_l) \leq L C_1^n \delta^n = C_1^n \mu(E_\epsilon) < C_1^n \epsilon \; .$$

ii) Sei $x_0 \in \Omega$, und sei $A\colon \mathbb{R}^n \to \mathbb{R}^n$ die Abbildung

$$Ax = \Phi(x_0) + d\Phi(x_0)(x - x_0) \; .$$

Da Φ nach Annahme an der Stelle x_0 differenzierbar ist, gilt

$$\frac{\Phi(x) - A(x)}{|x - x_0|} \to 0 \;\; (x \to x_0).$$

Für einen Würfel W um x_0 der Kantenlänge $\delta > 0$ folgt

$$\frac{\sup_{x\in W} |\Phi(x) - A(x)|}{\delta} \to 0 \;\; (\delta \to 0) \; .$$

Da die Seiten des Parallelepipeds AW einen $(n-1)$-dimensionalen Inhalt proportional zu δ^{n-1} haben, können wir daher abschätzen

$$\frac{|\mu(\Phi(W)) - |det(d\Phi(x_0))|\,\mu(W)|}{\delta^n} = \frac{|\mu(\Phi(W)) - \mu(AW)|}{\delta^n} \to 0 \;\; (\delta \to 0) \; .$$

iii) Für jedes $\delta > 0$ zerlege $\mathbb{R}^n$ in Würfel der Kantenlänge δ mittels eines achsenparallen Gitters. Zu $\epsilon > 0$ wähle $\delta > 0$ so klein, dass die Elementarfigur $E_\epsilon = \bigcup_{l=1}^{L} W_l$, bestehend aus den Würfeln W des Gitters mit $W \cap \partial\Omega \neq \emptyset$, die Bedingungen

$$\partial\Omega \subset E_\epsilon \subset K, \;\; \mu(E_\epsilon) < \epsilon, \;\; \partial(\Phi(\Omega)) \subset \Phi(E_\epsilon) \subset F_\epsilon, \;\; \mu(F_\epsilon) < C\epsilon$$

gemäss i) erfüllt. Die Menge $\Omega_\epsilon = \Omega \backslash E_\epsilon$ ist dann ebenfalls Vereinigung von Würfeln W_l, $l = L+1, \ldots, M$, wobei $M \leq C\delta^{-n}$, und gemäss ii) gilt

$$\mu(\Phi(\Omega_\epsilon)) = \sum_{l=L+1}^{M} \mu(\Phi(W_l)) = \sum_{l=L+1}^{M} \Big(\big|det\big(d\Phi(x_l)\big)\big|\,\mu(W_l) + R_l \Big)$$

mit

$$R := \sum_{l=L+1}^{M} |R_l| \leq C\delta^{-n} \max_{L+1\leq l\leq M} |R_l| \to 0 \;\; (\delta \to 0).$$

Bei festem E_ϵ können wir δ durch $2^{-k}\delta$ mit beliebigem $k \in \mathbb{N}$ ersetzen. Im Limes $k \to \infty$ ergibt Satz 8.1.1 dann die Beziehung

$$\mu\big(\Phi(\Omega_\epsilon)\big) = \int_{\Omega_\epsilon} \big|det\big(d\Phi(x)\big)\big| \;\, d\mu,$$

und wir erhalten

$$\begin{aligned}\mu(\Phi(\Omega)) - \int_\Omega |det(d\Phi(x))|\ d\mu(x) \\ = \mu(\Phi(\Omega)) - \mu\big(\Phi(\Omega_\epsilon)\big) - \int_{\Omega\setminus\Omega_\epsilon} \big|det\big(d\Phi(x)\big)\big|\ .\end{aligned}$$

Da

$$\Phi(\Omega_\epsilon) \subset \Phi(\Omega) \subset \Phi(\Omega_\epsilon) \cup F_\epsilon\ ,$$

erhalten wir für die Fehlerterme einerseits die Abschätzung

$$0 < \mu\big(\Phi(\Omega)\big) - \mu\big(\Phi(\Omega_\epsilon)\big) \le \mu(F_\epsilon) < C\ \epsilon\ ;$$

andererseits gilt

$$0 < \int_{\Omega\setminus\Omega_\epsilon} \big|det\big(d\Phi(x)\big)\big|\ d\mu \le \int_{E_\epsilon} \underbrace{\big|det\big(d\Phi(x)\big)\big|}_{\le C}\ d\mu \le C\mu(E_\epsilon) \le C\epsilon\ .$$

Der Satz folgt nach Übergang zum Limes $\epsilon \downarrow 0$.

□

Beispiel 8.5.3. i) Sei $g \in C^1(]a,b[)$ mit $g' > 0$ ein Diffeomorphismus auf $g(]a,b[) =]c,d[$ gemäss Satz 5.2.2, und seien $a < x_0 < x_1 < b$, so dass $\Omega :=]x_0, x_1[\subset \overline{\Omega} \subset U =]a,b[$. Dann gilt

$$g(\Omega) =]g(x_0), g(x_1)[,$$

und

$$\mu(g(\Omega)) = g(x_1) - g(x_0) = \int_{x_0}^{x_1} g'(x)\ dx = \int_{x_0}^{x_1} |det(dg(x))|\ dx\ .$$

Im Limes $x_0 \downarrow a$ oder $x_1 \uparrow b$ können die auftretenden Terme divergieren. Wählen wir $g(x) = \log(x)\colon \mathbb{R}_+ \to \mathbb{R}$, so erhalten wir beispielsweise "$\mu(g(]0,1[)) = \infty$".

