

Analysis I und II

Prof. Michael Struwe

Herbstsemester 2011/Frühlingssemester 2012

ETH Zürich

Vorwort

Das vorliegende Skript entstand parallel zu meiner gleichnamigen Vorlesung im akademischen Jahr 2011/12, wobei ich mich auf meine Aufzeichnungen zur gleichen Vorlesung im akademischen Jahr 2005/06 und die von Frau Eveline Hardmeier in LaTeX erfasste Mitschrift dieser Vorlesung stützen konnte, ebenso wie auf deren spätere Umarbeitung für meine Vorlesung “Analysis für Informatik” in den akademischen Jahren 2008/09 und 2009/10, bei deren Erfassung in LaTeX mir Frau Manuela Dübendorfer geholfen hat. Auch etliche der im Jahr 2005/06 von Herrn Mathias Weyland sowie im Sommer 2010 von Herrn Jorim Jaggi erstellten Graphiken konnte ich im vorliegenden Skript verwenden und das unter Mithilfe von Herrn Jaggi gestaltete lay-out des Skripts der “Analysis für Informatik” übernehmen. An dieser Stelle möchte ich daher Frau Hardmeier, Frau Dübendorfer, Herrn Jaggi sowie Herrn Weyland erneut von Herzen danken für ihren grossen geleisteten Einsatz; ebenso Frau Prisca Greminger für ihre Mithilfe beim Korrigieren der Mitschrift von Frau Hardmeier und Herrn Joachim Näf für das überaus sorgfältige Korrekturlesen des vorliegenden Skripts. Herr Näf hat diese Arbeit so gründlich erledigt, dass von den Studierenden dieser Vorlesung nur noch wenige Korrekturhinweise kamen; dennoch möchte ich nicht versäumen, auch den Studierenden des Jahrgangs 2011/12 ebenso wie denjenigen des Jahrgangs 2005/06 für ihre vielen wertvollen Anregungen zu danken.

Natürlich wollen wir im vorliegenden Text den Stoff der Analysis nicht neu erfinden. Das Skript stützt sich auf die umfangreiche Literatur zu diesem Thema; insbesondere verdanke ich den Lehrbüchern meines Kollegen Christian Blatter den Hinweis auf den eleganten Zugang zum Transformationssatz für das Jordansche Mass mittels Lemma 9.3.1.

Zürich, 6. September 2012

Michael Struwe

Inhaltsverzeichnis

1	Logik und Grundlagen	3
1.1	Logik	3
1.1.1	Logische Verknüpfungen	4
1.2	Mengenlehre	5
1.2.1	Mengen und Quantoren	5
1.2.2	Verknüpfungen	7
1.3	Funktionen	8
1.3.1	Komposition von Abbildungen	9
1.3.2	Urbildfunktion	11
1.4	Relationen	12
1.4.1	Äquivalenzrelation	12
1.4.2	Ordnungsrelation	14
2	Die reellen Zahlen	17
2.1	Elementare Zahlen	17
2.2	Die reellen Zahlen	18
2.2.1	Axiome für \mathbb{R}	18
2.2.2	Supremum und Infimum	23
2.3	Die natürlichen Zahlen	26
2.4	Kardinalität	28
2.5	Der euklidische Raum	31
2.6	Komplexe Zahlen	35
3	Folgen und Reihen	39
3.1	Beispiele	39

3.2	Grenzwert einer Folge	39
3.3	Monotone Konvergenz	44
3.4	Teilfolgen, Häufungspunkte	46
3.5	Cauchy-Kriterium	50
3.6	Folgen in \mathbb{R}^d oder \mathbb{C}	53
3.7	Reihen	54
3.8	Absolute Konvergenz	59
3.9	Die Exponentialreihe und die Funktion e^x	62
4	Topologische Grundbegriffe	65
4.1	Topologie des \mathbb{R}^n	65
4.2	Topologie und Konvergenz	70
4.3	Kompakte Mengen	72
4.4	Zusammenhängende Mengen	75
4.5	Relativtopologie	77
4.6	Hausdorffräume	79
4.7	Ausblick	80
5	Stetigkeit	81
5.1	Grenzwerte von Funktionen	81
5.1.1	Lipschitz stetige Funktionen	84
5.2	Stetigkeitskriterien	86
5.3	Stetige Funktionen	88
5.3.1	Stetigkeit und Kompaktheit.	89
5.3.2	Eine Norm für $C^0(K; \mathbb{R}^n)$	90
5.3.3	Homöomorphismen	91
5.4	Äquivalenz von Normen auf \mathbb{R}^d	92
5.5	Stetige Fortsetzung	94
5.6	Zwischenwertsatz und Folgerungen	96
5.6.1	Monotone Funktionen	98
5.6.2	Topologische Anwendungen	100
5.7	Folgen stetiger Funktionen	102
5.7.1	Punktweise und gleichmässige Konvergenz	102

5.7.2	Monotonie und Konvergenz	104
5.7.3	Der Raum $C^0(\overline{\Omega})$	105
6	Differentialrechnung auf \mathbb{R}	109
6.1	Differential und Differentiationsregeln	109
6.1.1	Differentiationsregeln	111
6.2	Der Mittelwertsatz und Folgerungen	113
6.3	Die trigonometrischen Funktionen	119
6.3.1	Zyklometrische Funktionen (Arcus-Funktionen).	121
6.3.2	Hyperbel- und Areafunktionen.	122
6.4	Funktionen der Klasse C^1	123
6.5	Höhere Ableitungen	127
6.5.1	Der Raum $C^m(\Omega)$	127
6.5.2	Taylor-Formel	129
6.5.3	Lokale Extrema.	130
6.5.4	Konvexe Funktionen.	132
6.5.5	“Glatte” Funktionen, analytische Funktionen	136
6.6	Gewöhnliche Differentialgleichungen	140
6.7	Inhomogene Differentialgleichungen	150
7	Integration	155
7.1	Stammfunktionen	155
7.2	Das Riemannsches Integral	166
7.3	Integrationsregeln, Hauptsatz	174
7.3.1	Das R-Integral vektorwertiger Funktionen	179
7.4	Uneigentliches Riemann-Integral	180
7.5	Differentialgleichungen	182
7.5.1	Der Banachsche Fixpunktsatz	186
7.5.2	Beweis von Satz 7.5.1	188
7.5.3	Stetige Abhängigkeit von den Daten	191
7.5.4	Globale Fortsetzbarkeit	192
8	Differentialrechnung im \mathbb{R}^n	197
8.1	Partielle Ableitungen und Differential	197

8.1.1	Funktionen der Klasse C^1	201
8.1.2	Landau-Symbole	202
8.2	Differentiationsregeln	203
8.3	Differentialformen und Vektorfelder	210
8.4	Wegintegrale	212
8.4.1	Konservative Vektorfelder	217
8.5	Höhere Ableitungen	218
8.5.1	Taylor-Näherung m -ter Ordnung	220
8.5.2	Der Raum $C^m(\overline{\Omega})$	224
8.5.3	Reell analytische Funktionen	225
8.6	Vektorwertige Funktionen	226
8.7	Der Umkehrsatz	228
8.8	Implizite Funktionen	233
8.9	Extrema mit Nebenbedingungen	239
8.9.1	Notwendige Bedingungen	240
8.9.2	Geometrische Deutung	243
8.9.3	Hinreichende Bedingungen	247
8.10	Immersionen	249
9	Integration im \mathbb{R}^n	253
9.1	Riemannsches Integral über einem Quader	253
9.1.1	Zerlegungen und Treppenfunktionen	253
9.1.2	Das Riemann-Integral	255
9.2	Der Satz von Fubini	257
9.3	Jordan-Bereiche	260
9.3.1	Das Jordansche Mass	260
9.3.2	Das Riemann-Integral über Jordan-Bereiche	266
9.4	Der Satz von Green	268
9.5	Substitutionsregel	275
9.5.1	Lineare Transformationen	275
9.5.2	Transformation Jordan-messbarer Mengen	276
9.5.3	Substitutionsregel	280
9.6	Oberflächenmass und Fluss-Integral	282

9.6.1	Der Inhalt eines parametrischen Flächenstücks	283
9.6.2	Der Fluss eines Vektorfeldes	286
9.7	Der Satz von Stokes im \mathbb{R}^3	286
9.8	Der Satz von Gauss im \mathbb{R}^3	292
9.8.1	Partielle Integration	296
9.8.2	Harmonische Funktionen	297
9.8.3	Transformation des Dirichlet-Integrals	299
9.9	k -Formen im \mathbb{R}^n	301
9.9.1	Integration von k -Formen	304
9.9.2	Der Satz von Stokes im \mathbb{R}^n	307
9.9.3	Der allgemeine Satz von Stokes	310

Kapitel 1

Logik und Grundlagen

1.1 Logik

Beispiele für mathematische **Aussagen**:

- i) “ $4 > 2$ ” (wahr)
- ii) “ $\forall n \in \mathbb{N} : n > 4 \Rightarrow n > 2$ ” (wahr)
- iii) “ $5 < 3$ ” (falsch)

In der Mathematik stützen wir uns auf gewisse Grundannahmen, sogenannte “*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 rückzügliche 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.

1.1.1 Logische Verknüpfungen

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 ($A \rightarrow B$) und die Äquivalenz ($A \leftrightarrow B$) sind definiert durch die **Wahrheitstafel**.

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$
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

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

Beispiel 1.1.1. i) “ $(n > 4) \rightarrow (n > 2)$ ”. *Beachte:* Weder die Annahme (Voraussetzung) “ $n > 4$ ” noch die Folgeaussage “ $n > 2$ ” ist für alle $n \in \mathbb{N}$ erfüllt, die Implikation ist jedoch stets wahr.

ii) Ist die Aussage A falsch, so ist für eine beliebige Aussage B die Implikation $A \rightarrow B$ immer wahr. In der Politik macht man sich dies gern zunutze: Die Aussage “Wenn das Volk damals anders entschieden hätte, dann ...” ist bei beliebiger Fortsetzung korrekt. (Conjunctivus irrealis, “*ex falso quodlibet*”).

Eine **wahre** Implikation $A \rightarrow B$ 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:

$$(A \rightarrow B) \wedge (B \rightarrow C) \Rightarrow (A \rightarrow C).$$

Wir können daher über eine Kette von Folgerungen

$$A \Rightarrow B \Rightarrow \dots \Rightarrow S$$

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

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

Satz 1.1.1. *Es gilt*

$$(A \rightarrow B) \Leftrightarrow (\neg B \rightarrow \neg A).$$

Beweis. Die Aussage folgt sofort mit der Wahrheitstafel:

A	B	$A \rightarrow B$	$\neg A$	$\neg B$	$\neg B \rightarrow \neg A$
w	w	w	f	f	w
w	f	f	f	w	f
f	w	w	w	f	w
f	f	w	w	w	w

Die Spalten von $A \rightarrow B$ und $\neg B \rightarrow \neg A$ sind gleich. \square

Aus Satz 1.1.1 können wir einige nützliche Regeln ableiten.

Umkehrschluss (Kontraposition): Die (wahre) Aussage " $A \Rightarrow B$ " ist nach Satz 1.1.1 gleichbedeutend mit " $(\neg B) \Rightarrow (\neg A)$ ". Falls $A \Rightarrow B$, so kann A nicht wahr sein, wenn B falsch ist. (" B ist notwendig für A .")

Weiter folgt das **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! \square

Satz 1.1.2. Es gelten die Äquivalenzen

i) $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$,

ii) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$.

Beweis. i) Wieder arbeiten wir mit der Wahrheitstafel:

A	B	$A \vee B$	$\neg(A \vee B)$	$\neg A$	$\neg B$	$(\neg A) \wedge (\neg B)$
w	w	w	f	f	f	f
w	f	w	f	f	w	f
f	w	w	f	w	f	f
f	f	f	w	w	w	w

Die Behauptung folgt aus der Gleichheit der entsprechenden Spalten. Analog folgt ii). \square

1.2 Mengenlehre

1.2.1 Mengen und Quantoren

Georg Cantor: "Eine Menge ist die ungeordnete Zusammenfassung verschiedener Elemente zu einem Ganzen."

Beispiel:

i) Für $a \neq b$ gilt $\{a, b\} = \{b, a\} = \{a, b, a\}$;

ii) $\mathbb{N} = \{1, 2, 3, 4, \dots\}$;

iii) $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\} = \mathbb{N} \cup \{0\}$;

iv) $\emptyset = \{\}$: leere Menge;

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

Nicht alle Bildungsgesetze sind zulässig. Insbesondere müssen wir uns vor rückbezüglichen Definitionen hüten, wie das folgende Beispiel zeigt:

vi) (Bertrand Russell) Die Menge M aller Mengen, die sich selbst als Element nicht enthalten, gibt es nicht. (Wäre $M \in M$, so gehörte M nach Definition nicht zu M ; falls jedoch $M \notin M$, so müsste M zu M gehören.)

Das Russellsche 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.

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

\forall : der Allquantor (“für alle”),
 \exists : der Existenzquantor (“es gibt”).

Beispiel 1.2.1. 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 $n_0 \in \mathbb{N}$ ” aus 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 Aussage ii).)

Man kann den All- und Existenzquantor mit Mengen wie folgt definieren:

$$\begin{aligned} \forall x \in M : A(x) &\Leftrightarrow \{x \in M; A(x)\} = M, \\ \exists x \in M : A(x) &\Leftrightarrow \{x \in M; A(x)\} \neq \emptyset. \end{aligned}$$

Satz 1.2.1. *Es gilt:*

$$\text{i) } \neg(\forall x \in M : A(x)) \Leftrightarrow (\exists x \in M : \neg A(x)) ;$$

$$\text{ii) } \neg(\exists x \in M : A(x)) \Leftrightarrow (\forall x \in M : \neg A(x)) .$$

Beweis. i) Nach dem Satz vom ausgeschlossenen Dritten haben wir die disjunkte Zerlegung

$$M = \{x \in M; A(x)\} \cup \{x \in M; \neg A(x)\} .$$

Somit gilt

$$\begin{aligned} \neg(\forall x \in M : A(x)) &\Leftrightarrow \{x \in M; A(x)\} \neq M \\ &\Leftrightarrow \{x \in M; \neg A(x)\} \neq \emptyset \Leftrightarrow \exists x \in M : \neg A(x). \end{aligned}$$

Analog erhält man ii). □

1.2.2 Verknüpfungen

Seien X, Y beliebige Mengen:

$X \cup Y = \{x; x \in X \vee x \in Y\}$: Vereinigungsmenge,

$X \cap Y = \{x; x \in X \wedge x \in Y\}$: Durchschnitt,

$X \setminus Y = \{x; x \in X, x \notin Y\}$: Differenz,

$X \subset Y$: Teilmenge (Inklusion).

Bemerkung 1.2.1. i) Sei M eine Menge und seien $X = \{x \in M; A(x)\}$, $Y = \{x \in M; B(x)\}$. Dann gilt z.B.

$$X \cup Y = \{x \in M; A(x) \vee B(x)\},$$

$$X \cap Y = \{x \in M; A(x) \wedge B(x)\};$$

man kann daher logische Verknüpfungen durch Verknüpfungen von Mengen veranschaulichen, und umgekehrt.

ii) Es gilt

$$X = Y \Leftrightarrow (X \subset Y) \wedge (Y \subset X);$$

die Gleichheit von Mengen entspricht also zwei Inklusionen.

iii) Es gilt das "*Distributivgesetz*"

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z), \text{ etc.};$$

vergleiche die Übungen.

Beweis. iii) Mit der Äquivalenz $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$ (mittels Wahrheitstafel) erhalten wir für beliebiges x :

$$\begin{aligned} x \in X \cap (Y \cup Z) &\Leftrightarrow x \in X \wedge x \in (Y \cup Z) \\ &\Leftrightarrow (x \in X) \wedge (x \in Y \vee x \in Z) \\ &\Leftrightarrow (x \in X \wedge x \in Y) \vee (x \in X \wedge x \in Z) \\ &\Leftrightarrow (x \in X \cap Y) \vee (x \in X \cap Z) \\ &\Leftrightarrow x \in ((X \cap Y) \cup (X \cap Z)). \end{aligned}$$

□

iv) Die Teilmengen einer Menge X bilden deren *Potenzmenge*

$$\mathcal{P}(X) = \{Y; Y \subset X\}.$$

iv) Die *geordneten Paare* (x, y) von Elementen der Mengen X , bzw. Y bilden deren *Produktmenge*

$$X \times Y = \{(x, y); x \in X, y \in Y\}.$$

Satz 1.2.2. (de Morgan) Sei I eine beliebige Indexmenge und für $i \in I$ sei $A_i \subset X$ mit Komplement $A_i^c = X \setminus A_i$. Dann gilt:

- i) $(\cup_{i \in I} A_i)^c = \cap_{i \in I} A_i^c$,
 ii) $(\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c$.

Beweis. i) Es gilt für jedes $x \in X$:

$$\begin{aligned} x \in (\cup_{i \in I} A_i)^c &\Leftrightarrow x \notin \cup_{i \in I} A_i \Leftrightarrow \forall i \in I : x \notin A_i \\ &\Leftrightarrow \forall i \in I : x \in A_i^c \Leftrightarrow x \in \cap_{i \in I} A_i^c. \end{aligned}$$

Die Aussage ii) beweist man analog; vergleiche die Übungen. \square

Beispiel 1.2.2. i) Im Falle $I = \{1, 2\}$, $A_1 =: A \subset X$, $A_2 =: B \subset X$ erhalten wir

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c;$$

beachte die Ähnlichkeit mit Satz 1.1.2.

ii) Sei $X = \mathbb{R} = I$. Für $a \in \mathbb{R}$ setze $A_a =]-\infty, a[= \{x \in \mathbb{R}; x < a\}$ mit $A_a^c = [a, \infty[= \{x \in \mathbb{R}; x \geq a\}$. Dann gilt

$$\cup_{a \in \mathbb{R}} A_a = \mathbb{R}, \quad \cap_{a \in \mathbb{R}} A_a^c = \emptyset, \quad \text{etc.}$$

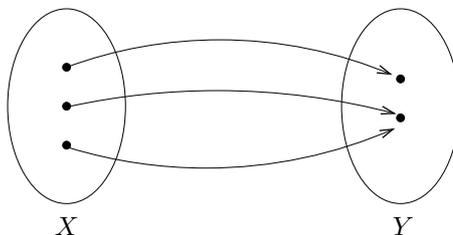
1.3 Funktionen

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

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

Allgemein seien X, Y Mengen.

Definition 1.3.1. Eine Funktion (oder Abbildung) $f : X \rightarrow Y$ ordnet jedem Punkt $x \in X$ genau ein "Bild" $y = f(x) \in Y$ zu. Jedes $z \in X$ mit $y = f(z)$ heisst dann ein "Urbild" von y .