ii) Polarkoordinaten. Die Abbildung

$$\Phi(r,\theta) = \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix},\quad r > 0,\ 0 < \theta < 2\pi$$

erfüllt gemäss Beispiel 7.7.1.iii) die Gleichung

$$det\big(d\Phi(r,\theta)\big) = r.$$

Für $B_R(0) = \Phi(]0,R[\times]0,2\pi[) \cup \{(x,0);\ -R < x < 0\}$ folgt

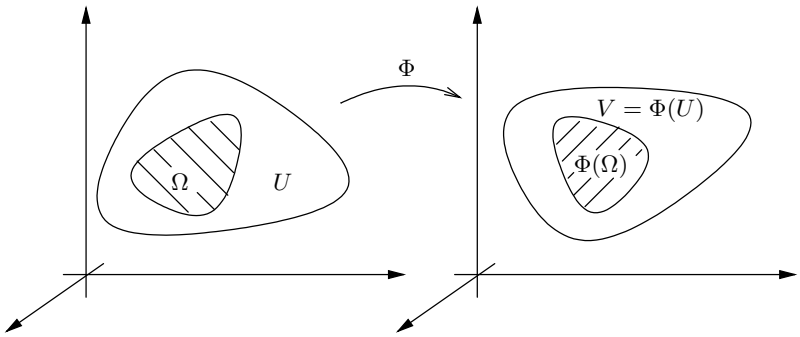
$$\mu(B_R(0)) = \lim_{\epsilon\downarrow 0} \int_{]\epsilon,R[\times]\epsilon,2\pi-\epsilon[} r\ d\mu(r,\theta) = \int_0^{2\pi} \Big(\underbrace{\int_0^R r\ dr}_{=R^2/2}\Big)\ d\theta = \pi R^2.$$

Analog zu Satz 8.5.1 erhalten wir auch eine zu Satz 6.1.5 analoge Regel für Integrale.

Satz 8.5.2. (Substitutionsregel) *Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^1(U;\mathbb{R}^n)$ ein Diffeomorphismus von U auf $V = \Phi(U) \subset \mathbb{R}^n$, $\Omega \subset \overline{\Omega} \subset U$ beschränkt und Jordan messbar, und sei $f\colon \Phi(\Omega) \to \mathbb{R}$ beschränkt und R-integrabel.*

Dann ist die Funktion $(f \circ \Phi) \cdot |det(d\Phi)| : \Omega \to \mathbb{R}$ ebenfalls R-integrabel, und es gilt

$$\int_{\Phi(\Omega)} f \; d\mu = \int_{\Omega} (f \circ \Phi) \cdot |det(d\Phi)| \; d\mu.$$



Beweis. **i)** Sei f eine Treppenfunktion (TF)

$$f = \sum_{j=1}^{J} c_j \chi_{Q_j}$$

mit disjunkten Quadern $Q_j \subset \mathbb{R}^n$, wobei

$$\Phi(\Omega) \subset \bigcup_{j=1}^{J} Q_j \subset \bigcup_{j=1}^{J} \overline{Q_j} \subset V.$$

Gemäss Satz 8.5.1, angewandt auf $\Psi = \Phi^{-1}$, ist für jedes j die Menge

$$\Omega_j = \Phi^{-1}(Q_j) \cap \Omega$$

Jordan-messbar. Weiter sind die Mengen $\Omega_1, \ldots, \Omega_L$ disjunkt und überdecken Ω. Mit Satz 8.5.1 erhalten wir nun

$$\begin{aligned} \int_{\Phi(\Omega)} f \; d\mu &= \sum_{j=1}^{J} c_j \mu\big(\underbrace{Q_j \cap \Phi(\Omega)}_{=\Phi(\Omega_j)}\big) \\ &= \sum_{j=1}^{J} c_j \int_{\Omega_j} |det(d\Phi)| \; d\mu = \int_{\Omega} (f \circ \Phi)\, |det(d\Phi)| \; d\mu. \end{aligned}$$

ii) Für eine beliebige beschränkte, R-integrable Funktion $f\colon \Phi(\Omega) \to \mathbb{R}$ folgt

nun mit wenig Mühe

$$\begin{aligned}\int_{\Phi(\Omega)} f\ d\mu &= \sup\Big\{\int_{\Phi(\Omega)} e\ d\mu;\ e\le f,\ e \text{ TF}\Big\}\\ &\overset{i)}{=} \sup\Big\{\int_{\Omega}(e\circ\Phi)\,|det(d\Phi)|\ d\mu;\ e\le f,\ e \text{ TF}\Big\}\\ &\le \underline{\int_{\Omega}}(f\circ\Phi)\,|det(d\Phi)|\ d\mu \le \overline{\int_{\Omega}}(f\circ\Phi)\,|det(d\Phi)|\ d\mu\\ &\le \inf\Big\{\int_{\Omega}(g\circ\Phi)\,|det(d\Phi)|\ d\mu;\ f\le g,\ g \text{ TF}\Big\} = \cdots = \int_{\Phi(\Omega)} f\ d\mu\ .\end{aligned}$$

Die Ausführung der Details sei dem Leser überlassen.

□

Beispiel 8.5.4. Die Funktion $f(x) = e^{-x^2}\colon \mathbb{R}\to\mathbb{R}$ hat keine elementar berechenbare Stammfunktion. Mit Satz 8.5.2 können wir jedoch das Integral $\int_{-\infty}^{\infty} e^{-x^2}\ dx$ explizit bestimmen.

Nach Fubini (Satz 8.2.1) gilt

$$\begin{aligned}\Big(\int_{-\infty}^{\infty} e^{-x^2}\ dx\Big)^2 &= \Big(\int_{-\infty}^{\infty} e^{-x^2}\ dx\Big)\Big(\int_{-\infty}^{\infty} e^{-y^2}\ dy\Big)\\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\ dx\ dy = \int_{\mathbb{R}^2} e^{-r^2}\ d\mu,\end{aligned}$$

wobei $r^2 = x^2+y^2$. Gemäss Satz 8.5.2 erhalten wir nach Einführung von Polarkoordinaten entsprechend Beispiel 8.5.3.ii) für den letzten Term den Ausdruck

$$\int_{\mathbb{R}^2} e^{-r^2}\ d\mu = \int_0^{2\pi}\underbrace{\Big(\int_0^{\infty} re^{-r^2}\ dr\Big)}_{=1/2}\ d\theta = \pi\ ;$$

also gilt

$$\int_{-\infty}^{\infty} e^{-x^2}\ dx = \sqrt{\pi}.$$

8.6 Oberflächenmass und Fluss-Integral

Die Sätze 8.5.1 und 8.5.2 zeigen einen natürlichen Weg auf zur Definition des Inhalts von regulären Flächenstücken im $S\subset\mathbb{R}^3$ und Oberflächenintegralen. Dazu benötigen wir lediglich ein Konzept, welches den Begriff des Diffeomorphismus in geeigneter Weise auf Abbildungen $\Phi\colon \mathbb{R}^2\supset\Omega\to S\subset\mathbb{R}^3$ verallgemeinert (und analog in höheren Dimensionen).

Definition 8.6.1. *Sei $U\subset\mathbb{R}^2$ offen, $\Phi\in C^1(U;\mathbb{R}^3)$ injektiv. Falls $d\Phi(x)$ für alle $x\in U$ den (maximalen) Rang 2 hat, so heisst Φ eine* **lokale Immersion**.

Beispiel 8.6.1. i) Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U;\mathbb{R}^2)$ Diffeomorphismus auf $V = \Phi(U) \subset \mathbb{R}^2$. Fassen wir $\mathbb{R}^2$ auf als den Unterraum $\mathbb{R}^2 \times \{0\} \subset \mathbb{R}^3$, so ist $\Phi\colon U \to \mathbb{R}^3$ lokale Immersion.

ii) Sei $\Phi \in C^1(\mathbb{R}^2;\mathbb{R}^3)$ gegeben mit

$$\Phi(x,y) = \frac{1}{1+x^2+y^2}\begin{pmatrix} 2x \\ 2y \\ 1-x^2-y^2 \end{pmatrix},\ (x,y) \in \mathbb{R}^2.$$

Da

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

gilt $|\Phi(x,y)| \equiv 1$; d.h.,

$$\Phi\colon \mathbb{R}^2 \to S^2 = \{(x,y,z);\ x^2+y^2+z^2 = 1\}\ .$$

Behauptung Φ ist Immersion.

Beweis. Betrachte für festes $(x,y) \in \mathbb{R}^2$ die Matrix

$$A = \Big(\Phi, \frac{\partial\Phi}{\partial x}, \frac{\partial\Phi}{\partial y}\Big)(x,y).$$

Wegen Antisymmetrie der Determinante gilt

$$det(A) = det\begin{pmatrix} 2x & 2 & 0 \\ 2y & 0 & 2 \\ 1-x^2-y^2 & -2x & -2y \end{pmatrix} \cdot \frac{1}{(1+x^2+y^2)^3}\ .$$

Es folgt

$$det(A) = \frac{4(1-x^2-y^2)+8x^2+8y^2}{(1+x^2+y^2)^3} = \frac{4}{(1+x^2+y^2)^2} \neq 0\ ,$$

und $Rang\big(d\Phi(x,y)\big) = 2,\ \forall (x,y) \in \mathbb{R}^2$.

Zudem ist Φ injektiv. Die Kreise $S^1_r = \{(x,y);\ x^2+y^2 = r^2\}$, $r > 0$, werden nämlich bijektiv auf die Breitenkreise mit der konstanten z-Komponente $z = (1-r^2)/(1+r^2)$ abgebildet, welche streng monoton mit r fällt. □

Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U;\mathbb{R}^3)$ lokale Immersion. Weiter sei $\Omega \subset \overline{\Omega} \subset U$ beschränkt und Jordan messbar, $S = \Phi(\Omega) \subset \mathbb{R}^3$ das durch Φ parametrisierte **Flächenstück**.

Zur besseren Unterscheidung wählen wir im folgenden meist die Bezeichnungen $(u,v) \in U \subset \mathbb{R}^2$, bzw. $(x,y,z) \in \mathbb{R}^3$ für die Koordinaten im Urbild- und Zielraum.

Definition 8.6.2. *Der 2-dimensionale* **Flächeninhalt** *von S (bzgl. Φ) ist*

$$\mu_2(\Omega) := \int_S do := \int_\Omega |\Phi_u \times \Phi_v|\ d\mu(u,v),$$

wobei $\Phi_u = \frac{\partial \Phi}{\partial u}$*, etc, und mit*

$$do := |\Phi_u \times \Phi_v| \; d\mu(u,v),$$

dem **skalaren Flächeninhalt** *bzgl.* Φ.

Bemerkung 8.6.1. i) $|\Phi_u \times \Phi_v|$ ist der Inhalt des von den Vektoren Φ_u und Φ_v aufgespannten Parallelogramms.

ii) Falls $\Phi(U) \subset \mathbb{R}^2$ (oder falls wir zu gegebenem $w = (u,v) \in U$ Koordinaten (x,y,z) für $\mathbb{R}^3$ wählen, so dass $\Phi_u(w)$, $\Phi_v(w) \in \mathbb{R}^2 \times \{0\}$), so gilt

$$|\Phi_u \times \Phi_v|\,(w) = |det(d\Phi(w))|\,;$$

die Definition 8.6.2 ist somit konsistent mit Satz 8.5.1.

iii) Allgemein gilt

$$|\Phi_u \times \Phi_v| = \sqrt{det(g)},$$

wobei

$$g = d\Phi^t \underbrace{d\Phi}_{(\Phi_u, \Phi_v)} = \begin{pmatrix} |\Phi_u|^2 & \Phi_u \cdot \Phi_v \\ \Phi_u \cdot \Phi_v & |\Phi_u|^2 \end{pmatrix}$$

die von Φ induzierte **Gramsche Matrix** (Metrik) bezeichnet.

Beweis: In geeigneten Koordinaten gilt $d\Phi(w)\colon \mathbb{R}^2 \to \mathbb{R}^2 \times \{0\} \cong \mathbb{R}^2$.

$$\begin{aligned} |\Phi_u \times \Phi_v|\,(w) &= |det(d\Phi(w))| = \big|det(d\Phi^t(w))\big| \\ &= \sqrt{det\big(d\Phi^t(w)\big) \cdot det\big(d\Phi(w)\big)} = \sqrt{det(g)}. \end{aligned}$$

Bemerkung 8.6.2. Aus Satz 8.5.2 folgt mit Bemerkung 8.6.1.iii), dass der Flächeninhalt $\mu_2(S)$ von der Parametrisierung unabhängig ist.

Beispiel 8.6.2. Sei $\Phi \in C^1(\mathbb{R}^2; S^2)$ mit

$$\Phi(u,v) = \frac{1}{1+u^2+v^2} \begin{pmatrix} 2u \\ 2v \\ 1-u^2-v^2 \end{pmatrix}, \quad (u,v) \in \mathbb{R}^2,$$

wie in Beispiel 8.6.1.ii) mit

$$|\Phi_u \times \Phi_v| = \big|det\big(\Phi, \Phi_u, \Phi_v\big)\big| \equiv \frac{4}{(1+u^2+v^2)^2}.$$

Es folgt

$$\begin{aligned} \mu_2(S^2) &= \int_{\mathbb{R}^2} |\Phi_u \times \Phi_v| \; d\mu(u,v) = \int_{\mathbb{R}^2} \frac{4}{(1+u^2+v^2)^2} \; d\mu(u,v) \\ &= \int_0^{2\pi} \Big(\underbrace{\int_0^\infty \frac{4r \; dr}{(1+r^2)^2}}_{=2\int_0^\infty \frac{ds}{(1+s)^2} = -\frac{2}{1+s}\Big|_{s=0}^\infty = 2} \Big) \, d\theta = 4\pi. \end{aligned}$$

Wir können nun auch stetige Funktionen über reguläre Flächenstücke integrieren.