Das heisst, eine Funktion wird erklärt durch die Angabe

- des *Definitionsbereiches* (hier X)

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

Beispiel 1.3.1. i) $f : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto x - x^3$;

ii) $f : [-1, 1] \rightarrow \mathbb{R}; x \mapsto x - x^3$;

iii) $g : \mathbb{R} \rightarrow [-1, 1], x \mapsto \sin(x)$;

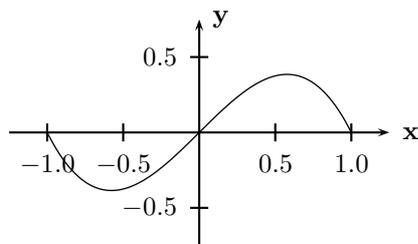
iv) $h : \mathbb{R} \rightarrow [0, \infty[, x \mapsto x^2$;

v) $\text{id}_X : X \rightarrow X, x \mapsto x = \text{id}_X(x)$; die Identität auf X .

Wir können Funktionen durch ihren *Graphen* darstellen:

$$G(f) = \{(x, f(x)); x \in X\} \subset X \times Y.$$

Zum Beispiel erhalten wir für f aus Beispiel 1.3.1.ii) den Graphen



Dies geht auch allgemein (jedoch abstrakt).

1.3.1 Komposition von Abbildungen

Seien $f : X \rightarrow Y, g : Y \rightarrow Z$ Abbildungen.

Durch Komposition erhalten wir eine neue Abbildung

$$g \circ f : X \rightarrow Z, \quad x \mapsto g(f(x)).$$

$$\begin{array}{ccc}
 X & \xrightarrow{f} & Y & \xrightarrow{g} & Z \\
 & & \searrow & \nearrow & \\
 & & & & F = g \circ f
 \end{array}$$

Satz 1.3.1. Für Abbildungen $f : X \rightarrow Y, g : Y \rightarrow Z, h : Z \rightarrow W$ gilt:

$$h \circ (g \circ f) = (h \circ g) \circ f \quad (\text{Assoziativgesetz}).$$

Beweis. i) Die Definitionsbereiche X sind identisch.

ii) Ebenso sind die Wertebereiche W identisch.

iii) Schliesslich prüfen wir die Gleichheit der Zuordnungsvorschrift. Für alle $x \in X$ gilt:

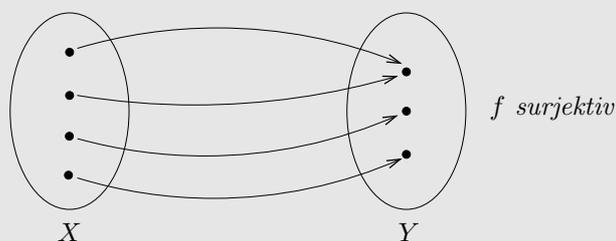
$$\begin{aligned} (h \circ (g \circ f))(x) &= h((g \circ f)(x)) = h(g(f(x))) \\ &= (h \circ g)(f(x)) = ((h \circ g) \circ f)(x). \end{aligned}$$

□

Definition 1.3.2. Sei $f : X \rightarrow Y$ eine Abbildung.

i) f heisst **surjektiv**, falls jedes $y \in Y$ **mindestens** ein Urbild hat, das heisst, falls

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

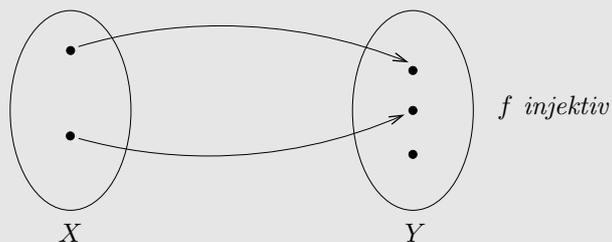


ii) f heisst **injektiv** oder **eineindeutig**, falls jedes $y \in Y$ **höchstens** ein Urbild hat, das heisst, falls

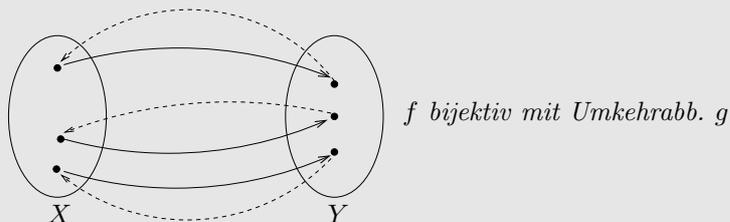
$$\forall x_1, x_2 \in X : f(x_1) = f(x_2) \Rightarrow x_1 = x_2,$$

oder, hierzu äquivalent,

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



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



Falls f bijektiv ist (und nur in diesem Fall), können wir eine Abbildung $g : Y \rightarrow X$ 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 , bzw. die zu f *inverse Abbildung*, und wir schreiben $g = f^{-1}$.

Beispiel 1.3.2. i) Die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}; x \mapsto x - x^3$ ist surjektiv, aber nicht injektiv.

ii) Die Abbildung $f :] - \pi/2, \pi/2[\rightarrow] - 1, 1[$ mit $x \mapsto \sin(x)$ ist bijektiv.

Satz 1.3.2. Sei $f : X \rightarrow Y$ eine Abbildung. Dann gilt:

i) f injektiv $\Rightarrow \exists g : Y \rightarrow X$ mit $g \circ f = id_X$;

ii) f surjektiv $\Rightarrow \exists g : Y \rightarrow X$ mit $f \circ g = id_Y$;

iii) f bijektiv $\Rightarrow \exists g : Y \rightarrow X$ mit $f \circ g = id_Y, g \circ f = id_X$.

Beweis. i) Für alle $y \in f(X) = \{f(x); x \in X\} \subset Y$ gibt es genau ein Urbild $x =: g(y)$. Fixiere $x_0 \in X$ und setze $g(y) = x_0$ für $y \notin f(X)$. Dann ist $g : Y \rightarrow X$ wohldefiniert und $g \circ f = id_X$.

ii) Zu $y \in Y$ ist $A(y) = \{x \in X; f(x) = y\} \neq \emptyset$. Wähle ein beliebiges $x \in A(y)$ und setze $g(y) = x$. (Hier benutzen wir das Auswahlaxiom, s.u.). Dann ist $g : Y \rightarrow X$ wohldefiniert und $f \circ g = id_Y$.

□

Bemerkung 1.3.1. Offenbar gelten auch die Umkehrungen der Aussagen i) - iii) in Satz 1.3.2.

1.3.2 Urbildfunktion

Sei $f : X \rightarrow Y$ eine Abbildung, und seien $A \subset X, B \subset Y$. Analog zum Bild

$$f(A) = \{f(x); x \in A\} \subset Y$$

der Menge A unter f können wir das Urbild

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

der Menge B unter f einführen. Auf diese Weise wird eine Funktion

$$f^{-1} : \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$$

definiert, die *Urbildfunktion*. Hierzu muss f nicht bijektiv sein.

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

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

Satz 1.3.3. Für $B, C \subset Y$ gilt:

- i) $f^{-1}(B \cup C) = f^{-1}(B) \cup f^{-1}(C)$;
- ii) $f^{-1}(B \cap C) = f^{-1}(B) \cap f^{-1}(C)$.

Beweis. i) Für beliebiges $x \in X$ gilt

$$\begin{aligned} x \in f^{-1}(B \cup C) &\Leftrightarrow f(x) \in (B \cup C) \\ &\Leftrightarrow f(x) \in B \vee f(x) \in C \\ &\Leftrightarrow x \in f^{-1}(B) \vee x \in f^{-1}(C) \\ &\Leftrightarrow x \in (f^{-1}(B) \cup f^{-1}(C)). \end{aligned}$$

Analog folgt ii); vergleiche die Übungen. □

Satz 1.3.4. f ist bijektiv genau dann, wenn $f^{-1}(\{y\})$ für jedes $y \in Y$ genau ein Element enthält.

Falls f bijektiv ist, so bezeichnen wir nach Definition 1.3.2 mit $f^{-1} : Y \rightarrow X$ auch die *Umkehrfunktion* von f . Offenbar gilt in diesem Fall

$$\underbrace{f^{-1}(\{y\})}_{\text{Urbildfkt.}} = \{ \underbrace{f^{-1}(y)}_{\text{Umkehrabb.}} \},$$

was die Mehrdeutigkeit der Bezeichnungen rechtfertigt.

1.4 Relationen

Elemente einer Menge können zueinander in vielfältiger Weise in Beziehung stehen. Zum Beispiel können Aussagen zueinander *äquivalent* sein. Ein weiteres Beispiel ist die *Ordnung* natürlicher Zahlen nach ihrer Grösse.

1.4.1 Äquivalenzrelation

Sei X eine beliebige Menge.

Definition 1.4.1. Eine Beziehung " \sim " auf X heisst *Äquivalenzrelation* auf X , falls gilt:

- i) *Reflexivität:* $\forall x \in X : x \sim x$;
- ii) *Symmetrie:* $\forall x, y \in X : x \sim y \Rightarrow y \sim x$;
- iii) *Transitivität:* $\forall x, y, z \in X : x \sim y \wedge y \sim z \Rightarrow x \sim z$.

- Beispiel 1.4.1.** i) “=” auf beliebiger Menge;
 ii) die logische Äquivalenz von Aussagen (vergleiche Übung);
 iii) “Reste modulo p ”. Sei $p \in \mathbb{N}$ fest. Für $m, n \in \mathbb{Z}$ setze $m \sim n$, falls $m = n + kp$ für ein $k \in \mathbb{Z}$. Wir schreiben: $m = n \pmod{p}$.

Beweis. Wir verifizieren die Eigenschaften einer Äquivalenzrelation.

- i) Reflexivität: Offenbar gilt $m \sim m$ (wähle $k = 0$).
 ii) Symmetrie: Es gelte $m \sim n$, das heisst, es existiert $k \in \mathbb{Z}$ mit $m = n + kp$. Dann erhalten wir $n = m - kp = m + (-k)p$, also $n \sim m$.
 iii) Transitivität: Seien $m \sim n$, $n \sim \ell$, das heisst, $m = n + kp$, $n = \ell + jp$, für gewisse $j, k \in \mathbb{Z}$. Es folgt

$$m = n + kp = \ell + jp + kp = \ell + p(j + k);$$

das heisst, $m \sim \ell$.

□

Sei \sim eine Äquivalenzrelation auf X , $x \in X$. Die Menge

$$[x] := \{y \in X \mid x \sim y\}$$

heisst *Äquivalenzklasse* von X . Es gilt:

Behauptung 1. $\forall y \in [x] : [y] = [x]$.

Beweis. Wir zeigen “ \subset ” und “ \supset ”.

- “ \subset ” Sei $z \in [y]$; das heisst $y \sim z$. Da $x \sim y$ folgt mit Transitivität $x \sim z$, also $z \in [x]$.
 “ \supset ” Da $y \in [x]$, also $x \sim y$, folgt mit Symmetrie auch $y \sim x$, also $x \in [y]$, und “ \supset ” folgt mit i).

□

Behauptung 2. $\forall y \in X : y \notin [x] \Rightarrow [y] \cap [x] = \emptyset$.

Beweis (indirekt). Sei $y \notin [x]$, und sei $z \in [x] \cap [y]$. Mit Behauptung 1 folgt $[x] = [z] = [y] \ni y$. Widerspruch. □

Wir folgern:

Satz 1.4.1. Eine Äquivalenzrelation auf X definiert eine disjunkte Zerlegung von X in Äquivalenzklassen.

Bemerkung 1.4.1. Umgekehrt definiert eine Zerlegung

$$X = \bigcup_{i \in I} X_i$$

von X in disjunkte Mengen X_i eine Äquivalenzrelation \sim mittels

$$x \sim y \Leftrightarrow \exists i \in I : x, y \in X_i .$$

1.4.2 Ordnungsrelation

Definition 1.4.2. Eine Beziehung " \leq " auf einer Menge X heisst **partielle Ordnung**, falls gilt:

- i) Reflexivität: $\forall x \in X : x \leq x$;
- ii) Transitivität: $\forall x, y, z \in X : x \leq y \wedge y \leq z \Rightarrow x \leq z$;
- iii) Identivität: $\forall x, y \in X : x \leq y \wedge y \leq x \Rightarrow x = y$.

Beispiel 1.4.2. i) \mathbb{N} mit der natürlichen Ordnung " \leq ";

ii) $\mathcal{P}(M)$ mit $A \leq B :\Leftrightarrow A \subset B$.

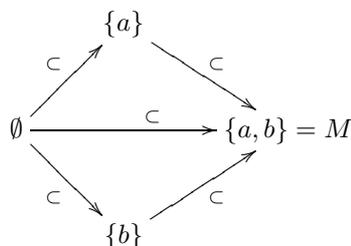
(Betrachte zum Beispiel $M = \{a, b\}$ mit $\mathcal{P}(M) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.)

Bemerkung 1.4.2. Man kann eine partiell geordnete Menge (X, \leq) durch einen gerichteten Graphen, den *Ordnungsgraphen* veranschaulichen.

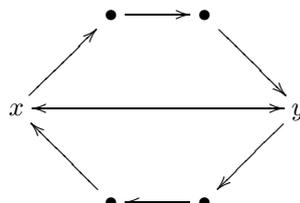
i) (\mathbb{N}, \leq) :

$$1 \longrightarrow 2 \longrightarrow 3 \longrightarrow \dots$$

ii) $(\mathcal{P}(\{a, b\}), \subset)$:



Ein Ordnungsgraph darf keine Zyklen enthalten, sonst folgt mit $x \leq y$ und $y \leq x$ sofort $x = y$ für alle Elemente des Zyklus.



Sei (X, \leq) partiell geordnet.

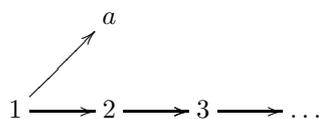
Definition 1.4.3. i) $m \in X$ heisst **maximal**, falls gilt:

$$\forall x \in X : m \leq x \Rightarrow x = m.$$

Das heisst, genau diejenigen $m \in X$ sind maximal, zu denen keine grösseren Elemente existieren. Jedoch muss ein maximales $m \in X$ nicht grösser sein als alle $x \neq m$.

ii) $m \in X$ heisst **obere Schranke** für eine Menge $L \subset X$, falls $\ell \leq m$ für alle $\ell \in L$.

Beispiel 1.4.3. Sei $X = \mathbb{N} \cup \{a\}$ mit dem Ordnungsgraphen



a ist maximales Element, aber $3 \not\leq a$.

Definition 1.4.4. (X, \leq) heisst **total** oder **linear geordnet**, falls gilt:

$$\forall x, y \in X : x \leq y \text{ oder } y \leq x.$$

Beispiel 1.4.4. i) (\mathbb{N}, \leq) ist total geordnet.

ii) $(\mathcal{P}(M), \subset)$ ist im Allgemeinen nicht total geordnet.

Zornsches Lemma: Sei (X, \leq) partiell geordnet und es gelte

$$\begin{array}{l} \text{Zu jeder total geordneten Teilmenge } L \text{ von } X \\ \text{gibt es eine obere Schranke } m \in X. \end{array} \quad (1.4.1)$$

Dann gibt es zu jedem $x \in X$ ein maximales Element $m \in X$ mit $x \leq m$.

Beispiel 1.4.5. (\mathbb{N}, \leq) hat kein maximales Element:

$$1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow \dots$$

Die Menge \mathbb{N} selbst ist total geordnet, es gibt aber keine obere Schranke für \mathbb{N} ; vergleiche Beispiel 1.1.2.

Das Zornsche Lemma ist ein Axiom, äquivalent zum *Auswahlaxiom*.

Auswahlaxiom: Sei $(A_\iota)_{\iota \in I}$ eine Familie nichtleerer Mengen $A_\iota \subset M$, $\iota \in I$. Dann gibt es eine Funktion $g : I \rightarrow M$, so dass für alle $\iota \in I$ gilt $g(\iota) \in A_\iota$. (Zermelo; vergleiche NZZ 29.9.2004)

Wir zeigen eine Richtung der Äquivalenz.

Satz 1.4.2. *Zornsches Lemma \Rightarrow Auswahlaxiom.*

Beweis. Sei

$$X = \{h : D(h) \subset I \rightarrow M; \forall \iota \in D(h) : h(\iota) \in A_\iota\},$$

wobei $D(h)$ jeweils den Definitionsbereich von h bezeichnet. Wir versehen X mit einer partiellen Ordnung, indem wir setzen

$$\forall h, \ell \in X : h \leq \ell \iff D(h) \subset D(\ell) \wedge \forall \iota \in D(h) : \ell(\iota) = h(\iota).$$

Behauptung 1. “ \leq ” definiert eine partielle Ordnung auf X .

Beweis. i) Reflexivität: Offenbar gilt $h \leq h$ für alle $h \in X$.

ii) Transitivität: Seien $h \leq k$ und $k \leq \ell$ mit $h, k, \ell \in X$. Dann gilt $D(h) \subset D(k) \subset D(\ell)$, und für alle $\iota \in D(h) \subset D(k)$ gilt $\ell(\iota) = k(\iota) = h(\iota)$, also $h \leq \ell$.

iii) Identivität: Seien $h \leq \ell$ und $\ell \leq h$ mit $h, \ell \in X$. Dann gilt $D(h) \subset D(\ell) \subset D(h)$, also $D(h) = D(\ell)$. Weiter gilt für alle $\iota \in D(h) = D(\ell)$ dann auch $\ell(\iota) = h(\iota)$, also $h = \ell$. □

Behauptung 2. (X, \leq) erfüllt die Annahme (1.4.1).

Beweis. Sei $L \subset X$ total geordnet. Setze

$$J = \bigcup_{\ell \in L} D(\ell) \subset I.$$

Für $\iota \in J$ existiert $\ell \in L$ mit $\iota \in D(\ell)$, und für jedes $h \in L$ mit $\iota \in D(h)$ gilt entweder $h \leq \ell$, also $\ell(\iota) = h(\iota)$, oder $\ell \leq h$, also ebenfalls $h(\iota) = \ell(\iota)$.

Also ist $m : J \rightarrow M$ mit $m(\iota) = \ell(\iota)$ für $\iota \in D(\ell)$, wohldefiniert; weiter gilt offenbar für alle $\ell \in L$:

$$D(\ell) \subset J = D(m), \quad m(\iota) = \ell(\iota) \text{ für jedes } \iota \in D(\ell).$$

Also ist m obere Schranke für L . □

Mit dem Zornschen Lemma folgt nun, dass X ein bezüglich \leq maximales Element g besitzt.

Behauptung 3. $D(g) = I$.