Definition 8.6.3. *Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U;\mathbb{R}^3)$ lokale Immersion, $\Omega \subset \overline{\Omega} \subset U$ beschränkt und Jordan-messbar, und sei $S = \Phi(\Omega)$ das zugehörige Flächenstück im $\mathbb{R}^3$, $f\colon \overline{S} \to \mathbb{R}$ stetig. Dann ist*

$$\int_S f \; do := \int_\Omega (f \circ \Phi) \; |\Phi_u \times \Phi_v| \; d\mu(u,v)$$

wohldefiniert (unabhängig von Φ).

Beispiel 8.6.3. Für $\Phi, S = S^2$ wie in Beispiel 8.6.2 und $f(x,y,z) = z^2$ erhalten wir

$$\begin{aligned}
\int_{S^2} f \; d\mu &= \int_{\mathbb{R}^2} (f \circ \Phi) \cdot |\Phi_u \times \Phi_v| \; d\mu \\
&= \int_{\mathbb{R}^2} \underbrace{\Big(\frac{1-u^2-v^2}{1+u^2+v^2}\Big)^2}_{=\left(\frac{2}{1+u^2+v^2}-1\right)^2=\left(\frac{2}{1+u^2+v^2}\right)^2-\frac{4}{1+u^2+v^2}+1} \cdot \frac{4}{(1+u^2+v^2)^2} \; d\mu \\
&= \int_{\mathbb{R}^2} \Big(\frac{16}{(1+u^2+v^2)^4} - \frac{16}{(1+u^2+v^2)^3} + \frac{4}{(1+u^2+v^2)^2} \Big) \; d\mu \\
&= \int_0^{2\pi} \int_0^\infty \Big(\frac{8\; ds}{(1+s)^4} - \frac{8\; ds}{(1+s)^3} + \frac{2\; ds}{(1+s)^2} \Big) \; d\theta = \frac{4\pi}{3} \; .
\end{aligned}$$

Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U;\mathbb{R}^3)$ lokale Immersion, $\Omega, S = \Phi(\Omega)$ wie oben, und sei $w \in \Omega$. Da $d\Phi(w)$ den Rang 2 hat, spannen die Vektoren $\Phi_u(w)$ und $\Phi_v(w)$ den Tangentialraum auf an S im Punkt $p = \Phi(w)$, d.h.

$$T_pS = \{d\Phi(w)\zeta; \; \zeta \in T_w\mathbb{R}^2 \cong \mathbb{R}^2\} = span\{\Phi_u(w), \Phi_v(w)\}.$$

Durch Auswahl eines **Normalenvektors** $n = n(p) \in \mathbb{R}^3$ mit $n(p) \perp T_pS$ können wir T_pS **orientieren**, d.h. "oben" und "unten" definieren. Offenbar gibt es genau zwei Möglichkeiten, am Punkt p einen auf Länge 1 normierten Normalenvektor festzulegen; mittels der Parametrisierung Φ erhalten wir insbesondere

$$n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$$

als eine kanonische Wahl.

Für ein (in einer Umgebung W von S erklärtes) Vektorfeld $K = (P,Q,R)^t \in C^1(W;\mathbb{R}^3)$ deuten wir die Normalkomponente $(K \cdot n)(p)$ als **Flussdichte** von K durch S am Punkt p.

Definition 8.6.4. *Das Integral*

$$\int_S K \cdot n \; do = \int_\Omega (K \circ \Phi) \cdot \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|} \, |\Phi_u \times \Phi_v| \; d\mu(u,v)$$

heisst **Fluss** *des Vektorfeldes K durch die mit $n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$ orientierte Fläche*

S; dabei heisst

$$n\ do = \Phi_u \times \Phi_v\ d\mu(u,v)$$

das durch n **orientierte Flächenelement** *auf S.*

Beispiel 8.6.4. i) Sei $\Phi \in C^1(\mathbb{R}^2; S^2)$ wie in Beispiel 8.6.1.ii), $S = \Phi(\mathbb{R}^2) = S^2$, und sei K das Ortsvektorfeld mit $K(p) = p$, $p \in S^2$. Für die duch Φ definierte Orientierung $n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$ gilt offenbar $n(p) = p$; also

$$\int_{S^2} \underbrace{K \cdot n}_{\equiv 1}\ do = \int_{S^2} do = \mu(S^2) = 4\pi.$$

ii) Mit Φ, $S = S^2$ mit $n(p) = p$ wie in i) und $K(x,y,z) = (0,0,z)^t$ erhalten wir

$$\int_{S^2} K \cdot n\ do = \int_{S^2} z^2\ do = \frac{4\pi}{3}\ ;$$

vgl. Beispiel 8.6.3.

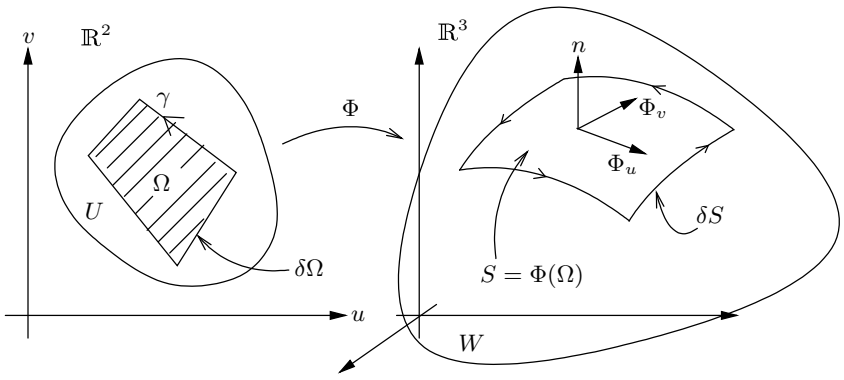
8.7 Der Satz von Stokes im $\mathbb{R}^3$

Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U; \mathbb{R}^2)$ lokale Immersion, $\Omega \subset \overline{\Omega} \subset U$ beschränkt und von der Klasse C^1_{pw}, und sei $S = \Phi(\Omega)$ das durch Φ parametriserte Flächenstück.

Sei $\gamma \in C^1_{pw}([0,1]; \mathbb{R}^2)$ eine Parametrisierung von $\partial\Omega$, und sei γ positiv orientiert in dem Sinne, dass Ω beim Durchlaufen von γ zur Linken liegt. Sei weiter $\Gamma := \Phi \circ \gamma \in C^1_{pw}([0,1]; \mathbb{R}^3)$ die von γ induzierte Parametrisierung des (intrinsischen) "Randes" $\partial S := \Phi(\partial\Omega)$ von S.

Mit der durch $n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$ definierten Orientierung von S ist dann Γ ebenfalls positiv orientiert: Wenn wir "auf" Γ die Fläche S umfahren, liegt S stets zur Linken.

Sei $K \in C^1(W; \mathbb{R}^3)$ in einer Umgebung W von S in $\mathbb{R}^3$ erklärt, $K = (P,Q,R)^t$.



Definition 8.7.1. *Das Vektorfeld*

$$\vec{rot}\ K = \begin{pmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{pmatrix} = \nabla \times K = det \begin{pmatrix} \partial_x & P & e_1 \\ \partial_y & Q & e_2 \\ \partial_z & R & e_3 \end{pmatrix}$$

heisst **Rotation** *("Wirbelstärke", engl. curl) von K.*

Satz 8.7.1. (Stokes) *Sind S und ∂S wie oben orientiert, so gilt*

$$\int_S \vec{rot}\, K \cdot n\ do = \int_{\partial S} K \cdot \vec{ds}.$$

Bemerkung 8.7.1. i) Das Integral $\int_{\partial S} K \cdot \vec{ds}$ heisst auch "Zirkulation" des Vektorfeldes K längs ∂S.

ii) Für $\Phi = id$, $S = \Omega \subset \mathbb{R}^2$, $n = e_3$ gilt

$$\vec{rot}\ K \cdot n = Q_x - P_y$$

sowie

$$\int_{\partial S} K \cdot \vec{ds} = \int_{\partial \Omega} P\ dx + Q\ dy,$$

und wir erhalten Satz 8.4.1.

Beispiel 8.7.1. Sei $K \in C^1(\mathbb{R}^3; \mathbb{R}^3)$ mit

$$K(x,y,z) = \begin{pmatrix} x(y^2+z^2) \\ y(x^2+z^2) \\ z(x^2+y^2) \end{pmatrix},$$

E das Ellipsoid

$$E = \{(x,y,z);\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}$$

mit Halbachsen $0 < a, b, c \in \mathbb{R}$, orientiert durch die äussere Normale, S das Flächenstück

$$S = \{(x,y,z) \in E;\ \frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} \leq 1,\ z > 0\}$$

mit $0 < \alpha < a$, $0 < \beta < b$, und $\Gamma = \partial S$ der (orientierte) Rand von S.

Beachte

$$\vec{rot}\ K = \begin{pmatrix} \cdots \\ \cdots \\ \frac{\partial(y(x^2+z^2))}{\partial x} - \frac{\partial(x(y^2+z^2))}{\partial y} \end{pmatrix} = 0;$$

also

$$\int_\Gamma K \cdot \vec{ds} = \int_S \vec{rot}\ K \cdot n\ do = 0.$$

Auf direktem Wege ist die Berechnung der Zirkulation offenbar viel aufwendiger.

Beweis von Satz 8.7.1. Wir führen Satz 8.7.1 zurück auf Satz 8.4.1. Setze dazu

$$(K \circ \Phi) \cdot d\Phi =: \lambda = g\ du + h\ dv$$

mit $g = (K \circ \Phi) \cdot \Phi_u$, $h = (K \circ \Phi) \cdot \Phi_v$, so dass

$$\begin{aligned}\int_{\partial S} K \cdot \vec{ds} &= \int_0^1 \underbrace{(K \circ \Gamma)(t)}_{=(K\circ\Phi)(\gamma(t))} \cdot \underbrace{\frac{d\Gamma}{dt}(t)}_{=d\Phi(\gamma(t))\dot\gamma(t)}\ dt \\ &= \int_0^1 \lambda(\gamma(t))\dot\gamma(t)\ dt = \int_{\partial\Omega} g\ du + h\ dv.\end{aligned}$$

Behauptung Es gilt

$$\big(\vec{rot}\ K \circ \Phi\big) \cdot \Phi_u \times \Phi_v = h_u - g_v.$$

Beweis: Die rechte Seite ergibt

$$h_u - g_v = (K \circ \Phi)_u \cdot \Phi_v - (K \circ \Phi)_v \cdot \Phi_u,$$

da sich die beiden übrigen Terme aufheben. Die weitere Rechnung können wir mit der folgenden Überlegung stark vereinfachen. Offenbar sind alle Ausdrücke unabhängig von der Wahl der Koordinaten im $\mathbb{R}^3$. Für festes $w \in \Omega$, $p = \Phi(w) \in S$ dürfen wir daher annehmen, dass

$$\Phi_u(w) = a\ e_1, \quad \Phi_v(w) = b\ e_1 + c\ e_2$$

mit $a, c > 0$, $b \in \mathbb{R}$.