Beweis (indirekt). Sei $\iota_0 \in I \setminus D(g) \neq \emptyset$. Wähle $x_0 \in A_{\iota_0}$ und definiere $g_0 : D(g_0) = D(g) \cup \{\iota_0\} \rightarrow M$ durch

$$g_0(\iota) = \begin{cases} g(\iota), & \iota \in D(g), \\ x_0, & \iota = \iota_0. \end{cases}$$

Dann ist $g \leq g_0 \neq g$, im Widerspruch zur Maximalität von g . □

Die Abbildung g ist somit eine Auswahlfunktion. □

Im folgenden werden wir das Zornsche Lemma, beziehungsweise das Auswahlaxiom stets als eines unserer Axiome voraussetzen.

Kapitel 2

Die reellen Zahlen

2.1 Elementare Zahlen

Mit den **natürlichen Zahlen**

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

kann man Objekte abzählen. Zahlen in \mathbb{N} kann man addieren und multiplizieren.

In den **ganzen Zahlen**

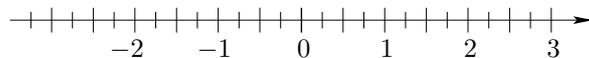
$$\mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$$

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

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

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

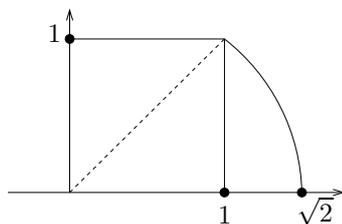
Offenbar kann man diese **elementaren Zahlen** $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$ der Grösse nach auf dem **Zahlenstrahl** anordnen.



Durch Aneinanderfügen von Längen beziehungsweise mit Hilfe des Strahlensatzes kann man die Operationen in \mathbb{Q} auch “geometrisch” ausführen.

Bemerkung 2.1.1. Zwischen je zwei rationalen Zahlen $r_1 < r_2$ liegt eine weitere, zum Beispiel $\frac{r_1 + r_2}{2} \in \mathbb{Q}$, welche den halben Abstand zu r_1 hat wie r_2 . Die rationalen Zahlen scheinen demnach die Zahlengerade lückenlos zu überdecken.

Irrationale Zahlen Dies ist jedoch nicht der Fall. Bereits die Pythagoräer erkannten, dass es keine Zahl $r \in \mathbb{Q}$ gibt mit $r^2 = 2$. Die Länge der Diagonalen im Quadrat steht in keinem rationalen Verhältnis zur Länge der Grundseite.



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 (also 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. \square

\mathbb{Q} weist also “Lücken” auf. Man kann die rationalen Zahlen auf verschiedene Weisen *vervollständigen*, beispielsweise

- mit dem *Dedekindschen Schnitt* (Dedekind 1858/59; siehe z.B. Königsberger, Aufgabe 2.5.13);
- oder durch *Fundamentalfolgen*. (Z.B. kann man die Zahl $\sqrt{2}$ darstellen durch Folgen $(r_k)_{k \in \mathbb{N}} \subset \mathbb{Q}$ mit $1 < r_k < 2$ und $|2 - r_k^2| < \frac{1}{k}$, $k \in \mathbb{N}$.)

Im folgenden gehen wir jedoch axiomatisch vor und *postulieren* die Existenz eines geordneten Zahlkörpers \mathbb{R} , der \mathbb{Q} umfasst und in einem zu präzisierenden Sinn *ordnungsvollständig* ist. Die eigentliche Konstruktion von \mathbb{R} verschieben wir auf später.

2.2 Die reellen Zahlen

2.2.1 Axiome für \mathbb{R}

Die reellen Zahlen \mathbb{R} bilden einen linear geordneten Zahlkörper mit den Operationen Addition und Multiplikation.

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

- A.i) Assoziativität: $\forall x, y, z \in \mathbb{R} : x + (y + z) = (x + y) + z$;
- A.ii) Neutrales Element: $\exists 0 \in \mathbb{R} \forall x \in \mathbb{R} : x + 0 = x$;
- A.iii) Inverses Element: $\forall x \in \mathbb{R} \exists y \in \mathbb{R} : x + y = 0$;
- A.iv) Kommutativität: $\forall x, y \in \mathbb{R} : x + y = y + x$.

Das heisst, \mathbb{R} bildet eine *abelsche* (kommutative) *Gruppe* bezüglich der Addition.

Bemerkung 2.2.1. i) Das neutrale Element ist eindeutig bestimmt.

ii) Das zu $x \in \mathbb{R}$ additiv inverse Element $y = -x$ ist eindeutig bestimmt.

iii) Die Kommutativität wird zum Beweis von i) und ii) nicht benötigt.

Beweis. i) Seien $0, \bar{0} \in \mathbb{R}$ neutrale Elemente, so folgt:

$$\bar{0} \stackrel{A.ii)}{=} \bar{0} + 0 \stackrel{A.iv)}{=} 0 + \bar{0} \stackrel{A.ii)}{=} 0.$$

ii) Falls y und z zu x invers, so folgt

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

iii) Übung. □

M) Es gibt eine weitere Operation \cdot : $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $(x, y) \mapsto x \cdot y = xy$, auf \mathbb{R} , genannt **Multiplikation**, 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: $\exists 1 \in \mathbb{R} \setminus \{0\} \forall x \in \mathbb{R} : x \cdot 1 = x$,
- M.iii) Inverses Element: $\forall x \in \mathbb{R} \setminus \{0\} \exists y \in \mathbb{R} : x \cdot y = 1$,
- M.iv) Kommutativität: $\forall x, y \in \mathbb{R} : x \cdot y = y \cdot x$.

Die Multiplikation ist mit der Addition verträglich im Sinne des **Distributivgesetzes**:

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

Bemerkung 2.2.2. i) $\forall x \in \mathbb{R} : x \cdot 0 = 0$; insbesondere ist zu jedem $x \neq 0$ das (eindeutig bestimmte) multiplikativ Inverse $x^{-1} \neq 0$.

ii) Falls $x \cdot y = 0$, so gilt $x = 0$ oder $y = 0$.

Das heisst, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$ bildet bezüglich der Multiplikation eine abelsche Gruppe mit dem neutralen Element 1.

Beweis. i) Für $x \in \mathbb{R}$ gilt

$$x \cdot 0 = x \cdot (0 + 0) \stackrel{D}{=} x \cdot 0 + x \cdot 0.$$

Addiere $-(x \cdot 0)$.

ii) Sei $x \neq 0$. Dann ist $x^{-1} \neq 0$ nach i). Falls $x \cdot y = 0$, so folgt mit i)

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

□

O) Schliesslich fordern wir die Existenz einer **totalen Ordnung** “ \leq ” im Sinne der Definitionen 1.4.2 und 1.4.4, konsistent mit den Operationen Addition und Multiplikation im Sinne der **Konsistenzaxiome**

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

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

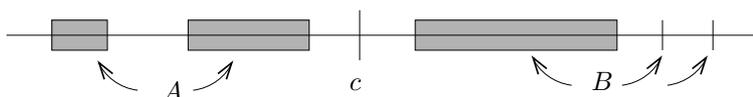
Die Axiome A.i)-iv), M.i)-iv), D, K.i),ii) gelten bereits in \mathbb{Q} . Die entscheidende zusätzliche Forderung an \mathbb{R} ist nun das **Vollständigkeitsaxiom**:

V) Der Zahlkörper \mathbb{R} ist **ordnungsvollständig**; das heisst, zu je zwei nicht leeren Mengen $A, B \subset \mathbb{R}$ mit

$$\forall a \in A, b \in B : a \leq b$$

gibt es eine Zahl $c \in \mathbb{R}$ mit

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



Beachte: Der Körper \mathbb{Q} ist nicht ordnungsvollständig; siehe Bemerkung 2.2.3.

Einige elementare Folgerungen aus den Axiomen:

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

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

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

iv) $0 < 1 < 2 < \dots;$

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

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

vii) Es gibt $c \in \mathbb{R}$ mit $c^2 = 2, c > 0$.

Beweis. i) Es gilt

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

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

ii) Mit $(-1) + 1 = 0$ folgt $1 = -(-1)$. Setze nun $x = -1$ in i).

iii) Sei $x \in \mathbb{R}$ beliebig gewählt. Da die Ordnung total ist, gilt $x \geq 0$ oder $x \leq 0$.

a) Falls $x \geq 0$, so folgt $x^2 \geq 0 \cdot x = 0$ mit K.ii).

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

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

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

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

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

“ \Rightarrow ”: Sei $y \geq x$, also $y - x \geq 0$. Da $y + x > 0$ folgt $y^2 \geq x^2$ mit K.ii).

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

vii) Gemäss iv) ist \mathbb{N} und damit auch \mathbb{Q} in \mathbb{R} enthalten. Setze

$$A = \{a \in \mathbb{Q}; 1 \leq a \leq a^2 < 2\}, B = \{b \in \mathbb{Q}; 1 \leq b \leq 2, b^2 \geq 2\}.$$

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

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

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

$$\forall a \in A, b \in B : a \leq c \leq b. \quad (2.2.1)$$

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

Behauptung 1. c ist eindeutig bestimmt.

Beweis. Nimm an, es existieren Zahlen $c_1 < c_2$ in \mathbb{R} , welche die Bedingung (2.2.1) erfüllen. Nach Bemerkung 2.1.1 gibt es dann auch derartige Zahlen $c_1, c_2 \in \mathbb{Q}$. Es folgt $c_0 = \frac{c_1 + c_2}{2} \in \mathbb{Q}$, und

$$\forall a \in A, b \in B : a \leq c_1 < c_0 = \frac{c_1 + c_2}{2} < c_2 \leq b.$$

Beachte, dass $c_0 \in A \cup B$. Falls $c_0 \in A$, so ist (2.2.1) für c_1 verletzt; falls $c_0 \in B$, so ist (2.2.1) für c_2 verletzt. Widerspruch. \square

Behauptung 2. $c^2 = 2$.

Beweis. Für alle $a \in A$, $b \in B$ gilt gemäss K.i):

$$2 - c^2 \leq b^2 - c^2 \leq b^2 - a^2 = (b - a) \underbrace{(b + a)}_{\leq 4} \leq 4(b - a).$$

Da man gemäss Behauptung 1 und Bemerkung 2.2.1 zu beliebig kleinem $d > 0$ Zahlen $a \in A$, $b \in B$ mit Abstand $b - a < d$ finden kann, folgt $c^2 \geq 2$. Analog sieht man, dass $c^2 \leq 2$. \square

Damit haben wir alle Folgerungen aus den Axiomen hergeleitet. \square

Bemerkung 2.2.3. Im Beweis von vii) gilt $A, B \subset \mathbb{Q}$; die Mengen A und B werden aber durch kein $c \in \mathbb{Q}$ getrennt. Wie wir gesehen haben, ist das die Mengen A, B trennende $c \in \mathbb{R}$ nämlich eindeutig bestimmt und erfüllt $c^2 = 2$. Nach Satz 2.1.1 gehört c 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

- i) $\forall x \in \mathbb{R} : x \leq |x|$;
- ii) $\forall x, y \in \mathbb{R} : |x \cdot y| = |x| \cdot |y|$.

Satz 2.2.1. (Dreiecks-Ungleichung). Es gilt

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

Beweis. Offenbar gelten für jedes $x \in \mathbb{R}$ die Beziehungen $x \leq |x|$, $-x \leq |x|$. Es folgt

$$x + y \leq |x| + |y|, \quad -(x + y) = (-x) + (-y) \leq |x| + |y|.$$

\square

Satz 2.2.2. (Young) Für alle $a, b \in \mathbb{R}$, $\delta > 0$ gilt

$$2|a \cdot b| \leq \delta a^2 + \frac{b^2}{\delta}.$$

Beweis. Es genügt, $a \geq 0$, $b \geq 0$ zu betrachten. Setze $\varepsilon = \sqrt{\delta} > 0$. Die Behauptung folgt aus

$$0 \leq \left(\varepsilon a - \frac{b}{\varepsilon}\right)^2 = \varepsilon^2 a^2 + \frac{b^2}{\varepsilon^2} - 2a \cdot b.$$

\square

2.2.2 Supremum und Infimum

Definition 2.2.2. Eine Menge $A \subset \mathbb{R}$ heisst **nach oben beschränkt**, falls eine Zahl $b \in \mathbb{R}$ existiert mit

$$\forall a \in A : a \leq b.$$

Jedes derartige b heisst eine **obere Schranke** für A .

(Analog: **nach unten beschränkt**, **untere Schranke**.)

Beispiel 2.2.1. i) 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.

ii) $\mathbb{N} = \{1, 2, 3, \dots\}$ ist nach unten beschränkt, jedoch nach oben unbeschränkt.

Sei nun $\emptyset \neq A \subset \mathbb{R}$ nach oben beschränkt. Dann gilt

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

und

$$\forall a \in A, b \in B : a \leq b.$$

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

$$\forall a \in A, b \in B : a \leq c \leq b. \quad (2.2.2)$$

Bemerkung 2.2.4. 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 , und c ist durch (2.2.2) eindeutig bestimmt.

Definition 2.2.3. Zu einer nach oben beschränkten Menge $\emptyset \neq A \subset \mathbb{R}$ heisst die durch (2.2.2) definierte Zahl $c =: \sup A$, das **Supremum** von A .

Wir fassen zusammen:

Satz 2.2.3. i) Jede nach oben beschränkte Menge $\emptyset \neq A \subset \mathbb{R}$ besitzt eine kleinste obere Schranke $c = \sup A$.

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

Beispiel 2.2.2. i) Sei $A =]-1, 1[\subset \mathbb{R}$. Dann gilt

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

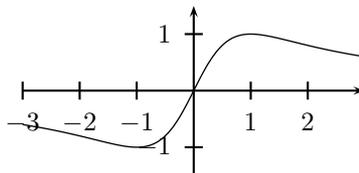
ii) $\mathbb{N} = \{1, 2, \dots\}$ ist nach oben unbeschränkt, besitzt also kein Supremum. Andererseits gilt für jedes $k \in \mathbb{N}$ offenbar $k \geq 1$. Da $1 \in \mathbb{N}$, ist 1 auch die grösste untere Schranke; das heisst,

$$\inf \mathbb{N} = 1.$$

iii) Sei A die Menge

$$A = \left\{ \frac{2x}{1+x^2}; x \in \mathbb{R} \right\},$$

veranschaulicht durch den Graphen der Funktion $x \mapsto 2x/(1+x^2)$:



Mit Satz 2.2.2 folgt

$$\forall x \in \mathbb{R} : 2x \leq 1 + x^2;$$

also $\sup A \leq 1$. Andererseits liefert der Vergleich mit $x = 1$ die Ungleichung $\sup A \geq 1$; also

$$\sup A = 1.$$

Beispiel 2.2.2.i) zeigt, 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

$$\sup A = \max A, \quad \text{das Maximum von } A.$$

Gehört analog die Zahl $\inf A$ zu A , so sagen wir, "das Infimum wird in A angenommen", und schreiben

$$\inf A = \min A, \quad \text{das Minimum von } A.$$

Beispiel 2.2.3. i) Sei $A = [-1, 1] = \{x \in \mathbb{R}; -1 \leq x \leq 1\}$. Dann gilt

$$\sup A = 1 = \max A, \quad \inf A = -1 = \min A.$$

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

$$\sup A = a_k = \max A, \quad \inf A = a_1 = \min A.$$

Falls die Zahl $\sup A$ oder $\inf A$ nicht angenommen wird, benötigen wir für das Rechnen mit \sup und \inf das folgende Lemma.

Lemma 2.2.1. Sei $\emptyset \neq A \subset \mathbb{R}$ nach oben beschränkt, und sei $\varepsilon > 0$ beliebig. Dann gibt es ein $a \in A$ mit

$$a \geq \sup A - \varepsilon.$$

Beweis (indirekt). Falls für ein $\varepsilon > 0$ gilt

$$\forall a \in A : a \leq \sup A - \varepsilon,$$

so ist die Zahl $\sup A - \varepsilon < \sup A$ obere Schranke für A . Jedoch ist $\sup A$ nach Definition die kleinste obere Schranke für A . Widerspruch! \square

Als Anwendung von Lemma 2.2.1 zeigen wir beispielhaft die folgende Aussage.

Satz 2.2.4. *Seien $A, B \subset \mathbb{R}$ nicht leer und nach oben beschränkt. Dann ist*

$$A + B = \{a + b; a \in A, b \in B\}$$

nicht leer und nach oben beschränkt, und es gilt

$$\sup(A + B) = \sup A + \sup B.$$

Offenbar ist die Aussage wahr, falls $\sup A$ und $\sup B$ angenommen werden. (Der Beweis ist in diesem Falle “trivial”.)

Beweis. i) Wir zeigen:

$$\sup(A + B) \leq \sup A + \sup B.$$

Nach Definition des Supremums gilt

$$\forall a \in A, b \in B : a \leq \sup A, b \leq \sup B;$$

also folgt mit K.i)

$$\forall a \in A, b \in B : a + b \leq \sup A + \sup B.$$

Das heisst, $\sup A + \sup B$ ist eine obere Schranke für $A + B$.

ii) Wir zeigen:

$$\sup A + \sup B \leq \sup(A + B).$$

Nimm widerspruchswise an,

$$0 < \sup A + \sup B - \sup(A + B) =: 3\varepsilon, \varepsilon > 0.$$

Wähle $a \in A, b \in B$ gemäss Lemma 2.2.1 mit

$$\sup A \leq a + \varepsilon, \sup B \leq b + \varepsilon.$$

Dann folgt mit

$$\begin{aligned} \sup A + \sup B &\leq a + b + 2\varepsilon \leq \sup(A + B) + 2\varepsilon \\ &= (\sup A + \sup B - 3\varepsilon) + 2\varepsilon < \sup A + \sup B \end{aligned}$$

der gewünschte Widerspruch. □

Analoge Sätze gelten für das Infimum.

Vereinbarung: Falls $\emptyset \neq A$ nach oben, bzw. nach unten unbeschränkt ist, setzen wir

$$\sup A = \infty, \text{ bzw. } \inf A = -\infty,$$

mit der Konvention $\infty + \infty = \infty$ sowie

$$\forall x \in \mathbb{R} : \infty + x = \infty.$$

2.3 Die natürlichen Zahlen

Gemäss Folgerung 2.2.1.iv) ist \mathbb{N} mit der natürlichen Ordnung in \mathbb{R} eingebettet.