Wir erhalten

$$\begin{aligned}(h_u - g_v)(w) &= aK_x(p) \cdot (b\ e_1 + c\ e_2) - (bK_x(p) + cK_y(p)) \cdot (ae_1) \\ &= \big(abP_x + acQ_x - abP_x - acP_y\big)(p) = ac(Q_x - P_y)(p).\end{aligned}$$

Die linke Seite ergibt andererseits, mit $(\Phi_u \times \Phi_v)(w) = ac\ e_3$,

$$\Big(\big(\vec{rot}\ K \circ \Phi\big) \cdot \Phi_u \times \Phi_v\Big)(w) = \vec{rot}\ K(p) \cdot ac\ e_3 = ac\ (Q_x - P_y)(p),$$

und die Behauptung folgt. □

Definition 8.7.2. *Sei $W \subset \mathbb{R}^3$ offen. W heisst* **einfach zusammenhängend** *(1-zshg), falls jede einfach geschlossene Kurve $\Gamma \in C^1_{pw}([0,1]; W)$ in W "zusammenziehbar" ist; d.h., falls jede derartige Kurve Γ die einzige Randkomponente eines orientierten Flächenstücks $S \subset W$ ist, welches darstellbar ist als Vereinigung von endlich vielen immersierten Flächenstücken $S_j = \Phi_j(\Omega_j)$, $1 \le j \le J$, mit kompatiblen Orientierungen.*

Beispiel 8.7.2. i) Der Ball $B_1(0; \mathbb{R}^3)$ ist 1-zshg,

ii) Die Kugelschale $B_2(0) \backslash B_1(0) \subset \mathbb{R}^3$ ist 1-zshg,

iii) $B_2(0; \mathbb{R}^3) \backslash \big(B_1(0; \mathbb{R}^2) \times \mathbb{R}\big)$ ist nicht 1-zshg.

iv) Das Möbiusband hat nur eine einzige Randkomponente, ist aber kein orientiertes Flächenstück.

Analog zu Satz 8.4.3 erhalten wir:

Satz 8.7.2. (Poincaré) *Sei $W \subset \mathbb{R}^3$ 1-zshg, $K \in C^1(W;\mathbb{R}^3)$. Es sind äquivalent:*

i) *K ist konservativ;*

ii) $\vec{rot}\, K = 0$.

Beweis. *i) $\Rightarrow$ ii)* Sei K konservativ. Nach Satz 7.4.3 gibt es $f \in C^2(W)$ mit $K = \nabla f$; also

$$\vec{rot}\, K = \begin{pmatrix} \cdots \\ \cdots \\ \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \end{pmatrix} = 0.$$

ii) $\Rightarrow$ i) Sei $\Gamma \subset W$ einfach geschlossen, $S \subset W$ 1-zshg mit $\partial S = \Gamma$. Falls $S = \Phi(\Omega)$ für eine injektive Immersion $\Phi \in C^1(U;\mathbb{R}^3)$ mit einem einfach zusammenhängenden $\Omega \subset \overline{\Omega} \subset U \subset \mathbb{R}^2$, so gilt nach Satz 8.7.1

$$\int_\Gamma K \cdot \vec{ds} = \int_S \vec{rot}\, K \cdot n \; do = 0.$$

Falls S nicht in dieser Weise darstellbar ist als $S = \Phi(\Omega)$, so zerlege S disjunkt in endlich viele derartige Teile S_j, $1 \le j \le J$, mit $S = \bigcup_{j=1}^{J} S_j$.

Da die S_j nach Annahme kompatibel orientiert sind, sind gemeinsame Randkurven von aneinander angrenzenden Flächenstücken S_j und S_k entgegengesetzt orientiert. Es folgt

$$\int_\Gamma K \cdot \vec{ds} = \int_{\partial S} K \cdot \vec{ds} = \sum_{j=1}^{J} \int_{\partial S_j} K \cdot \vec{ds} = \sum_{j=1}^{J} \int_{S_j} \vec{rot}\, K \cdot n \; do = 0;$$

d.h., K ist konservativ. □

Beispiel 8.7.3. Sei K wie in Beispiel 8.7.1. Nach Satz 8.7.2 ist K konservativ. In der Tat ist

$$f(x,y,z) = \frac{x^2y^2 + x^2z^2 + y^2z^2}{2}$$

eine Potentialfunktion.

Bemerkung 8.7.2. Ein Möbiusband S lässt sich so in $W = \mathbb{R}^3 \backslash \big(\{(0,0)\} \times \mathbb{R}\big)$ einbetten, dass die Projektion von ∂S auf $\mathbb{R}^2 \times \{0\}$ nach radialer Projektion $w \mapsto w/|w|$ dem doppelt durchlaufenen Einheitskreis entspricht.

Für $K \in C^1(W;\mathbb{R}^3)$ mit

$$K(x,y,z) = \frac{1}{x^2+y^2} \begin{pmatrix} -y \\ x \\ 0 \end{pmatrix} = \begin{pmatrix} P \\ Q \\ R \end{pmatrix}$$

gilt

$$\vec{rot}\, K = \begin{pmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} ,$$

jedoch erhalten wir

$$\int_{\partial S} K \cdot \vec{ds} = 4\pi;$$

vgl. Beispiel 8.4.7.ii).

Beachte, dass ∂S nur **eine** Zusammenhangskomponente hat; jedoch ist S nicht orientierbar. Tatsächlich ist S **nicht** einfach zusammenhängend.

Das Kriterium für einfach zusammenhängende Mengen $\Omega \subset \mathbb{R}^2$ nach Definition 8.4.3 ist also nur auf beschränkte Mengen $\Omega \subset \mathbb{R}^2$ anwendbar (oder allgemeiner auf Teilmengen von orientierbaren 2-dimensionalen Untermannigfaltigkeiten).

8.8 Der Satz von Gauss

Sei $W \subset \mathbb{R}^3$ offen, $K \in C^1(W; \mathbb{R}^3)$, $K = (K^i)_{1 \le i \le 3}$, $x = (x^i)_{1 \le i \le 3} \in W$.

Definition 8.8.1. *Der Ausdruck*

$$div(K) = \sum_{i=1}^{3} \frac{\partial K^i}{\partial x^i}$$

heisst **Divergenz** *(Quellstärke) von K.*

Satz 8.8.1. *(Gauss) Sei $\Omega \subset \overline{\Omega} \subset W$ beschränkt und von der Klasse C^1_{pw}. Dann gilt*

$$\int_\Omega div(K)\ d\mu = \int_{\partial\Omega} K \cdot n\ do,$$

wobei n die äussere Normale ist.

Bemerkung 8.8.1. Nach Satz 8.8.1 ist also der Fluss des Vektorfeldes K durch $\partial\Omega$ gleich dem Integral über die Quellstärke von K.