Satz 2.3.1. (Archimedisches Prinzip) *Zu jedem $x \in \mathbb{R}$ gibt es eine natürliche Zahl n mit $x < n$. Insbesondere ist jedes $x \in \mathbb{R}$ endlich.*

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

$$\forall n \in \mathbb{N} : n \leq x.$$

Das heisst, x ist obere Schranke für \mathbb{N} . Nach dem Vollständigkeitsaxiom existiert $c = \sup \mathbb{N} \in \mathbb{R}$, und

$$\forall n \in \mathbb{N} : n \leq c.$$

Da mit n auch $n + 1$ zu \mathbb{N} gehört, folgt damit auch

$$\forall n \in \mathbb{N} : n + 1 \leq c;$$

also

$$\forall n \in \mathbb{N} : n \leq c - 1.$$

Das heisst, $c - 1 < c$ ist obere Schranke für \mathbb{N} , im Widerspruch zur Charakterisierung von c als *kleinste* obere Schranke für \mathbb{N} . \square

Korollar 2.3.1. *Zu jedem $\varepsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit*

$$0 < \frac{1}{n} < \varepsilon.$$

Beweis. Wähle $n \in \mathbb{N}$ mit $n > \varepsilon^{-1}$ gemäss Satz 2.3.1. Nach Multiplikation mit $\varepsilon n^{-1} > 0$ folgt mit K.ii) die Behauptung. \square

Prinzip der vollständigen Induktion: Offenbar bilden die natürlichen Zahlen \mathbb{N} die kleinste Menge $X \subset \mathbb{R}$ mit den Eigenschaften

- I.i) $1 \in X$,
- I.ii) $\forall x \in X : x + 1 \in X$.

Definition 2.3.1. *Eine Menge $X \subset \mathbb{R}$ mit I.i),ii) heisst **induktiv**.*

Es gilt also:

$$\mathbb{N} = \bigcap_{\substack{X \subset \mathbb{R} \\ X \text{ induktiv}}} X.$$

Satz 2.3.2. *Es seien $A(k)$ Aussagen, $k \in \mathbb{N}$. Es gelte:*

- i) $A(1)$ ist wahr (“Induktionsverankerung”)
- ii) $\forall k \in \mathbb{N} : A(k) \Rightarrow A(k + 1)$ (“Induktions-Schluss”)

Dann ist $A(k)$ wahr für alle $k \in \mathbb{N}$.

Beweis. Setze

$$X = \{k \in \mathbb{N}; A(k) \text{ ist wahr}\}.$$

Dann erfüllt X die Bedingungen I.i), ii); also $X = \mathbb{N}$. □

Wir illustrieren Satz 2.3.2 mit einigen Anwendungen.

Beispiel 2.3.1. i) Für beliebiges $n \in \mathbb{N}$ gilt

$$\sum_{\ell=1}^n (2\ell - 1) = 1 + 3 + \dots + (2n - 1) = n^2.$$

Beweis (vollständige Induktion). a) Induktionsverankerung: $n = 1$ ✓

b) Induktions-Schluss $n \mapsto n + 1$: Nach Induktions-Voraussetzung gelte für ein $n \in \mathbb{N}$ die Aussage

$$\sum_{\ell=1}^n (2\ell - 1) = n^2 \quad (\text{“Induktions-Annahme”}).$$

Es folgt:

$$\sum_{\ell=1}^{n+1} (2\ell - 1) = \sum_{\ell=1}^n (2\ell - 1) + (2(n+1) - 1) = n^2 + 2n + 1 = (n+1)^2,$$

wie gewünscht. □

ii) Für beliebiges $n \in \mathbb{N}$ gilt

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

Beweis (vollständige Induktion). a) Induktionsverankerung: $n = 1$ ✓

b) Induktions-Schluss $n \mapsto n + 1$. Die Aussage gelte für n . Dann folgt:

$$\begin{aligned} 1 + 2 + \dots + (n+1) &= (1 + 2 + \dots + n) + (n+1) \\ &= \frac{n(n+1)}{2} + (n+1) = \frac{n+2}{2}(n+1). \end{aligned}$$

□

iii) Für beliebiges $n \in \mathbb{N}$ gilt

$$(1 + 2 + \dots + n)^2 = 1 + 2^3 + \dots + n^3.$$

Beweis (vollständige Induktion). a) Induktionsverankerung: $n = 1$ ✓

b) Induktions-Schluss $n \mapsto n + 1$: Es gelte nach Induktions-Voraussetzung

$$\left(\sum_{\ell=1}^n \ell \right)^2 = (1 + 2 + \dots + n)^2 = 1 + 2^3 + \dots + n^3 = \sum_{\ell=1}^n \ell^3.$$

Dann folgt mit Beispiel ii):

$$\begin{aligned} \left(\sum_{\ell=1}^{n+1} \ell\right)^2 &= (1 + 2 + \dots + n + (n+1))^2 = \left(\left(\sum_{\ell=1}^n \ell\right) + (n+1)\right)^2 \\ &= \left(\sum_{\ell=1}^n \ell\right)^2 + (n+1)^2 + 2(n+1)\left(\sum_{\ell=1}^n \ell\right) \\ &= \sum_{\ell=1}^n \ell^3 + (n+1)^2 + 2(n+1)\frac{n(n+1)}{2} = \sum_{\ell=1}^n \ell^3 + (n+1)^3, \end{aligned}$$

wie gewünscht. □

iv) *Geometrische Reihe*. Sei $0 < q < 1$. Dann gilt für jedes $n \in \mathbb{N}$

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

Beweis (vollständige Induktion). a) Verankerung: $n = 0 \checkmark$, $n = 1 \checkmark$

b) Induktions-Schluss $n \mapsto n+1$: Die Aussage gelte für n . Dann folgt:

$$\begin{aligned} S_{n+1} = S_n + q^{n+1} &= \frac{1 - q^{n+1}}{1 - q} + q^{n+1} \\ &= \frac{1 - q^{n+1} + (1 - q)q^{n+1}}{1 - q} = \frac{1 - q^{n+2}}{1 - q}. \end{aligned}$$

□

v) *Bernoullische Ungleichung*. Sei $x > -1$. Dann gilt:

$$\forall n \in \mathbb{N} : (1 + x)^n \geq 1 + nx.$$

Beweis. a) $n = 1 \checkmark$

b) $n \mapsto n+1$: Nimm an, es gilt $(1 + x)^n \geq 1 + nx$. Rechne

$$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n \underbrace{(1 + x)}_{\geq 0} \stackrel{K.ii)}{\geq} (1 + nx)(1 + x) \\ &= 1 + (n+1)x + \underbrace{nx^2}_{\geq 0} \geq 1 + (n+1)x. \end{aligned}$$

□

2.4 Kardinalität

Gibt es “mehr” reelle als rationale Zahlen? Ebenso wie \mathbb{Q} gemäss Bemerkung 2.1.1 so liegen offenbar auch die irrationalen Zahlen “dicht” auf dem Zahlenstrahl, sogar bereits die Menge

$$\mathbb{Q} + \sqrt{2} = \{r + \sqrt{2}; r \in \mathbb{Q}\};$$

insbesondere gibt es also auch unendlich viele irrationale Zahlen. Wie kann man entscheiden, welche der Mengen \mathbb{Q} oder $\mathbb{R} \setminus \mathbb{Q}$ “grösser” ist?

Betrachten wir zunächst ein einfacheres Problem: Welche der Mengen \mathbb{N} oder \mathbb{Z} ist grösser? - Naiv würde man meinen, \mathbb{Z} wäre “doppelt so gross” wie \mathbb{N} ; jedoch kann man zum Beispiel mittels der Abbildung $f : \mathbb{N} \rightarrow \mathbb{Z}$ mit

$$n \mapsto \begin{cases} \frac{n-1}{2}, & \text{falls } n \text{ ungerade,} \\ -\frac{n}{2}, & \text{falls } n \text{ gerade,} \end{cases}$$

\mathbb{N} bijektiv auf \mathbb{Z} abbilden und so die Elemente von \mathbb{Z} “abzählen”.

Allgemein vereinbaren wir:

Definition 2.4.1. Zwei Mengen X und Y heissen **gleichmächtig**, falls es eine bijektive Abbildung $f : X \rightarrow Y$ gibt.

Beispiel 2.4.1. \mathbb{N} und \mathbb{Z} sind gleichmächtig.

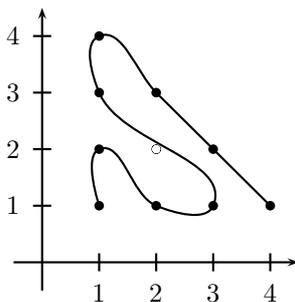
Definition 2.4.2. Eine Menge X heisst **abzählbar**, falls sie gleichmächtig ist zu \mathbb{N} , das heisst, falls eine bijektive Abbildung $f : X \rightarrow \mathbb{N}$ existiert.

Satz 2.4.1. \mathbb{Q} ist abzählbar.

Beweis (1. Cantorsches Diagonalverfahren). Stelle

$$\mathbb{Q}_+ = \{r \in \mathbb{Q}; r > 0\} = \left\{ \frac{p}{q}; p, q \in \mathbb{N} \text{ teilerfremd} \right\}$$

dar durch ganzzahlige Koordinatenpaare im 1. Quadranten, und durchlaufe diese Punkte wie in der nachstehenden Skizze angegeben:



Analog zur Abzählung von \mathbb{Q}_+ erhält man dann die Abzählung für ganz \mathbb{Q} . \square

Bemerkung 2.4.1. Offenbar kann man mit dem 1. Cantorschen Verfahren auch $\mathbb{N} \times \mathbb{N}$ abzählen. Induktion liefert, dass das Produkt endlich vieler abzählbarer Mengen wieder abzählbar ist:

$$\forall k \in \mathbb{N} : \mathbb{N} \times \underbrace{(\mathbb{N} \times \dots \times \mathbb{N})}_k \stackrel{\text{gleichmächtig}}{=} \mathbb{N} \times \mathbb{N} \stackrel{\text{gleichmächtig}}{=} \mathbb{N}.$$

Satz 2.4.2. \mathbb{R} ist nicht abzählbar.

Der Beweis benötigt etwas Vorbereitung.

Lemma 2.4.1. $\mathcal{P}(\mathbb{N})$ ist nicht abzählbar.

Beweis. Wir ordnen jeder Teilmenge $S \subset \mathbb{N}$ eindeutig eine “binäre Folge” $f^S = (f_1^S, f_2^S, \dots)$ zu, das heisst, eine Abbildung $f^S = (f_1^S, f_2^S, \dots) : \mathbb{N} \rightarrow \{0, 1\}$ mit

$$f_k^S = \begin{cases} 1, & \text{falls } k \in S \\ 0, & \text{falls } k \notin S \end{cases} \quad (2.4.1)$$

Der “Raum” dieser binären Folgen wird mit dem Symbol $2^{\mathbb{N}}$ bezeichnet. Somit gilt:

$$\mathcal{P}(\mathbb{N}) \stackrel{\text{gleichmächtig}}{=} 2^{\mathbb{N}}.$$

Behauptung. $2^{\mathbb{N}}$ ist nicht abzählbar.

Beweis (indirekt). Sei widerspruchswise $f : \mathbb{N} \rightarrow 2^{\mathbb{N}}$ eine Abzählung mit

$$n \mapsto f(n) = (f_1(n), f_2(n), \dots), \quad n \in \mathbb{N}.$$

Definiere die Folge $a = (a_1, a_2, \dots) : \mathbb{N} \rightarrow \{0, 1\}$ mit

$$a_k = f_k(k) + 1 \pmod{2}, \quad k \in \mathbb{N}.$$

Für jedes $k \in \mathbb{N}$ gilt dann $a_k \neq f_k(k)$, also $a \neq f(k)$. Daher ist f nicht surjektiv im Widerspruch zu unserer Annahme. \square

Den Beweis können wir uns gut veranschaulichen mit dem **2. Cantorschen Diagonalverfahren**:

$f_\ell(k)$	ℓ	\rightarrow			
k	$f_1(1)$	$f_2(1)$	$f_3(1)$	\dots	
\downarrow	$f_1(2)$	$f_2(2)$	$f_3(2)$	\dots	
	$f_1(3)$	$f_2(3)$	$f_3(3)$		
	\vdots	\vdots	\vdots		

Die im Beweis konstruierte Folge a unterscheidet sich für jedes $k \in \mathbb{N}$ von $f(k)$ an der k -ten Stelle. \square

Beweis (von Satz 2.4.2). Die Abbildung

$$f : \mathcal{P}(\mathbb{N}) \rightarrow \mathbb{R}, \quad S \mapsto x = 0, f_1^S, f_2^S, \dots \quad (\text{unendlicher Dezimalbruch})$$

gemäss (2.4.1) ist wohldefiniert und injektiv. Die Bildmenge $X = f(\mathcal{P}(\mathbb{N})) \subset \mathbb{R}$ ist nach Lemma 2.4.1 nicht abzählbar; also ist auch \mathbb{R} nicht abzählbar. \square

Bemerkung 2.4.2. Fassen wir die Bilder der im Beweis von Satz 2.4.2 konstruierten Abbildung auf als binäre Darstellungen von Zahlen

$$a = f^S = \sum_{\ell=1}^{\infty} f_\ell^S 2^{-\ell} \in [0, 1] \subset \mathbb{R},$$

so erkennen wir, dass $\mathcal{P}(\mathbb{N})$ auch mindestens so mächtig ist wie die Menge $[0, 1]$, und damit mindestens so mächtig wie \mathbb{R} (Übung); $\mathcal{P}(\mathbb{N})$ und \mathbb{R} sind demnach gleichmächtig.

Für allgemeine Mengen X gilt:

Satz 2.4.3. Für jede Menge X ist $\mathcal{P}(X)$ “mächtiger” als X ; das heisst, es gibt keine surjektive Abbildung $f : X \rightarrow \mathcal{P}(X)$.

Beweis (indirekt). Nimm an, es existiert eine surjektive Abbildung $f : X \rightarrow \mathcal{P}(X)$. Betrachte die Menge

$$A = \{x \in X; x \notin f(x)\}$$

Da f nach Annahme surjektiv, gibt es ein $x_0 \in X$ mit $A = f(x_0)$. Wir führen diese Aussage zum Widerspruch.

a) Nehmen wir an $x_0 \in A$, so folgt nach Definition von A , dass $x_0 \notin f(x_0) = A$.

b) Also muss gelten $x_0 \notin A = f(x_0)$. In diesem Falle ergibt die Definition von A jedoch, dass gilt $x_0 \in A$. Widerspruch.

□

Kontinuumshypothese. Ausgehend von $X_0 = \mathbb{N}$ erhält man also Mengen

$$X_1 = \mathcal{P}(\mathbb{N}), X_2 = \mathcal{P}(X_1), \dots, X_{k+1} = \mathcal{P}(X_k), k \in \mathbb{N}$$

strikt aufsteigender Mächtigkeit. Die Mächtigkeit von \mathbb{N} bezeichnet man mit \aleph_0 (“Aleph 0”). Die Frage, ob es “zwischen” den natürlichen Zahlen \mathbb{N} und dem “Kontinuum” $\mathcal{P}(\mathbb{N}) = \mathbb{R}$ noch eine weitere Menge anderer Mächtigkeit gibt, ob also eine überabzählbare Teilmenge der reellen Zahlen existiert, die in ihrer Mächtigkeit kleiner ist als \mathbb{R} , hat die Mathematik lange beschäftigt. Schliesslich gelang es Gödel (1937) und Cohen (1964) zu zeigen, dass die Frage nicht aus den Axiomen entscheidbar ist. Vergleiche dazu Davis-Hersch: *Erfahrung Mathematik*, S.336.

2.5 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 2.5.1. Den Graphen einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ 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 2.5.2. i) Die Bewegung eines Massepunktes kann man durch dessen Orts- und Geschwindigkeitsvektor beschreiben.

ii) Eine Schar von N Massepunkten (Atome in einem Gas, Planeten im Sonnensystem) können wir gleichzeitig mit ihrem jeweiligen Ort $x^{(i)} = (x_1^{(i)}, x_2^{(i)}, x_3^{(i)})$, $1 \leq i \leq N$, 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**

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

mit komponentenweiser **Addition** $+$: $\mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$

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

Dann ist \mathbb{R}^n bezüglich der Addition eine abelsche Gruppe mit neutralem Element

$$0 = (0, \dots, 0) \in \mathbb{R}^n : \text{der Nullvektor.}$$

Zusätzlich kann man Vektoren $x \in \mathbb{R}^n$ mit einem Skalierungsfaktor (Skalar) $\alpha \in \mathbb{R}$ strecken:

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

Für diese **Skalarmultiplikation** $\mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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: $1 \cdot x = x$,

wobei $x, y \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ beliebig sind.

Somit hat der Raum \mathbb{R}^n die Struktur eines **Vektorraums** mit dem **Skalkörper** \mathbb{R} , eines **\mathbb{R} -Vektorraums**; vergleiche die Vorlesung "*Lineare Algebra*".

Kanonische Basis: Bezüglich der **Standardbasis**

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

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

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

darstellen.

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

$$x \cdot y := \langle x, y \rangle := x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i \in \mathbb{R}.$$

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 $x, y \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$. Zur besseren Unterscheidung lassen wir hier und im folgenden meist den Multiplikationspunkt bei der Skalarmultiplikation und bei der Multiplikation reeller Zahlen weg.

Beispiel 2.5.3. i) Für $x = (2, 0, 3)$, $y = (-3, 1, 2)$ gilt

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

das heisst, x und y stehen **senkrecht** aufeinander, sie sind **orthogonal**, $x \perp y$.

ii) Dies gilt auch für verschiedene Standardbasisvektoren

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

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

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

Beispiel 2.5.4. i) Es gilt $\|e_i\| = 1$, $1 \leq i \leq n$. Die Standardbasisvektoren sind also paarweise **orthogonal** und auf Länge 1 **normiert**; sie sind **orthonormal**.

ii) *Satz von Pythagoras.* Falls $x, y \in \mathbb{R}^n$ orthogonal mit $x \cdot y = 0$, so folgt

$$\|x + y\|^2 = (x + y) \cdot (x + y) = x \cdot x + 2x \cdot y + y \cdot y = \|x\|^2 + \|y\|^2.$$

iii) Der Abstand des Punktes (x_1, x_2) vom Nullpunkt ist

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

konsistent mit der elementargeometrischen Interpretation.

iv) Die Diagonale im Einheitsquadrat hat 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 2.5.1. (Cauchy-Schwarz) Für alle $x, y \in \mathbb{R}^n$ gilt

$$|x \cdot y| \leq \|x\| \cdot \|y\|.$$

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

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

□

Wir können Satz 2.5.1 auch geometrisch herleiten: OBdA sei $x \neq 0$. Zerlege y orthogonal

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

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

Satz 2.5.2. Die euklidische Norm hat die Eigenschaften

- i) *Definitheit:* $\forall x \in \mathbb{R}^n : \|x\| \geq 0$, und $\|x\| = 0 \Rightarrow x = 0$,
- 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) Mit Satz 2.5.1 schätzen wir ab

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

Die Behauptung folgt mit Folgerung 2.2.1.vi). □

Beispiel 2.5.5. Für $x = (3, 4) = 3e_1 + 4e_2 \in \mathbb{R}^2$ gilt

$$5 = \|x\| \leq \|3e_1\| + \|4e_2\| = 7.$$

Euklidische Metrik: Via $\|\cdot\|$ können wir schliesslich den *Abstand* zweier durch die “Ortsvektoren” $x, y \in \mathbb{R}^n$ gegebener Punkte einführen als:

$$d(x, y) = \|x - y\| \quad (d: \text{“Distanz”}).$$

Aus Satz 2.5.2 folgen die Eigenschaften:

- i) Positive Definitheit: $\forall x, y \in \mathbb{R}^n : d(x, y) \geq 0$, und $d(x, y) = 0 \Leftrightarrow x = y$;

- ii) Symmetrie: $\forall x, y \in \mathbb{R}^n : d(x, y) = d(y, x)$;
- iii) Dreiecksungleichung: $\forall x, y, z \in \mathbb{R}^n : d(x, z) \leq d(x, y) + d(y, z)$.

Beweis. i) und ii) sind offensichtlich.

iii) Schreibe $x - z = (x - y) + (y - z)$ und benutze Satz 2.5.2.iii). \square

2.6 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

$$(a, b) \cdot \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) = (1, 0); \quad (2.6.1)$$

das heisst,

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

ist zu (a, b) invers.

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

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

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

Bemerkung 2.6.1. 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 Skalarmultiplikation in \mathbb{R}^2 , denn

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

In diesem Sinne können wir den Standardbasisvektor $e_1 = (1, 0) \in \mathbb{R}^2$ mit der Zahl $1 \in \mathbb{R}$ "identifizieren". Weiter setzen wir

$$e_2 = (0, 1) =: i, \quad \text{die imaginäre Einheit,}$$

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** $x = \operatorname{Re}(z)$ und **Imaginärteil** $y = \operatorname{Im}(z)$.

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

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

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

$$z \cdot \bar{z} = (x + iy) \cdot (x - iy) = x^2 - i^2 y^2 = x^2 + y^2 = \|z\|^2. \quad (2.6.2)$$

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

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2, \quad \overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2. \quad (2.6.3)$$

Beweis. Rechne nach für $z_k = x_k + iy_k \in \mathbb{C}$, $k = 1, 2$:

$$\overline{z_1 z_2} = \overline{(x_1 + iy_1) \cdot (x_2 + iy_2)} = \overline{(x_1 x_2 - y_1 y_2) - i(x_1 y_2 + x_2 y_1)} = \bar{z}_1 \cdot \bar{z}_2.$$

□

Folgerung 2.6.1. i) Mit (2.6.2) folgt

$$\forall z \in \mathbb{C} \setminus \{0\} : z^{-1} = \frac{\bar{z}}{\|z\|^2},$$

in Übereinstimmung mit (2.6.1).

ii) Mit (2.6.2) erhalten wir

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

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

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

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

Beispiel 2.6.1. i)

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

ii)

$$\frac{2 + i}{2 - i} = \frac{2 + i}{2 - i} \cdot \frac{2 + i}{2 + i} = \frac{4 - 1 + 4i}{5} = \frac{3}{5} + \frac{4}{5}i.$$

Bemerkung 2.6.2. i) Später werden wir noch die Polarform $z = re^{i\phi}$ einer Zahl $z = x + iy \in \mathbb{C}$ mit $r = |z|$ und $x = r \cos \phi$, $y = r \sin \phi$ kennenlernen, die sich für Rechnungen besonders gut eignet.

ii) Nicht nur die in \mathbb{R} nicht lösbare Gleichung $z^2 + 1 = 0$ hat in \mathbb{C} Lösungen $z = \pm i$. Allgemein 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. Das heisst, \mathbb{C} ist im Unterschied zu \mathbb{R} **algebraisch vollständig**. Den Beweis müssen wir jedoch auf später verschieben.

iii) Leider gibt es für \mathbb{C} jedoch keine mit den Körperoperationen verträgliche **Ordnung**; sonst wäre gemäss Folgerung 2.2.1.iii)

$$-1 = i^2 \geq 0.$$

im Widerspruch zu Folgerung 2.2.1.iv)

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,

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

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

$$e = 2.718\dots,$$

den “Limes der kontinuierlichen Verzinsung”.

iii) Die **geometrische Reihe**

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

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 $n \rightarrow \infty$, falls gilt

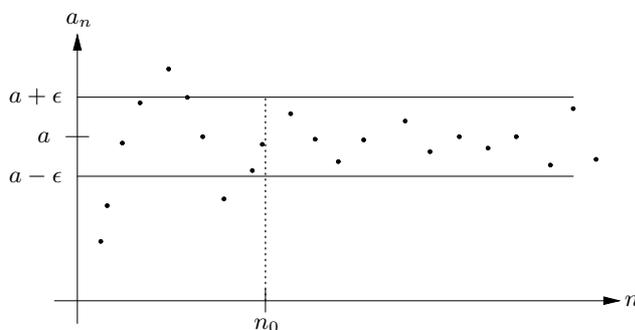
$$\forall \epsilon > 0 \exists n_0 = n_0(\epsilon) \in \mathbb{N} \forall n \geq n_0 : |a_n - a| < \epsilon;$$

das heisst, 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

$$a = \lim_{n \rightarrow \infty} a_n \quad \text{oder} \quad a_n \rightarrow a \quad (n \rightarrow \infty)$$

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 $a_n \rightarrow 0$ ($n \rightarrow \infty$).

Beweis. Nach Korollar 2.3.1 gibt es zu jedem $\epsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $n_0 > \frac{1}{\epsilon}$; das heisst, $\frac{1}{n_0} < \epsilon$. Es folgt

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

wie gewünscht. □

ii) Sei $q \in \mathbb{R}$ mit $0 < q < 1$. Dann gilt $q^n \rightarrow 0$ ($n \rightarrow \infty$).

Beweis. Schreibe $\frac{1}{q} = 1 + \delta$ mit $\delta > 0$. Mit der Bernoullischen Ungleichung $(1 + x)^n \geq 1 + nx$ für alle $x \geq -1$, $n \in \mathbb{N}$ aus Beispiel 2.3.1.v) folgt

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

also

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

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

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

□

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

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

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

somit folgt

$$\forall a, b > 0, n \in \mathbb{N} : a \geq b \Leftrightarrow a^n \geq b^n. \quad (3.2.1)$$

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

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

falls n so gross gewählt, dass

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

Wähle

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

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 (3.2.1) folgt

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

das heisst,

$$\forall n \geq n_0 : |\sqrt[n]{n} - 1| < \epsilon.$$

□

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

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

das heisst, 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

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

Damit erhalten wir

$$\forall n \geq n_0 : 0 < a_n = \left(\frac{\sqrt[n]{n}}{(1 + \epsilon)^2} \right)^{np} < \left(\frac{1}{1 + \epsilon} \right)^{pn} = r^n$$

mit

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

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

Beispiel 3.2.2. Nicht jede Folge $(a_n)_{n \in \mathbb{N}}$ konvergiert, wie die folgenden Beispiele zeigen.

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

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

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

iii) Fibonacci-Zahlen: Mit $F_0 = 1, F_1 = 1, F_{n+1} = F_n + F_{n-1}$ ($n \geq 2$) folgt durch Induktion $F_n \geq n$, $n \in \mathbb{N}$, und analog zu ii) kann kein $a \in \mathbb{R}$ Grenzwert von $(F_n)_{n \in \mathbb{N}}$ sein.

Rechnen mit Grenzwerten: Mit Grenzwerten kann man rechnen, selbst wenn man den genauen Wert nicht kennt.

Satz 3.2.1. Sei $(a_n)_{n \in \mathbb{N}}$ sowohl gegen $a \in \mathbb{R}$ als auch gegen $b \in \mathbb{R}$ konvergent. Dann folgt $a = b$.

Beweis (indirekt). Annahme: $a \neq b$. Wähle $\epsilon = \frac{|b-a|}{2} > 0$ und dazu $n_0 = n_0(\epsilon)$ mit

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

Dann folgt für $n \geq n_0$ mit der Dreiecksungleichung:

$$2\epsilon = |a - b| = |(a - a_n) - (b - a_n)| \leq |a - a_n| + |b - a_n| < 2\epsilon.$$

Widerspruch! \square

Satz 3.2.2. Seien die Folgen $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ konvergent mit $\lim_{n \rightarrow \infty} a_n = a$, $\lim_{n \rightarrow \infty} b_n = b$. Dann konvergieren die Folgen $(a_n + b_n)_{n \in \mathbb{N}}, (a_n \cdot b_n)_{n \in \mathbb{N}}$, und

- i) $\lim_{n \rightarrow \infty} (a_n + b_n) = a + b = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n,$
 ii) $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = a \cdot b = \lim_{n \rightarrow \infty} a_n \cdot \lim_{n \rightarrow \infty} b_n.$
 iii) Falls zusätzlich $b \neq 0 \neq b_n$ für alle n , so gilt auch $\lim_{n \rightarrow \infty} (a_n/b_n) = a/b.$
 iv) Falls $a_n \leq b_n$ für $n \in \mathbb{N}$, so auch $a \leq b.$

Bemerkung 3.2.1. Falls $a_n < b_n, n \in \mathbb{N}$, so gilt im Allgemeinen nicht $a < b$. Betrachte dazu das Beispiel

$$a_n := 0 < \frac{1}{n} =: b_n, n \in \mathbb{N}, \text{ mit } \lim_{n \rightarrow \infty} a_n = 0 = \lim_{n \rightarrow \infty} b_n.$$

Beweis von Satz 3.2.2. Zu $\epsilon > 0$ sei im folgenden $n_0 = n_0(\epsilon) \in \mathbb{N}$ stets so gewählt, dass

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

i) Schätze ab

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

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

ii) OBdA sei $\epsilon < 1$. Damit erhalten wir zunächst

$$\forall n \geq n_0 : |b_n| \leq |b_n - b| + |b| \leq |b| + 1$$

und weiter

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

Da wiederum $\epsilon > 0$ 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

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

Es folgt

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

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

iv) (indirekt). Falls $a > b$, so folgt bei Wahl von $a - b =: 2\epsilon > 0$ aus der Ungleichungskette

$$\forall n \geq n_0 : a < a_n + \epsilon \leq b_n + \epsilon < b + 2\epsilon = a$$

der gewünschte Widerspruch. \square

Beispiel 3.2.3. Es gilt

$$a_n = \frac{3n^4 - 7n^3 + 5}{2n^4 + 6n^2 + 3n} = \frac{3 - 7n^{-1} + 5n^{-4}}{2 + 6n^{-2} + 3n^{-3}} \rightarrow \frac{3}{2} \quad (n \rightarrow \infty).$$

3.3 Monotone Konvergenz

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

Definition 3.3.1. $(a_n)_{n \in \mathbb{N}}$ heisst nach oben (unten) beschränkt, falls gilt

$$\exists b \in \mathbb{R}, \forall n \in \mathbb{N} : a_n \leq b \quad (\text{bzw. } b \leq a_n);$$

das heisst, falls die Menge $A = \{a_n; n \in \mathbb{N}\}$ nach oben (unten) beschränkt ist.

Satz 3.3.1. Falls $(a_n)_{n \in \mathbb{N}}$ konvergent ist, dann ist $(a_n)_{n \in \mathbb{N}}$ beschränkt.

Beweis. Sei $a = \lim_{n \rightarrow \infty} a_n$, und zu $\epsilon = 1$ sei $n_0 \in \mathbb{N}$ so gewählt, dass $|a_n - a| < 1$ für $n \geq n_0$. Für $n \geq n_0$ gilt dann

$$|a_n| = |a_n - a + a| \leq |a| + |a_n - a| \leq |a| + 1.$$

Es folgt

$$\forall n \in \mathbb{N} : |a_n| \leq \max\{|a| + 1, |a_1|, |a_2|, \dots, |a_{n_0}|\}.$$

□

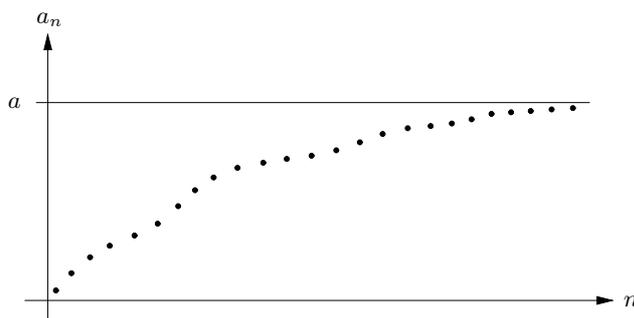
Beschränktheit ist also notwendig, jedoch nicht hinreichend für Konvergenz, wie das Beispiel der Folge $a_n = (-1)^n$, $n \in \mathbb{N}$, zeigt.

Satz 3.3.2. Sei $(a_n)_{n \in \mathbb{N}}$ nach oben beschränkt und **monoton wachsend**, das heisst, mit einer Zahl $b \in \mathbb{R}$ gilt

$$\forall n \in \mathbb{N} : a_1 \leq a_2 \leq \dots \leq a_n \leq a_{n+1} \leq \dots \leq b.$$

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

Analog, falls $(a_n)_{n \in \mathbb{N}}$ nach unten beschränkt und **monoton fallend**.



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 2.2.3.

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

Beweis. Sei $\epsilon > 0$ beliebig vorgegeben. Gemäss Lemma 2.2.1 gibt es $n_0 = n_0(\epsilon) \in \mathbb{N}$ mit $a_{n_0} > a - \epsilon$. Monotonie ergibt

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

also

$$\forall n \geq n_0 : |a_n - a| < \epsilon, \quad \forall n \geq n_0,$$

wie gewünscht. □

Beispiel 3.3.1. i) Jeder unendliche Dezimalbruch

$$x = x_0.x_1 \dots x_k \dots, \text{ mit } x_k \in \{0, \dots, 9\}, k \in \mathbb{N},$$

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

$$a_n = x_0.x_1 \dots x_n, n \in \mathbb{N}.$$

ii) *Eulersche Zahl:* Betrachte die Folgen

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

Behauptung. Es gilt

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

Beweis. Wir schätzen ab

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

wobei wir im vorletzten 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

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

Die Behauptung folgt. □

Gemäss Satz 3.3.2 und Satz 3.2.2 existieren $a, b \in \mathbb{R}$ mit

$$2 \leq a = \lim_{n \rightarrow \infty} a_n = \sup_{n \in \mathbb{N}} a_n \leq b = \lim_{n \rightarrow \infty} b_n = \inf_{n \in \mathbb{N}} b_n \leq 4.$$

Weiter gilt nach Satz 3.2.2

$$1 \leq \frac{b}{a} = \lim_{n \rightarrow \infty} \frac{b_n}{a_n} = 1;$$

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

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

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

Behauptung. $\lim_{n \rightarrow \infty} a_n = a$ existiert, und $a^2 = c$, $a > 0$.

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

$$\forall n \in \mathbb{N} : a_{n+1}^2 = \left(a_n + \frac{c - a_n^2}{2a_n} \right)^2 = a_n^2 + (c - a_n^2) + \underbrace{\left(\frac{c - a_n^2}{2a_n} \right)^2}_{\geq 0} \geq c; \quad (3.3.1)$$

insbesondere folgt somit auch $a_{n+1} \leq a_n$, $n \in \mathbb{N}$. Induktion liefert zudem $a_n > 0$ und daher mit (3.3.1) sogar die Ungleichung $a_n \geq 1$ für jedes n . Somit erhalten wir

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

und gemäss Satz 3.3.2 existiert $a = \lim_{n \rightarrow \infty} a_n$. Mit Satz 3.2.2 folgt schliesslich

$$1 \leq a = \lim_{n \rightarrow \infty} a_{n+1} = \lim_{n \rightarrow \infty} \frac{1}{2} \left(a_n + \frac{c}{a_n} \right) = \frac{1}{2} \left(a + \frac{c}{a} \right),$$

und a erfüllt die Gleichung $a^2 = c$, wie gewünscht. \square

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 $a_n = (-1)^n$, $n \in \mathbb{N}$, hat die konstanten Teilfolgen $(a_{2n})_{n \in \mathbb{N}}$, bzw. $(a_{2n+1})_{n \in \mathbb{N}}$.

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; das heisst, falls eine unendliche Teilmenge $\Lambda \subset \mathbb{N}$ existiert mit

$$a_l \rightarrow a \quad (l \rightarrow \infty, l \in \Lambda); \quad \text{das heisst: } a = \lim_{l \rightarrow \infty, l \in \Lambda} a_l.$$

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

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

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

Beweis. i) Falls $a_l \rightarrow a$ ($l \rightarrow \infty$, $l \in \Lambda$), so existiert offenbar zu jedem $\epsilon > 0$ ein $l_0 \in \Lambda$ mit

$$\forall l \geq l_0, l \in \Lambda : |a - a_l| < \epsilon.$$

Ist nun zusätzlich $n_0 \in \mathbb{N}$ vorgegeben, so wähle $l \in \Lambda$ beliebig mit $l \geq \max\{n_0, l_0\}$. Es folgt $|a - a_l| < \epsilon$, und $l \geq n_0$.

ii) Umgekehrt wähle $l(1) = 1$ und definiere $(l(n))_{n \in \mathbb{N}}$ induktiv, wie folgt. Sei $n \in \mathbb{N}$, und seien $l(1) < l(2) < \dots < l(n)$ bereits bestimmt. Es gilt dann automatisch $l(j) \geq j$, $1 \leq j \leq n$. Zu $\epsilon = \frac{1}{n} > 0$ und $n_0 = l(n) + 1$ existiert nach Annahme ein Index $l(n+1) := l \geq n_0 > l(n)$ mit $|a - a_l| < 1/n$. Die so konstruierte Teilfolge $(a_{l(n)})_{n \in \mathbb{N}}$ konvergiert offenbar gegen a . \square

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

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

Für $k \in \mathbb{N}$ existieren dann nach Satz 2.2.3

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

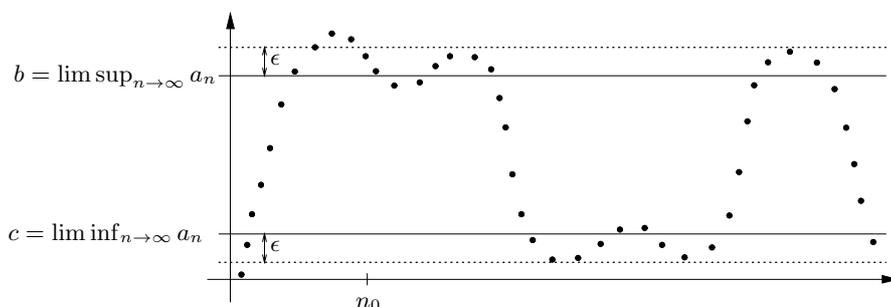
Offenbar gilt

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

Mit Satz 3.3.2 folgt die Existenz von

$$\begin{aligned} b &= \lim_{k \rightarrow \infty} b_k =: \limsup_{n \rightarrow \infty} a_n && \text{("Limes superior")}, \\ c &= \lim_{k \rightarrow \infty} c_k =: \liminf_{n \rightarrow \infty} a_n && \text{("Limes inferior")}, \end{aligned}$$

und $c \leq b$ wegen Satz 3.2.2.