Beweis. **i)** Sei zunächst Ω ein Normalbereich bzgl. aller Achsen $x^1, \dots, x^3$. Vereinfachend gelte mit einem Quader $Q \subset \mathbb{R}^2$, $0 \le \psi \in C^1(\overline{Q})$:

$$\Omega = \{x = (x^i);\ (x^1, x^2) \in Q,\ 0 \le x^3 \le \psi(x^1, x^2)\}.$$

Es folgt

$$\begin{aligned} \int_\Omega \frac{\partial K^3}{\partial x^3}\ d\mu_3 &= \int_Q \Big(\int_0^{\psi(x^1,x^1)} \frac{\partial K^3}{\partial x^3}\ dx^3 \Big) d\mu_2(x^1, x^2) \\ &= \int_Q \big(K^3(x^1, x^2, \psi(x^1, x^2)) - K^3(x^1, x^2, 0) \big)\ d\mu_2(x^1, x^2). \end{aligned}$$

Sei $\Phi(x^1, x^2) = (x^1, x^2, \psi(x^1, x^2)) \in C^1(Q; \mathbb{R}^3)$ die von ψ induzierte Parametrisierung des "oberen" Teils $S = \mathcal{G}(\psi) = \Phi(Q)$ von $\partial\Omega$ mit

$$d\Phi = (\Phi_{x^1}, \Phi_{x^2}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \partial\psi/\partial x^1 & \partial\psi/\partial x^2 \end{pmatrix}.$$

Es folgt

$$\Phi_{x^1} \times \Phi_{x^2} = \begin{pmatrix} -\partial\psi/\partial x^1 \\ -\partial\psi/\partial x^2 \\ 1 \end{pmatrix}, \tag{8.8.1}$$

und wir erhalten

$$\begin{aligned} \int_S \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \cdot n \; do &= \int_Q \begin{pmatrix} 0 \\ 0 \\ K^3 \circ \Phi \end{pmatrix} \cdot \Phi_{x^1} \times \Phi_{x^2} \; d\mu_2(x^1, x^2) \\ &= \int_Q K^3\big(x^1, x^2, \psi(x^1, x^2)\big) \; d\mu_2(x^1, x^2). \end{aligned}$$

Analog gilt am "unteren" Teil $Q \times \{0\}$ von $\partial\Omega$ $n = -e_3$, also

$$\int_{Q\times\{0\}} \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \cdot n \; do = -\int_Q K^3(x^1, x^2, 0) \; d\mu_2(x^1, x^2).$$

Entlang der "vertikalen" Teile von $\partial\Omega$ gilt $n^3 = 0$; diese liefern also keinen Beitrag zum Fluss des Vektorfeldes $(0, 0, K^3)^t \in C^1(\overline{\Omega}; \mathbb{R}^3)$. Für dieses Vektorfeld gilt somit die Identität

$$\int_\Omega div \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \; d\mu = \int_{\partial\Omega} \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \cdot n \; do;$$

analog für die übrigen Komponenten.

ii) Zerlege ein beliebiges beschränktes Gebiet $\Omega \subset \overline{\Omega} \subset W$ der Klasse C^1_{pw} gemäss Bemerkung 8.4.1 in endlich viele disjunkte Normalbereiche $\Omega_1, \ldots, \Omega_L$. Entlang gemeinsamer Begrenzungsflächen von Ω_k und Ω_l heben sich die Flüsse auf. Es folgt

$$\begin{aligned} \int_\Omega div(K) \; d\mu &= \sum_{l=1}^L \int_{\Omega_l} div(K) \; d\mu \\ &= \sum_{l=1}^L \int_{\partial\Omega_l} K \cdot n_l \; do = \int_{\partial\Omega} K \cdot n \; do, \end{aligned}$$

wie gewünscht.

□

Man kann Satz 8.8.1 insbesondere zur Volumenberechnung verwenden.

Beispiel 8.8.1. i) Sei $\Omega = B_1(0;\mathbb{R}^3)$. Für $K(x) = x$ mit $div(K) = 3$ folgt mit Satz 8.8.1

$$3\mu_3\big(B_1(0;\mathbb{R}^3)\big) = \int_{B_1(0;\mathbb{R}^3)} div\ x\ dx = \int_{S^2} \underbrace{x \cdot n(x)}_{\equiv 1}\ do = \mu_2(S^2) = 4\pi;$$

also

$$\mu_3\big(B_1(0;\mathbb{R}^3)\big) = \frac{4\pi}{3}.$$

ii) Sei Ω das Innere

$$\Omega = \{(x,y,z) \in \mathbb{R}^3;\ \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\}$$

des von dem Ellipsoiden S mit Halbachsen $a, b, c > 0$ umschlossenen Gebietes. Für das Vektorfeld $K(x,y,z) = (0,0,z)^t$, $(x,y,z) \in \mathbb{R}^3$, mit $div\ K \equiv 1$ ergibt Satz 8.8.1 und die dort durchgeführte Rechnung

$$\begin{aligned}\mu_3(\Omega) &= \int_\Omega div\ K\ d\mu_3 = \int_{\partial\Omega} K \cdot n\ do \\ &= 2c \int_{\{(x,y);\ \frac{x^2}{a^2}+\frac{y^2}{b^2}<1\}} \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}\ d\mu_2(x,y) = \frac{4\pi}{3}abc\,;\end{aligned}$$

vgl. Beispiel 8.3.2.

Andererseits kann man auch gewisse Oberflächeninhalte bestimmen.

Beispiel 8.8.2. i) Sei Ω der Kegelstumpf

$$\Omega = \{(x,y,z) \in \mathbb{R}^3;\ \sqrt{x^2+y^2} < 1 - z,\ z > 0\}$$

mit Mantel

$$M = \{(x,y,z);\ \sqrt{x^2+y^2} = 1 - z \le 1\} = \mathcal{G}(\psi),$$

wo $\psi(x,y) = 1 - \sqrt{x^2+y^2}$.