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 $b = \lim_{k \rightarrow \infty} b_k$, gibt es $k_0 = k_0(\epsilon)$ mit

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

das heisst,

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

OBdA sei $k_0 \geq n_0$. (Ersetze sonst k_0 durch n_0 .) Fixiere $k = k_0$. Offenbar gilt für jedes $l \geq k_0$ die Abschätzung

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

Gemäss Lemma 2.2.1 gibt es zudem ein $l \geq k_0 \geq n_0$ mit

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

also

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

Die Behauptung folgt mit Satz 3.4.1.

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

Es folgt

Satz 3.4.2. (Bolzano Weierstrass) *Jede beschränkte Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{R} besitzt eine konvergente Teilfolge, also auch einen Häufungspunkt.*

Bemerkung 3.4.1. 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

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

Insbesondere für $k = k_0$ folgt dann

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

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

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

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

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

iv) Umgekehrt konvergiert jede Teilfolge einer Folge $a_n \rightarrow a$ ($n \rightarrow \infty$) ebenfalls gegen a , und es gilt $a = b = c$.

Satz 3.4.3. Sei $(a_n)_{n \in \mathbb{N}}$ beschränkt, $a \in \mathbb{R}$. Es sind äquivalent:

- i) $a_n \rightarrow a$ ($n \rightarrow \infty$);
- ii) $\liminf_{k \rightarrow \infty} a_k = \limsup_{k \rightarrow \infty} a_k = a$;
- iii) Jede Teilfolge von $(a_n)_{n \in \mathbb{N}}$ besitzt eine Teilfolge $(a_\ell)_{\ell \in \Lambda}$ mit

$$a_\ell \rightarrow a \quad (\ell \rightarrow \infty, \ell \in \Lambda).$$

Beweis. i) \Rightarrow iii) Dies ist offensichtlich.

iii) \Rightarrow ii) Nach Annahme iii) hat insbesondere die gemäss Lemma 3.4.1 gegen $\limsup_{k \rightarrow \infty} a_k$ konvergente Teilfolge den Limes a ; analog $\liminf_{k \rightarrow \infty} a_k = a$.

ii) \Rightarrow i) Dies ist Bemerkung 3.4.1.iii). \square

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

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

für beliebiges $n \geq 3$ folgt

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

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 ist, existieren

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

Mit Satz 3.2.2 folgt $1 \leq a, b \leq 2$; weiter erhalten wir

$$a = \lim_{n \rightarrow \infty} g_{2n+1} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{g_{2n}}\right) = 1 + \frac{1}{b},$$

$$b = \lim_{n \rightarrow \infty} g_{2n} = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{g_{2n-1}}\right) = 1 + \frac{1}{a}.$$

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.

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

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

und $(g_n)_{n \in \mathbb{N}}$ ist konvergent gemäss Satz 3.4.3 mit $\lim_{n \rightarrow \infty} g_n = g$.

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

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

Satz 3.5.1. *Konvergente Folgen sind Cauchy-Folgen.*

Beweis. Es gelte $a_n \rightarrow a$ ($n \rightarrow \infty$). Zu $\epsilon > 0$ wähle $n_0 = n_0(\epsilon)$ mit

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

Dann gilt

$$\forall \ell, n \geq n_0 : |a_n - a_\ell| \leq |a_n - a| + |a - a_\ell| < 2\epsilon.$$

□

Gilt auch die Umkehrung?

Satz 3.5.2. *Cauchy-Folgen in \mathbb{R} sind konvergent.*

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

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

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

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

Fixiere $n = n_0$. Dann folgt

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

also

$$\forall l \in \mathbb{N} : |a_l| < \max\{|a_1|, \dots, |a_{n_0-1}|, |a_{n_0}| + 1\}.$$

□

Nach Satz 3.4.2 (Bolzano-Weierstrass) gibt es $\Lambda \subset \mathbb{N}$, $a \in \mathbb{R}$ mit

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

Zu $\epsilon > 0$ wähle $l_0 = l_0(\epsilon)$ mit

$$\forall l \geq l_0, l \in \Lambda : |a_l - a| < \epsilon$$

und weiter $n_0 = n_0(\epsilon)$ mit

$$\forall l, n \geq n_0 : |a_l - a_n| < \epsilon.$$

Für beliebiges $l \in \Lambda$ mit $l \geq \max\{l_0, n_0\}$ erhalten wir so

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

das heisst,

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

□

Gemäss Satz 3.5.1 und 3.5.2 ist das Cauchy-Kriterium (3.5.1) notwendig und hinreichend für die Konvergenz einer Folge in \mathbb{R} , wobei man den möglichen Grenzwert nicht bereits kennen muss.

Beispiel 3.5.1. i) Setzen wir $a_1 = 1$, $a_n = a_{n-1} + \frac{1}{n}$, $n \geq 2$, so erhalten wir die *harmonische Reihe* mit

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

Die Folge $(a_n)_{n \in \mathbb{N}}$ ist divergent, da für jedes $n \in \mathbb{N}$ gilt

$$a_{2n} - a_n = \frac{1}{n+1} + \dots + \frac{1}{2n} \geq \frac{n}{2n} = \frac{1}{2}.$$

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

ii) Setzen wir $a_1 = 1$, $a_n = a_{n-1} + \frac{(-1)^{n-1}}{n}$, $n \geq 2$, so erhalten wir die *alternierende harmonische Reihe* mit

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

Da für jedes $k \in \mathbb{N}$ gilt

$$a_{2k} - a_{2k-2} = \frac{1}{2k-1} - \frac{1}{2k} > 0 > \frac{1}{2k+1} - \frac{1}{2k} = a_{2k+1} - a_{2k-1}$$

sowie $a_{2k} < a_{2k+1}$, erhalten wir

$$\forall k \in \mathbb{N} : \frac{1}{2} = a_2 < a_{2k-2} < a_{2k} < a_{2k+1} < a_{2k-1} < a_1 = 1, \quad (3.5.2)$$

und die Teilfolgen $(a_{2k})_{k \in \mathbb{N}}$, bzw. $(a_{2k+1})_{k \in \mathbb{N}}$ sind nach Satz 3.3.2 monoton konvergent. Da

$$\forall n \in \mathbb{N} : |a_n - a_{n+1}| = \frac{1}{n+1}, \quad (3.5.3)$$

haben sie zudem denselben Limes a , und $(a_n)_{n \in \mathbb{N}}$ konvergiert nach Satz 3.4.3. Schliesslich erhalten wir mit (3.5.2) und (3.5.3) noch die Fehlerabschätzung

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

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

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

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

Bemerkung 3.5.1. Satz 3.5.2 benutzt entscheidend die Vollständigkeit von \mathbb{R} . Wir können das Cauchy-Kriterium (3.5.1) jedoch auch aus anderer Perspektive betrachten.

Metrische Vollständigkeit. Sei M eine beliebige Menge.

Definition 3.5.2. Ein $d : M \times M \rightarrow \mathbb{R}$ heisst **Metrik** auf M , falls gilt:

- i) $\forall x, y \in M : d(x, y) \geq 0, \quad d(x, y) = 0 \Leftrightarrow x = y;$
- ii) $\forall x, y \in M : d(x, y) = d(y, x);$
- iii) $\forall x, y, z \in M : d(x, z) \leq d(x, y) + d(y, z).$

Das Paar (M, d) heisst dann ein **metrischer Raum**.

Beispiel 3.5.2. i) \mathbb{R} mit $d(x, y) = |x - y|$ für $x, y \in \mathbb{R}$;

ii) \mathbb{Q} mit $d(x, y) = |x - y|$ für $x, y \in \mathbb{Q}$;

iii) \mathbb{R}^n mit $d(x, y) = \|x - y\|$ für $x, y \in \mathbb{R}^n$.

Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in M , und sei $a \in M$.

Definition 3.5.3. i) $(a_n)_{n \in \mathbb{N}}$ heisst **Cauchy-Folge** in M , falls gilt:

$$d(a_n, a_\ell) \rightarrow 0 \quad (\ell, n \rightarrow \infty).$$

ii) $(a_n)_{n \in \mathbb{N}}$ konvergiert gegen a , $\lim_{n \rightarrow \infty} a_n = a$, falls gilt:

$$d(a_n, a) \rightarrow 0 \quad (n \rightarrow \infty).$$

iii) (M, d) heisst **metrisch vollständig**, falls jede Cauchy-Folge $(a_n)_{n \in \mathbb{N}}$ in M konvergiert.

Beispiel 3.5.3. i) \mathbb{R} mit $d(x, y) = |x - y|$ für $x, y \in \mathbb{R}$ ist gemäss Satz 3.5.2 metrisch vollständig.

ii) \mathbb{Q} mit $d(x, y) = |x - y|$ für $x, y \in \mathbb{Q}$ ist nicht metrisch vollständig. Z.B. liefert Beispiel 3.3.1.iii) für $c = 2$ eine Cauchy-Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{Q} . Fassen wir $(a_n)_{n \in \mathbb{N}}$ als Folge in \mathbb{R} auf, so existiert $\lim_{n \rightarrow \infty} a_n = \sqrt{2}$, und wegen Satz 3.2.1 kann $(a_n)_{n \in \mathbb{N}}$ keinen von $a = \sqrt{2}$ verschiedenen Limes haben. Da $\sqrt{2} \notin \mathbb{Q}$ nach Satz 2.1.1, ist $(a_n)_{n \in \mathbb{N}}$ somit in \mathbb{Q} nicht konvergent.

Bemerkung 3.5.2. Wir können \mathbb{R} auch als die *metrische Vervollständigung* von \mathbb{Q} auffassen.

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

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

Satz 3.6.1. *Es sind äquivalent*

i) $a_n \rightarrow a$ ($n \rightarrow \infty$) (gemäss Definition 3.5.3);

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

Insbesondere ist \mathbb{R}^d metrisch vollständig.

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

$$\max_{1 \leq i \leq d} |x^i| \leq \|x\| = \sqrt{\sum_{i=1}^d |x^i|^2} \leq \sqrt{d} \max_{1 \leq i \leq d} |x^i|. \quad (3.6.1)$$

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

$$|a_n^i - a^i| \leq \|a_n - a\| \rightarrow 0 \quad (n \rightarrow \infty).$$

ii) \Rightarrow i): Falls $a_n^i \rightarrow a^i$ ($n \rightarrow \infty$) für $1 \leq i \leq d$, so folgt

$$\max_{1 \leq i \leq d} |a_n^i - a^i| \leq \sum_{i=1}^d |a_n^i - a^i| \rightarrow 0 \quad (n \rightarrow \infty)$$

gemäss Satz 3.2.2, also mit (3.6.1) auch $a_n \rightarrow a$ ($n \rightarrow \infty$). \square

Mit den Sätzen 3.5.1 und 3.5.2 sowie (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.

Beweis. i) \Rightarrow ii) wie Satz 3.5.1.

ii) \Rightarrow i) Mit $(a_n)_{n \in \mathbb{N}}$ ist für jedes $i \in \{1, \dots, d\}$ gemäss (3.6.1) auch $(a_n^i)_{n \in \mathbb{N}}$ eine Cauchy-Folge, also konvergent nach Satz 3.5.2. Die Behauptung folgt mit Satz 3.6.1. \square

Definition 3.6.1. $(a_n)_{n \in \mathbb{N}}$ ist beschränkt, falls gilt

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

Mit Satz 3.4.2 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 $1 \leq i \leq d$ schätze ab

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

Nach Satz 3.4.2 existieren Teilfolgen $N \supset \Lambda_1 \supset \dots \supset \Lambda_d =: \Lambda$ und Häufungspunkte $a^i \in \mathbb{R}$, $1 \leq i \leq d$, mit

$$a_n^i \rightarrow a^i \quad (n \rightarrow \infty, n \in \Lambda \subset \Lambda_i), \quad 1 \leq i \leq d.$$

Mit Satz 3.6.1 folgt

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

□

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*

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

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

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

existiert.

Beispiel 3.7.1. i) Für $|q| < 1$ gilt gemäss Beispiel 2.3.1.iv)

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

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

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

Die Sätze 3.5.1 und 3.5.2 liefern sofort:

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

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

Beweis. $|S_n - S_l| = \left| \sum_{k=l+1}^n a_k \right|$; benutze Satz 3.5.1, bzw. Satz 3.5.2. \square

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

ii) Die Bedingung $a_k \rightarrow 0$ ($k \rightarrow \infty$) ist aber *nicht hinreichend* für Konvergenz, vergleiche die harmonische Reihe, Beispiel 3.5.1.i).

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 $a_k \neq 0$, $k \in \mathbb{N}$.

i) Falls

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

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

ii) falls

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

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

Beweis. i) Setze

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

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

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

insbesondere bei Wahl von $n = n_0$ also

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

Es folgt für $k \geq n_0$ die Abschätzung

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

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

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

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

ii) Es gelte

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

Dann existiert n_0 mit

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

also

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

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

□

Beispiel 3.7.2. i) Die Exponentialreihe

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

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

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

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

erhalten wir mit Satz 3.7.2 Konvergenz.

□

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

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

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

$$\frac{a_{k+1}}{a_k} = z \left(\frac{k}{k+1} \right)^k = z \frac{1}{\left(1 + \frac{1}{k}\right)^k} \rightarrow \frac{z}{e} \quad (k \rightarrow \infty).$$

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 $|z| = e$, also 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 $\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} < 1$, so konvergiert $\sum_{k=1}^{\infty} a_k$;
 ii) falls $\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} > 1$, so divergiert $\sum_{k=1}^{\infty} a_k$.

Beweis. i) Für $q \in \mathbb{R}$ mit

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

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

$$\forall k \geq n_0 : \sqrt[k]{|a_k|} \leq q;$$

also

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

Für $n \geq l \geq n_0$ folgt

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

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

- ii) Falls $\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|} =: 1 + \epsilon > 1$, so gilt

$$\forall k_0 \in \mathbb{N} : \sup_{k \geq k_0} \sqrt[k]{|a_k|} \geq 1 + \epsilon.$$

Für jedes $k_0 \in \mathbb{N}$ gibt es dann gemäss Lemma 2.2.1 ein $k \geq k_0$ mit $\sqrt[k]{|a_k|} \geq 1$, also $|a_k| \geq 1$, und $\sum_{k=1}^{\infty} a_k$ ist divergent nach Bemerkung 3.7.1.i).

□

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

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

Setzen wir $a_k = c_k z^k$ mit

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

so erhalten wir mit Satz 3.7.3 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

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

sie ist divergent für alle $|z| > \rho$. Die Zahl ρ heisst **Konvergenzradius** der Reihe $p(z)$.

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 $|z| < 1$ und Divergenz für $|z| > 1$, während das Quotientenkriterium keine Aussage ermöglicht. Das Wurzelkriterium ist somit "stärker" als das Quotientenkriterium.

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

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

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

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

$$\frac{a_{k+1}}{a_k} = \left(\frac{k}{k+1} \right)^s \rightarrow 1 \quad (k \rightarrow \infty)$$

als auch

$$\sqrt[k]{a_k} = (\sqrt[k]{k})^s \rightarrow 1 \quad (k \rightarrow \infty);$$

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

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

mit

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

wobei $q = 2^{1-s} < 1$. Mit Beispiel 3.7.1.i) folgt, dass $\zeta(s)$ konvergiert für $s > 1$. Zum Beispiel fand Euler $\zeta(2) = \frac{\pi^2}{6}$; 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 $\sum_{k=1}^{\infty} a_k$ konvergiert **absolut**, falls die Reihe $\sum_{k=1}^{\infty} |a_k|$ konvergiert.

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

ii) Da die Partialsummen $S_n = \sum_{k=1}^n |a_k| \leq S_{n+1}$ monoton wachsen, genügt nach Satz 3.3.2 die Beschränktheit dieser Folge für Konvergenz.

Notation: $\sum_{k=1}^{\infty} |a_k| < \infty \Leftrightarrow \sum_{k=1}^{\infty} |a_k|$ 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? – Betrachten wir zunächst nur *eine* Reihe!

Beispiel 3.8.1. Die Reihe $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ 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 ordnet man die Folge $(\frac{(-1)^k}{k})_{k \in \mathbb{N}}$ nun so um, dass auf die ersten k_j positiven Folgenglieder jeweils das j -te negative Glied folgt, $j \in \mathbb{N}$, dann erhält man für die entsprechenden Partialsummen S_{k_l+l} der so umgeordneten alternierenden harmonischen Reihe den Wert

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

das heißt, die so umgeordnete alternierende harmonische Reihe ist divergent!

Anders verhält es sich für eine absolut konvergente Reihe.

Satz 3.8.1. Sei $\sum_{k=1}^{\infty} a_k$ absolut konvergent, und sei $\varphi: \mathbb{N} \rightarrow \mathbb{N}$ 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 $n_1 = \max\{\varphi^{-1}(1), \dots, \varphi^{-1}(n_0)\}$. 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

$$\begin{aligned} \left| \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} a_{\varphi(k)} \right| &\leq \left| \sum_{k=1}^{n_0} a_k - \sum_{k=1}^{n_1} a_{\varphi(k)} \right| + \left| \sum_{k=n_0+1}^{\infty} a_k \right| + \left| \sum_{k=n_1+1}^{\infty} a_{\varphi(k)} \right| \\ &\leq 3 \sum_{k=n_0}^{\infty} |a_k| < 3\epsilon. \end{aligned}$$

□

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.

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

(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 3.2.2 erhalten wir dann auch

$$\sum_{k,l=1}^{\infty} a_k b_l = \lim_{n \rightarrow \infty} \sum_{k,l=1}^n a_k b_l = \lim_{n \rightarrow \infty} \left(\sum_{k=1}^n a_k \cdot \sum_{l=1}^n b_l \right) = \sum_{k=1}^{\infty} a_k \cdot \sum_{l=1}^{\infty} b_l,$$

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

Behauptung. $\sum_{l=1}^{\infty} \left(\sum_{k=1}^{\infty} |a_k b_l| \right)$ 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 3.2.2

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

□

Beispiel 3.8.3. Sei

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

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

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

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

$$(*) = \sum_{k=0}^{\infty} \left(\sum_{n=k}^{\infty} \frac{x^k y^{n-k}}{k! (n-k)!} \right) = \sum_{k=0}^{\infty} \sum_{n=k}^{\infty} \binom{n}{k} \frac{x^k y^{n-k}}{n!}.$$

Erneutes Vertauschen der Summationsreihenfolge ergibt schliesslich

$$(*) = \sum_{n=0}^{\infty} \frac{1}{n!} \underbrace{\left(\sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \right)}_{=(x+y)^n} = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!}.$$

Es folgt

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

$$\text{Exp}(x) \cdot \text{Exp}(y) = \text{Exp}(x + y).$$

3.9 Die Exponentialreihe und die Funktion e^x

Die Exponentialreihe hat interessante Eigenschaften.