Wir bestimmen den 2-dimensionalen Inhalt von M, wie folgt. Wähle $K \equiv e_3$ mit $div(K) = 0$. Mit Satz 8.8.1 erhalten wir

$$0 = \int_\Omega div \begin{pmatrix}0\\0\\1\end{pmatrix}\ d\mu = \int_{\partial\Omega} \begin{pmatrix}0\\0\\1\end{pmatrix} \cdot n\ do = \int_M n^3\ do - \mu_2\big(B_1(0;\mathbb{R}^2)\big)\,.$$

Gemäss (8.1.1) gilt

$$n^3(x,y,\psi(x,y)) = \frac{1}{\sqrt{1 + |\nabla\psi(x,y)|^2}} = \frac{1}{\sqrt{2}}\,.$$

Es folgt die Gleichung

$$\int_M n^3\ do = \frac{1}{\sqrt{2}}\mu_2(M) = \mu_2\big(B_1(0;\mathbb{R}^2)\big) = \pi;$$

d.h.

$$\mu_2(M) = \sqrt{2} \cdot \pi.$$

ii) Was ist die Oberfläche der Sphärenkappe

$$S_{z_0} = \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 = 1,\ z > z_0\}$$

für gegebenes $z_0 > 0$? – Zusammen mit

$$D_{z_0} = \{(x,y,z) \in \mathbb{R}^3;\ x^2+y^2+z^2 \leq 1,\ z = z_0\}$$

berandet S_{z_0} ein Gebiet $\Omega \subset \mathbb{R}^3 \setminus \{0\}$. Setze

$$K(p) = \frac{p}{|p|^3},\quad p \in \mathbb{R}^3\backslash\{0\}\ .$$

Beachte die Gleichung

$$div(K) = \frac{\partial}{\partial x}\Big(\frac{x}{\sqrt{x^2+y^2+z^2}^3}\Big) + \cdots = 0 \ \text{ in } \mathbb{R}^3\backslash\{0\}.$$

Mit Satz 8.8.1 folgt

$$\begin{aligned}\mu_2(S_{z_0}) &= \int_{S_{z_0}} K\cdot n\ do = \int_{D_{z_0}} K\cdot e_3\ d\mu(x,y)\\ &= z_0\int_{D_{z_0}} \frac{d\mu(x,y)}{\sqrt{x^2+y^2+z_0^2}^3} = \pi z_0\int_0^{\sqrt{1-z_0^2}} \frac{r\ dr}{\sqrt{r^2+z_0^2}^3}\\ &= 2\pi z_0\int_0^{1-z_0^2}\frac{ds}{\sqrt{s+z_0^2}^3} = \pi z_0\cdot\Big(-\frac{2}{\sqrt{s+z_0^2}}\Big)\Big|_{s=0}^{1-z_0^2} = 2\pi\ (1-z_0),\end{aligned}$$

wobei

$$D_{z_0} = \{(x,y,z_0);\ x^2+y^2 < 1-z_0^2\}.$$

Index

äquivalente Normen, 64

abgeschlossen, 60, 146
Ableitung, 79
Abschluss, 51, 61
Absolutbetrag, 24
Absolute Konvergenz, 45
algebraisch vollständig, 26
Alternierende harmonische Reihe, 38
Anfangswertproblem, 104
Argument, 26
Aussagen, 3
Axiom, 3

Banachraum, 77
beschränkt, 39
beschränkt, nach oben, 16
beschränkt, nach unten, 16
bijektiv, 9

Cauchy-Folge, 37
charakteristisches Polynom, 107

Definitionsbereich, 7
Descartesches Blatt, 186
Diffeomorphismus, 211
Differential, 155
Differentialform, 165
Differentialgleichungen, 103
Differentialgleichungssystem, 104
Differenz, 6
differenzierbar, 79, 155
divergent, 28
Divergenz, 224
Durchschnitt, 6

einfach zusammenhängend, 209, 222
Einheitswurzel, 26
Elementarfigur, 202
Elementarinhalt, 195
Euklidische Norm, 21
Eulersche Zahl, 32

Exponentialreihe, 42

Feinheit, 195
Fibonacci Zahlen, 27
Flächeninhalt, 217
Fluss, 219
Fundamentallösung, 106

Geometrische Reihe, 27, 40
gerichtetes Längenelement, 171
Gleichgewichtslösung, 114
gleichmässig stetig, 74
gleichmässige Konvergenz, 75
Gradientenfeld, 165
Graph, 8
Grenzwert, 52
Gruppe, 12

Häufungspunkt, 34
Harmonische Reihe, 38, 40
Hessesche Form, 176

Imaginärteil, 23
Infimum, 17
injektiv, 9
Innerer Punkt, 59
Integral, 117

Jordan-messbar, 201
Jordansches Mass, 201

kompakt, 57
komplexe Multiplikation, 22
Komposition, 8
Konjugation, 23
konservatives Vektorfeld, 171
kontrahierend, 146
konvergent, 28
konvex, 100
kritischer Punkt, 176, 191

Lagrange-Multiplikator, 191

Lagrangefunktion, 191
Limes inferior, 34
Limes superior, 34
Linearkombination, 20
Lipschitz stetig, 55
Lipschitzkonstante, 55
Lokal Lipschitz-stetig, 56
lokale Immersion, 216
lokale Minimalstelle, 98

Maximum, 17
Menge, 5
Minimum, 17
monoton wachsend, 70

Natürlicher Logarithmus, 72
Norm, 64
Normalbereich, 206
normiert, 21

oberes Riemann-Integral, 130
offen, 59
offener Ball, 59
offener Kern, 60
orthogonal, 21
orthonormal, 21

Partialbruchzerlegung, 124
Partialsummen, 40
partikuläre Lösung, 112
Polarform, 24
Polarkoordinaten, 214
Potential, 171
Potentialfeld, 171
Potenzreihe, 43
punktweise Konvergenz, 75

Quader, 195

Rand, 61
Rang, 185
Realteil, 23
regulärer Punkt, 186
relativ abgeschlossen, 66
relativ offen, 66
Richtungsableitung, 160
Riemann-integrabel, 130, 197
Riemann-Integral, 130, 196
Rotation, 221

Schranke, obere, 16
Schranke, untere, 16
senkrecht, 21
Separation der Variablen, 123
Skalarer Flächeninhalt, 218
Skalarprodukt, 21
Stammfunktion, 117
Standardbasis, 20
stetig, 52, 56
stetig ergänzbar, 52
Supremum, 17
Supremumsnorm, 73
surjektiv, 9

Tangentialraum, 164, 192
Taylor-Polynom, 174
Teilfolge, 34
Treppenfunktion, 128, 195

Umgebung, 66
Umkehrabbildung, 9
unbestimmtes Integral, 118
uneigentlich R-integrabel, 140
unteres Riemann-Integral, 130
Urbild, 9

Vektorraum, 20
Vereinigung, 6
Verfeinerung, 196
Verzinsung, 27

Wahrheitstafel, 4
Wegintegral, 166
wegzusammenhängend, 72
Wertebereich, 7

Zahlen, ganze, 11
Zahlen, natürliche, 11
Zahlen, rationale, 11
Zahlen, reelle, 12
Zahlenstrahl, 11
Zahlkörper, 11
Zerlegung, 195