Satz 3.9.1. Es gilt

$$\text{Exp}(1) = \sum_{k=1}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2} + \dots = e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Beweis. Nach dem binomischen Lehrsatz gilt

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

mit

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

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

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

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

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

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

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

□

Mit Korollar 3.8.1 folgt aus Satz 3.9.1 induktiv

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

Unter Beachtung von

$$\text{Exp}(n) \cdot \text{Exp}(-n) = \text{Exp}(0) = 1$$

erhalten wir dann weiter

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

Analog gilt für $p, q \in \mathbb{N}$

$$\text{Exp}\left(\frac{p}{q}\right) = \left(\text{Exp}\left(\frac{1}{q}\right)\right)^p; \quad (3.9.1)$$

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

$$\text{Exp}\left(\frac{1}{q}\right) = (\text{Exp}(1))^{\frac{1}{q}} = e^{\frac{1}{q}},$$

und mit (3.9.1) dann auch

$$\forall p, q \in \mathbb{N} : \text{Exp}\left(\frac{p}{q}\right) = \left(e^{\frac{1}{q}}\right)^p = e^{\frac{p}{q}};$$

das heisst, wir erhalten:

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

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 $z = x + iy \in \mathbb{C}$

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

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

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

Satz 3.9.3. $e = \exp(1)$ ist irrational.

Beweis (indirekt). Nimm an, $e = \frac{m}{n}$, $m, n \in \mathbb{N}$. Setze

$$S_n = \sum_{k=0}^n \frac{1}{k!} < e$$

und schätze ab

$$\begin{aligned} 0 < e - S_n &= \sum_{k=n+1}^{\infty} \frac{1}{k!} = \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \dots \right) \\ &\leq \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \left(\frac{1}{n+2} \right)^2 + \dots \right) \\ &= \frac{1}{(n+1)!} \sum_{\ell=0}^{\infty} \left(\frac{1}{n+2} \right)^{\ell} \quad (\text{geometrische Reihe}) \\ &= \frac{1}{(n+1)!} \cdot \frac{1}{1 - \frac{1}{n+2}} = \frac{n+2}{(n+1)!(n+1)} = \frac{1}{n!} \cdot \frac{n+2}{(n+1)^2}. \end{aligned}$$

Es folgt:

$$0 < n!(e - S_n) \leq \frac{n+2}{(n+1)^2} < 1.$$

Andererseits gilt nach Annahme

$$n!(e - S_n) = (n-1)! m - \sum_{k=0}^n \frac{n!}{k!} \in \mathbb{Z}.$$

Widerspruch. □

Kapitel 4

Topologische Grundbegriffe

4.1 Topologie des \mathbb{R}^n

Sei $\|\cdot\|$ die euklidische Norm auf \mathbb{R}^n und

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, \quad x, y \in \mathbb{R}^n$$

die davon induzierte Metrik.

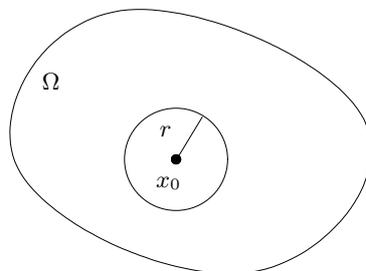
Definition 4.1.1. i) Sei $x_0 \in \mathbb{R}^n$. Der **offene Ball** vom Radius $r > 0$ um x_0 ist die Menge

$$B_r(x_0) = \{x \in \mathbb{R}^n; \|x - x_0\| < r\}.$$

ii) $x_0 \in \Omega$ heisst **innerer Punkt** von $\Omega \subset \mathbb{R}^n$, falls

$$\exists r > 0 : B_r(x_0) \subset \Omega.$$

iii) $\Omega \subset \mathbb{R}^n$ heisst **offen**, falls jedes $x_0 \in \Omega$ ein innerer Punkt ist.

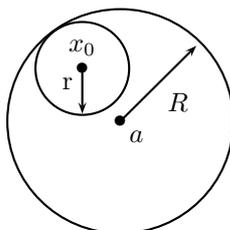


Beispiel 4.1.1. i) Jeder offene Ball $B_R(a) \subset \mathbb{R}^n$ ist offen.

Beweis. Sei $x_0 \in B_R(a)$. Setze $r = R - \|x_0 - a\| > 0$. Dann gilt für $x \in B_r(x_0)$

$$\|x - a\| \leq \underbrace{\|x - x_0\|}_{< r} + \|x_0 - a\| < R.$$

Also gilt $B_r(x_0) \subset B_R(a)$, und $B_R(a)$ ist offen. \square



ii) Sei $n = 1$. Für $x \in \mathbb{R}$, $r > 0$ gilt:

$$B_r(x) = \{y \in \mathbb{R}; x - r < y < x + r\} =]x - r, x + r[,$$

das "offene Intervall" der Länge $2r$ um x . Mit i) ist jedes offene Intervall $]a, b[$ offen, da $]a, b[= B_r(x) \subset \mathbb{R}$ mit $x = \frac{a+b}{2}$ und $r = \frac{b-a}{2} > 0$.

iii) Das "halboffene Intervall"

$$[a, b[= \{x \in \mathbb{R}; a \leq x < b\}$$

ist nicht offen, da a kein innerer Punkt ist. Für jedes $r > 0$ enthält nämlich $B_r(a) =]a - r, a + r[$ Punkte $y < a$; das heisst, $B_r(a) =]a - r, a + r[\not\subset [a, b[$.

Satz 4.1.1. *Es gilt:*

- i) \emptyset, \mathbb{R}^n sind offen;
- ii) $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ offen $\Rightarrow \Omega_1 \cap \Omega_2$ offen;
- iii) $\Omega_\iota \subset \mathbb{R}^n$ 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 $B_r(x) \subset \Omega_1 \cap \Omega_2$, und $\Omega_1 \cap \Omega_2$ ist offen.

iii) Sei $x \in \bigcup_{\iota \in I} \Omega_\iota$. Dann gibt es $\iota_0 \in I$ mit $x \in \Omega_{\iota_0}$. Da Ω_{ι_0} offen, gibt es $r > 0$ mit

$$B_r(x) \subset \Omega_{\iota_0} \subset \bigcup_{\iota \in I} \Omega_\iota.$$

\square

Bemerkung 4.1.1. Im allgemeinen ist der Durchschnitt unendlich vieler offener Mengen *nicht* offen; zum Beispiel gilt

$$\bigcap_{k=1}^{\infty} B_{1/k}(0) = \{0\}.$$

Satz 4.1.1 motiviert die folgende Definition für eine beliebige Menge M .

Definition 4.1.2. Eine Familie $\tau \subset \mathcal{P}(M)$ heisst eine **Topologie** auf M , falls gilt:

- i) $\emptyset, M \in \tau$;
- ii) $\Omega_1, \Omega_2 \in \tau \Rightarrow \Omega_1 \cap \Omega_2 \in \tau$;
- iii) $\Omega_\iota \in \tau, \iota \in I \Rightarrow \bigcup_{\iota \in I} \Omega_\iota \in \tau$.

Beispiel 4.1.2. i) Gemäss Satz 4.1.1 definieren die offenen Mengen in \mathbb{R}^n eine Topologie auf \mathbb{R}^n .

ii) Offenbar sind für jede Menge M die “diskrete” Topologie $\tau = \mathcal{P}(M)$ sowie die “triviale” Topologie $\tau = \{\emptyset, M\}$ Topologien auf M .

Wir kehren zurück zu \mathbb{R}^n mit der Topologie der offenen Mengen.

Definition 4.1.3. $A \subset \mathbb{R}^n$ heisst **abgeschlossen**, falls $\mathbb{R}^n \setminus A$ offen ist.

Beispiel 4.1.3. i) Das “abgeschlossene Intervall” $A := [a, b] \subset \mathbb{R}$ ist abgeschlossen, da $\mathbb{R} \setminus A =]-\infty, a[\cup]b, \infty[$ nach Beispiel 4.1.1.ii) und Satz 4.1.1.iii) offen ist.

ii) Der “abgeschlossene Ball”

$$\overline{B_R(x_0)} := \{x \in \mathbb{R}^n; \|x - x_0\| \leq R\}$$

ist abgeschlossen.

Beweis. Sei $x \in \mathbb{R}^n$ mit $\|x - x_0\| > R$. Setze $r = \|x - x_0\| - R > 0$. Für $y \in B_r(x)$ gilt nach Dreiecksungleichung

$$\|y - x_0\| \geq \|x - x_0\| - \|y - x\| > \|x - x_0\| - r = R;$$

also ist $\mathbb{R}^n \setminus \overline{B_R(x_0)}$ offen. □

iii) Sei $n \geq 2$. Betrachte die Menge

$$A = \{(x_1, \underbrace{0, \dots, 0}_{n-1}); x_1 \in \mathbb{R}\} = \mathbb{R} \times \{\underbrace{(0, \dots, 0)}_{n-1}\} \subset \mathbb{R}^n.$$

Behauptung. A ist abgeschlossen.

Beweis. Sei $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \setminus A = A^c$, das heisst

$$r = \max\{|x_2|, \dots, |x_n|\} > 0.$$

Dann ist $B_r(x) \subset \mathbb{R}^n \setminus A$; also $\mathbb{R}^n \setminus A$ offen. □ □

Mit den de Morganschen Regeln, Satz 1.2.2, ergibt sich aus Satz 4.1.1 unmittelbar:

Satz 4.1.2. *Es gilt:*

- i) \emptyset, \mathbb{R}^n sind abgeschlossen;
- ii) A_1, A_2 abgeschlossen $\Rightarrow A_1 \cup A_2$ abgeschlossen;
- iii) A_ι abgeschlossen, $\iota \in I \Rightarrow \bigcap_{\iota \in I} A_\iota$ abgeschlossen.

Beweis. i) klar.

ii) Sei A_i abgeschlossen, also $\Omega_i = \mathbb{R}^n \setminus A_i = A_i^c$ offen, $i = 1, 2$. Mit Satz 4.1.1.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.

iii) Falls A_ι abgeschlossen, $\iota \in I$, so ist

$$\left(\bigcap_{\iota \in I} A_\iota \right)^c = \bigcup_{\iota \in I} A_\iota^c$$

gemäss Satz 4.1.1.iii) offen, die Menge $\bigcap_{\iota \in I} A_\iota$ also abgeschlossen. □

Bemerkung 4.1.2. Im Allgemeinen ist die Vereinigung unendlich vieler abgeschlossener Mengen nicht abgeschlossen. Betrachte dazu als Beispiel die Mengen $A_k = [a_k, b_k] \subset \mathbb{R}$ mit $a_k = \frac{1}{k}$ und $b_k = 1 - \frac{1}{k}$, $k \in \mathbb{N}$. Dann ist

$$\bigcup_{k \in \mathbb{N}} A_k =]0, 1[$$

nicht abgeschlossen.

Definition 4.1.4. i) *Die Menge*

$$\text{int}(\Omega) = \bigcup_{U \subset \Omega, U \text{ offen}} U =: \Omega^\circ$$

heisst **offener Kern** oder das **Innere** (engl.: interior) von Ω .

ii) *Die Menge*

$$\text{clos}(\Omega) = \bigcap_{A \supset \Omega, A \text{ abgeschlossen}} A =: \bar{\Omega}$$

heisst **Abschluss** (engl.: closure) von Ω .

iii) *Die Menge*

$$\partial\Omega = \text{clos}(\Omega) \setminus \text{int}(\Omega) = \bar{\Omega} \setminus \Omega^\circ$$

heisst **Rand** von Ω .

Bemerkung 4.1.3. i) Ω° ist die Menge der inneren Punkte von Ω .

ii) Nach Satz 4.1.1.iii) ist $\Omega^\circ \subset \Omega$ offen. Offenbar ist $U = \Omega^\circ$ die *grösste* offene Menge $U \subset \Omega$, und Ω ist offen genau dann, wenn $\Omega = \Omega^\circ$.

iii) Analog ist nach Satz 4.1.2.iii) die Menge $A = \overline{\Omega} \supset \Omega$ die *kleinste* abgeschlossene Menge $A \supset \Omega$, und Ω ist abgeschlossen genau dann, wenn $\Omega = \overline{\Omega}$.

Beispiel 4.1.4. Sei zunächst $n = 1$.

i) Für $a < b$ in \mathbb{R} gilt:

$$([a, b])^\circ =]a, b[^\circ = ([a, b])^\circ =]a, b[^\circ =]a, b[^\circ \subset \mathbb{R}.$$

Beweis. $]a, b[$ ist offen, also $]a, b[\subset ([a, b])^\circ$. Weder a noch b sind innere Punkte von $[a, b]$, also $]a, b[= ([a, b])^\circ$. Der Rest ist analog. \square

ii) Für $a < b$ in \mathbb{R} gilt analog mit Beispiel 4.1.3.i):

$$\overline{]a, b[} = \overline{]a, b[} = \overline{]a, b[} = [a, b] = \overline{]a, b[}.$$

iii) Für $a < b$ folgt mit i) und ii):

$$\partial]a, b[= \partial ([a, b]) = \partial]a, b[= \partial]a, b[= \{a, b\}.$$

iv) $\mathbb{R} \setminus \mathbb{Q}$ enthält keinen inneren Punkt, da jedes Intervall $B_r(x) =]x - r, x + r[$ auch rationale Punkte enthält; das heisst $(\mathbb{R} \setminus \mathbb{Q})^\circ = \emptyset$.

v) $\mathbb{Q}^\circ = \emptyset$, da jedes $B_r(x)$ auch Punkte der Form $y = \sqrt{2} + q$ enthält mit $q \in \mathbb{Q}$, also $y \notin \mathbb{Q}$.

vi) $\overline{\mathbb{Q}} = \overline{\mathbb{R} \setminus \mathbb{Q}} = \mathbb{R}$, da für jedes abgeschlossene $A \subsetneq \mathbb{R}$ die Menge A^c ein offenes Intervall enthält und damit auch Punkte von \mathbb{Q} und $\mathbb{R} \setminus \mathbb{Q}$.

vii) Mit iv, v) und vi) folgt

$$\partial \mathbb{Q} = \overline{\mathbb{Q}} \setminus \mathbb{Q}^\circ = \mathbb{R} = \partial (\mathbb{R} \setminus \mathbb{Q}).$$

Der "Rand" einer Menge kann also erstaunlich gross sein!

viii) Sei $n \geq 1$. Für $x_0 \in \mathbb{R}^n$, $R > 0$ gilt gemäss Beispiel 4.1.1.i) und 4.1.3.ii):

$$\partial B_R(0) = \overline{B_R(0)} \setminus B_R(0) = \{x; \|x\| = R\}.$$

Den Rand einer Menge kann man auch wie folgt charakterisieren.

Satz 4.1.3. Für $\Omega \subset \mathbb{R}^n$ mit Komplement $\Omega^c = \mathbb{R}^n \setminus \Omega$ gilt

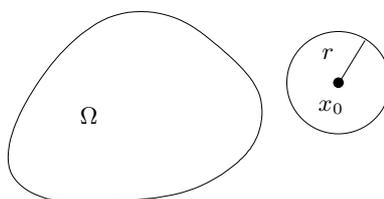
$$\partial \Omega = \{x \in \mathbb{R}^n; \forall r > 0 : B_r(x) \cap \Omega \neq \emptyset \neq B_r(x) \cap \Omega^c\}.$$

Beweis. Sei $x \in \mathbb{R}^n$ ein beliebiger Punkt. Falls $B_r(x) \cap \Omega = \emptyset$ für ein $r > 0$, so folgt $x \notin \overline{\Omega}$. (Wähle $A = \mathbb{R}^n \setminus B_r(x) \supset \Omega$.) Da $\overline{\Omega}$ abgeschlossen, $\mathbb{R}^n \setminus \overline{\Omega}$ offen, gilt auch die Umkehrung; das heisst,

$$x \notin \overline{\Omega} \Leftrightarrow \exists r > 0 : B_r(x) \cap \Omega = \emptyset,$$

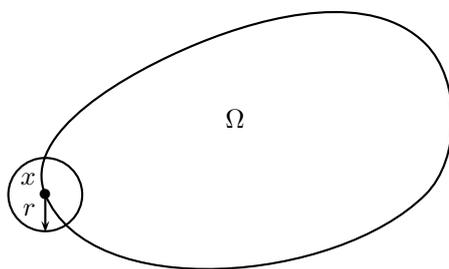
oder, dazu äquivalent,

$$x \in \overline{\Omega} \Leftrightarrow \forall r > 0 : B_r(x) \cap \Omega \neq \emptyset.$$



Analog gilt $B_r(x) \cap \Omega^c = \emptyset$, also $B_r(x) \subset \Omega$, für ein $r > 0$ genau dann, wenn $x \in \Omega^\circ$; das heisst,

$$x \in \Omega^\circ \Leftrightarrow \forall r > 0 : B_r(x) \cap \Omega^c = \emptyset.$$



□

Zwischen Ω , $\bar{\Omega}$ und $\partial\Omega$ bestehen folgende Beziehungen:

Satz 4.1.4. Für $\Omega \subset \mathbb{R}^n$ gilt:

- i) $\partial\Omega = \bar{\Omega} \setminus \Omega^\circ = \bar{\Omega} \cap (\mathbb{R}^n \setminus \Omega^\circ)$ ist abgeschlossen, $\partial\Omega = \partial(\Omega^c)$.
- ii) $\bar{\Omega} = \Omega^\circ \cup \partial\Omega$, und die Zerlegung ist disjunkt.
- iii) $\bar{\Omega} = \Omega \cup \partial\Omega$.
- iv) $\Omega^\circ = \Omega \setminus \partial\Omega$.

Beweis. i) Abgeschlossenheit von $\partial\Omega$ folgt mit Satz 4.1.2.iii), Satz 4.1.3 liefert die Beziehung $\partial\Omega = \partial(\Omega^c)$.

Die Aussagen ii) - iv) folgen mit $\Omega^\circ \subset \Omega \subset \bar{\Omega}$ und der Definition $\partial\Omega = \bar{\Omega} \setminus \Omega^\circ$. □

Zusammen mit Bemerkung 4.1.3 erhalten wir das Kriterium

Korollar 4.1.1. Für $\Omega \subset \mathbb{R}^n$ gilt:

- i) Ω ist abgeschlossen genau dann, wenn $\partial\Omega \subset \Omega$.
- ii) Ω ist offen genau dann, wenn $\partial\Omega \cap \Omega = \emptyset$.

4.2 Topologie und Konvergenz

Sei $a \in \mathbb{R}^n$.

Definition 4.2.1. Eine Menge $U \subset \mathbb{R}^n$ heisst Umgebung von a , falls a innerer Punkt von U ist, das heisst, falls ein $r > 0$ existiert mit $B_r(a) \subset U$.

Beispiel 4.2.1. Offenbar ist jeder Ball $B_r(a)$, $r > 0$, eine Umgebung von a .

Sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R}^n , und sei $a \in \mathbb{R}^n$. Mit Beispiel 4.2.1 folgt die Äquivalenz der Aussagen

i) $a_k \rightarrow a$ ($k \rightarrow \infty$), also

$$\forall \epsilon > 0 \exists k_0 = k_0(\epsilon) \in \mathbb{N} \forall k \geq k_0 : \|a_k - a\| < \epsilon.$$

ii) Für jede Umgebung U von a gilt:

$$\exists k_0 = k_0(U) \in \mathbb{N} \forall k \geq k_0 : a_k \in U.$$

Analoges gilt für Häufungspunkte. Umgekehrt können wir auch topologische Eigenschaften von Mengen mittels Konvergenz geeigneter Folgen charakterisieren.

Satz 4.2.1. (Folgenkriterium für Abgeschlossenheit) Für $A \subset \mathbb{R}^n$ sind äquivalent:

i) A ist abgeschlossen,

ii) $\forall (a_k)_{k \in \mathbb{N}} \subset A : a_k \rightarrow a$ ($k \rightarrow \infty$) $\Rightarrow a \in A$.

Beweis. i) \Rightarrow ii) (indirekt): Sei $(a_k)_{k \in \mathbb{N}} \subset A$ eine Folge in A mit $a_k \rightarrow a$ ($k \rightarrow \infty$), und nimm widerspruchswise an, $a \notin A$. Da A nach Annahme abgeschlossen, ist A^c offen, und es gibt ein $r > 0$ mit $B_r(a) \subset A^c$; das heisst,

$$B_r(a) \cap A = \emptyset.$$

Für $k \geq k_0(r)$ gilt jedoch $a_k \in B_r(a)$. Da andererseits $a_k \in A$ für jedes k , erhalten wir den gewünschten Widerspruch.

ii) \Rightarrow i): Gemäss Korollar 4.1.1 genügt es zu zeigen, dass $\partial A \subset A$. Sei dazu $x \in \partial A$. Nach Satz 4.1.3 gilt

$$\forall r > 0 : B_r(x) \cap A \neq \emptyset.$$

Für $r_k = \frac{1}{k}$ wähle $a_k \in B_{r_k}(x) \cap A$, $k \in \mathbb{N}$. Dann gilt $a_k \rightarrow x$ ($k \rightarrow \infty$). Da $(a_k)_{k \in \mathbb{N}} \subset A$ nach Konstruktion, folgt mit Annahme ii), dass auch $x \in A$. \square

Beispiel 4.2.2. i) Sei $0 < \ell < n$, und setze

$$A = \mathbb{R}^\ell \times \underbrace{\{(0, \dots, 0)\}}_{n-\ell} \subset \mathbb{R}^n.$$

Dann ist A abgeschlossen.

Beweis. Für jede Folge $(a_k)_{k \in \mathbb{N}} \subset A$ gilt

$$a_k = (a_k^1, \dots, a_k^\ell, 0, \dots, 0), \quad k \in \mathbb{N}.$$

Falls $a_k \rightarrow a$ ($k \rightarrow \infty$), so folgt $a = (a^1, \dots, a^\ell, 0, \dots, 0) \in A$, und A ist abgeschlossen nach Satz 4.2.1. \square

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

$$A = \{x \in \mathbb{R}^n; x \text{ ist Häufungspunkt von } (a_k)_{k \in \mathbb{N}}\}.$$

Dann ist A abgeschlossen.

Beweis. Sei $(x_\ell)_{\ell \in \mathbb{N}}$ eine Folge in A . Dann gibt es zu jedem $\ell \in \mathbb{N}$ ein $k_\ell \geq \ell$ mit

$$\|a_{k_\ell} - x_\ell\| < \frac{1}{\ell}.$$

Falls $x_\ell \rightarrow x$ ($\ell \rightarrow \infty$), so folgt:

$$\|a_{k_\ell} - x\| \leq \|a_{k_\ell} - x_\ell\| + \|x_\ell - x\| \rightarrow 0 \quad (\ell \rightarrow \infty);$$

das heisst, $a_{k_\ell} \rightarrow x$ ($\ell \rightarrow \infty$), und $x \in A$. Satz 4.2.1 ergibt die Behauptung. \square

4.3 Kompakte Mengen

Eine Menge $\Omega \subset \mathbb{R}^n$ ist beschränkt, falls $\sup_{x \in \Omega} \|x\| < \infty$.

Definition 4.3.1. Eine Menge $K \subset \mathbb{R}^n$ heisst **kompakt**, falls K beschränkt und abgeschlossen ist.

Beispiel 4.3.1. i) $[a, b] \subset \mathbb{R}$, $\overline{B_r(x_0)} \subset \mathbb{R}^n$ sind kompakt.

ii) \mathbb{R}^n ist unbeschränkt, also nicht kompakt.

iii) $]a, b[$, $B_r(x)$ sind nicht abgeschlossen, also nicht kompakt.

iv) Endliche Vereinigungen und beliebige Durchschnitte kompakter Mengen sind kompakt.

v) Sei $(a_k)_{k \in \mathbb{N}}$ beschränkte Folge in \mathbb{R}^n . Setze

$$K = \{a; a \text{ ist Häufungspunkt von } (a_k)_{k \in \mathbb{N}}\}.$$

Dann ist K beschränkt und abgeschlossen nach Beispiel 4.2.2.ii). also kompakt.

Satz 4.3.1. Für $K \subset \mathbb{R}^n$ sind äquivalent:

i) K ist kompakt.

ii) K ist **folgenkompakt**; das heisst, jede Folge $(x_k)_{k \in \mathbb{N}} \subset K$ besitzt einen Häufungspunkt in K .

iii) K ist **überdeckungskompakt**; das heisst, jede Überdeckung $K \subset \bigcup_{\iota \in I} \Omega_\iota$ mit offenen Mengen Ω_ι , $\iota \in I$, besitzt eine endliche Teilüberdeckung $(\Omega_{\iota_\ell})_{1 \leq \ell \leq L}$ mit $K \subset \bigcup_{\ell=1}^L \Omega_{\iota_\ell}$ (**Heine-Borel-Eigenschaft**).

Bevor wir Satz 4.3.1 beweisen, illustrieren wir das Heine-Borel Kriterium mit Beispielen.

Beispiel 4.3.2. i) Die Familie $(] \frac{1}{k}, 1[)_{k \in \mathbb{N}}$ ist eine offene Überdeckung von $]0, 1[$, die keine endliche Teilüberdeckung enthält. (Die Menge $]0, 1[$ ist nicht abgeschlossen, also nicht kompakt.)

ii) Die Familie $(]k, k+2])_{k \in \mathbb{Z}}$ ist eine offene Überdeckung von \mathbb{R} ohne endliche Teilüberdeckung. (Die Menge \mathbb{R} ist unbeschränkt, also nicht kompakt.)

Das Heine-Borel Kriterium hat verblüffende Anwendungen.

Korollar 4.3.1. Sei $K_\iota, \iota \in I$, eine Familie kompakter Mengen mit

$$K_{\iota_1} \cap K_{\iota_2} \cap \dots \cap K_{\iota_L} \neq \emptyset$$

für jede endliche Auswahl $\iota_1, \dots, \iota_L \in I$. Dann gilt:

$$\bigcap_{\iota \in I} K_\iota \neq \emptyset.$$

Beweis (indirekt). Fixiere ein $\iota_0 \in I$. Nimm widerspruchswise an

$$\bigcap_{\iota \in I} K_\iota = \emptyset.$$

Setze $\Omega_\iota = K_\iota^c = \mathbb{R}^n \setminus K_\iota, \iota \in I$. Dann ist Ω_ι offen, $\iota \in I$, und es gilt

$$\bigcup_{\iota \in I} \Omega_\iota = \left(\bigcap_{\iota \in I} K_\iota \right)^c = \mathbb{R}^n \supset K_{\iota_0}.$$

Das heisst, $(\Omega_\iota)_{\iota \in I}$ ist offene Überdeckung von K_{ι_0} . Da K_{ι_0} kompakt ist, existiert gemäss Satz 4.3.1 eine Auswahl $\iota_1, \dots, \iota_L \in I$ mit

$$K_{\iota_0} \subset \bigcup_{\ell=1}^L \Omega_{\iota_\ell} = \left(\bigcap_{\ell=1}^L K_{\iota_\ell} \right)^c,$$

also

$$\bigcap_{\ell=0}^L K_{\iota_\ell} = K_{\iota_0} \cap \left(\bigcap_{\ell=1}^L K_{\iota_\ell} \right) = \emptyset$$

im Widerspruch zu Voraussetzung. \square

Bemerkung 4.3.1. Korollar 4.3.1 ist falsch für offene und beschränkte Mengen. Zum Beispiel besitzen die Mengen $\Omega_n = (]0, \frac{1}{n}[)$, $n \in \mathbb{N}$, die endliche Durchschnittseigenschaft

$$\forall n \in \mathbb{N} : \bigcap_{k=1}^n \Omega_k =]0, \frac{1}{n}[\neq \emptyset,$$

jedoch gilt

$$\bigcap_{k=1}^{\infty} \Omega_k = \emptyset.$$

Die Kompaktheit der Mengen K_ι in Korollar 4.3.1 ist also eine unverzichtbare Voraussetzung.

Beweis von Satz 4.3.1. i) \Rightarrow iii) (indirekt). Sei K kompakt, und nimm an, es gibt eine offene Überdeckung $(\Omega_\iota)_{\iota \in I}$ von K ohne endliche Teilüberdeckung. Für jedes $k \in \mathbb{N}$ wird K überdeckt durch endlich viele Würfel eines achsenparallelen Gitters der Maschenweite $\frac{2}{\sqrt{n} \cdot k}$, also auch durch endlich viele abgeschlossene Kugeln

$$\overline{B_{1/k}(x_j^{(k)})}, \quad 1 \leq j \leq J_k.$$

Setze

$$K_j^{(k)} = K \cap \overline{B_{1/k}(x_j^{(k)})}, \quad 1 \leq j \leq J_k.$$

Behauptung 1. Für jedes $k \in \mathbb{N}$ gibt es $j_k \in \{1, \dots, J_k\}$ so, dass $K_{j_k}^{(k)}$ nicht von endlich vielen Ω_ι überdeckt wird.

Beweis. Falls für ein $k \in \mathbb{N}$ jede der Mengen $K_j^{(k)}$, $1 \leq j \leq J_k$, durch endlich viele Ω_ι überdeckt wird, liefert dies offenbar eine endliche Teilüberdeckung von K , im Widerspruch zu unserer Annahme. \square

Sei $(a_k)_{k \in \mathbb{N}}$ eine Folge mit $a_k \in K_{j_k}^{(k)} \subset K$, $k \in \mathbb{N}$, gemäss Behauptung 1. Da K beschränkt, gibt es nach Satz 3.4.2 (Bolzano-Weierstrass) eine Teilfolge $\Lambda \subset \mathbb{N}$ und $a \in \mathbb{R}^n$ mit $a_k \rightarrow a$ ($k \rightarrow \infty$, $k \in \Lambda$). Da K zudem abgeschlossen, folgt mit Satz 4.2.1, dass $a \in K$. Wegen

$$K \subset \bigcup_{\iota \in I} \Omega_\iota,$$

gibt es $\iota_0 \in I$ mit $a \in \Omega_{\iota_0}$; da weiter Ω_{ι_0} offen, existiert $r > 0$ mit

$$B_r(a) \subset \Omega_{\iota_0}.$$

Wähle $k_0 = k_0(r)$ mit

$$\|a_k - a\| + \frac{3}{k} < r,$$

falls $k \geq k_0$, $k \in \Lambda$. Da $a_k \in \overline{K_{j_k}^{(k)}} \subset \overline{B_{1/k}(x_{j_k}^{(k)})}$, gilt $\|a_k - x_{j_k}^{(k)}\| \leq 1/k$, und mit Dreiecksungleichung folgt $\overline{B_{1/k}(x_{j_k}^{(k)})} \subset B_{3/k}(a_k)$, $k \in \mathbb{N}$. Damit gilt für jedes $k \geq k_0$, $k \in \Lambda$, die Inklusionskette

$$K_{j_k}^{(k)} \subset \overline{B_{1/k}(x_{j_k}^{(k)})} \subset B_{3/k}(a_k) \subset B_r(a) \subset \Omega_{\iota_0},$$

im Widerspruch zur Auswahl von j_k nach Behauptung 1.

iii) \Rightarrow ii) (indirekt) Sei K Überdeckungskompakt, und sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in K , welche keinen Häufungspunkt $x \in K$ besitzt; das heisst,

$$\forall x \in K \exists r = r(x) > 0, k_0 = k_0(x) \in \mathbb{N} \forall k \geq k_0 : a_k \notin B_r(x).$$

Die Familie $(B_{r(x)}(x))_{x \in K}$ ist eine offene Überdeckung von K . Nach Annahme iii) genügen endlich viele Bälle $B_{r(x_\ell)}(x_\ell)$, $1 \leq \ell \leq L$, um K zu überdecken. Diese enthalten jedoch keines der Folgenglieder a_k mit $k \geq \max_{1 \leq \ell \leq L} k_0(x_\ell)$. Widerspruch.

ii) \Rightarrow i) Wir führen den Beweis in zwei Schritten.

Behauptung 2. K ist beschränkt.

Beweis (indirekt). Falls K unbeschränkt, gibt es zu $k \in \mathbb{N}$ ein $a_k \in K$ mit $\|a_k\| \geq k$. Offenbar kann $(a_k)_{k \in \mathbb{N}}$ keinen Häufungspunkt haben, im Widerspruch zur Annahme ii). \square

Behauptung 3. K ist abgeschlossen.

Beweis. Sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in K mit $a_k \rightarrow a$ ($k \rightarrow \infty$). Nach Annahme ii) gilt $a \in K$; also ist K abgeschlossen nach Satz 4.2.1. \square

Der Satz ist somit vollständig bewiesen. \square

Abschliessend geben wir noch ein Beispiel zum Folgenkriterium, Satz 4.3.1.ii).

Beispiel 4.3.3. Sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R}^n mit $a_k \rightarrow a$ ($k \rightarrow \infty$). Dann ist die Menge

$$K = \{a_k; k \in \mathbb{N}\} \cup \{a\}$$

kompakt gemäss Satz 4.3.1.ii).

4.4 Zusammenhängende Mengen

Sei $\Omega \subset \mathbb{R}^n$ beliebig.

Definition 4.4.1. Ω heisst **zusammenhängend**, falls für je zwei offene Mengen Ω_1, Ω_2 gilt:

$$\Omega \subset (\Omega_1 \cup \Omega_2), \quad \Omega \cap \Omega_i \neq \emptyset, \quad i = 1, 2 \Rightarrow \Omega \cap \Omega_1 \cap \Omega_2 \neq \emptyset.$$

Eine zusammenhängende Menge kann man also nicht disjunkt in “relativ offene”, nicht leere Teile zerlegen.

Satz 4.4.1. Die zusammenhängenden Teilmengen von \mathbb{R} sind genau die Intervalle.

Beweis. i) Sei $\Omega \subset \mathbb{R}$ kein Intervall. OBdA sei $\Omega \neq \emptyset$. Setze

$$a = \inf \Omega \geq -\infty, \quad b = \sup \Omega \leq \infty.$$

Da Ω kein Intervall, gibt es $c \in \mathbb{R}$ mit $a < c < b$ und $c \notin \Omega$. Die Mengen

$$\Omega_1 =] - \infty, c[, \quad \Omega_2 =]c, \infty[$$

sind dann offen mit $\Omega \subset (\Omega_1 \cup \Omega_2) = \mathbb{R} \setminus \{c\}$. Nach Definition des Infimums, beziehungsweise des Supremums gilt weiter

$$\Omega \cap \Omega_1 \neq \emptyset \neq \Omega \cap \Omega_2.$$

Schliesslich gilt $\Omega_1 \cap \Omega_2 = \emptyset$, also auch $\Omega \cap \Omega_1 \cap \Omega_2 = \emptyset$, und Ω ist nicht zusammenhängend.

ii) Sei Ω ein Intervall, und seien $\Omega_1, \Omega_2 \subset \mathbb{R}$ offen mit $\Omega \subset (\Omega_1 \cup \Omega_2)$, $\Omega \cap \Omega_i \neq \emptyset$, $i = 1, 2$. Wähle $x_i \in \Omega \cap \Omega_i$, $i = 1, 2$. OBdA $x_1 < x_2$, $x_2 \notin \Omega_1$. (Sonst $x_2 \in \Omega \cap \Omega_1 \cap \Omega_2$, und wir sind fertig). Setze

$$y_1 = \sup\{x \in \mathbb{R}; [x_1, x] \subset \Omega_1\}.$$

Dann gilt offenbar $[x_1, y_1[\subset \Omega_1$. Da Ω_1 offen, $x_2 \notin \Omega_1$, folgt zudem $x_1 < y_1 \leq x_2$.

Behauptung. $y_1 \notin \Omega_1$.

Beweis. Falls $y_1 \in \Omega_1$, so gibt es $r > 0$ mit $]y_1 - r, y_1 + r[\subset \Omega_1$, also

$$[x_1, y_1 + \frac{r}{2}] \subset [x_1, y_1[\cup]y_1 - r, y_1 + r[\subset \Omega_1$$

in Widerspruch zur Definition von y_1 . □

Da Ω nach Annahme ein Intervall ist, gilt $[x_1, x_2] \subset \Omega$. Mit $y_1 \notin \Omega_1$ und

$$y_1 \in [x_1, x_2] \subset \Omega \subset (\Omega_1 \cup \Omega_2)$$

folgt $y_1 \in \Omega_2$. Da Ω_2 offen, gibt es $r > 0$ mit $]y_1 - r, y_1 + r[\subset \Omega_2$, also

$$\Omega \cap \Omega_1 \cap \Omega_2 \supset ([x_1, x_2] \cap [x_1, y_1[\cap]y_1 - r, y_1 + r]) \neq \emptyset.$$

□

Kann man zusammenhängende Mengen in \mathbb{R}^n für $n \geq 2$ auf ähnliche Weise charakterisieren? Eine hinreichende Bedingung liefert der folgende Begriff:

Definition 4.4.2. $\Omega \subset \mathbb{R}^n$ heisst **konvex**, falls für beliebige $x_0, x_1 \in \Omega$ gilt:

$$\forall t \in [0, 1] : x_t = (1 - t)x_0 + tx_1 \in \Omega.$$

Satz 4.4.2. Eine konvexe Menge $\Omega \subset \mathbb{R}^n$ ist zusammenhängend.

Beweis. Sei Ω konvex, und seien $\Omega_0, \Omega_1 \subset \mathbb{R}^n$ offen mit $\Omega \subset (\Omega_0 \cup \Omega_1)$ sowie $\Omega \cap \Omega_i \neq \emptyset$, $i = 0, 1$. Wähle $x_i \in \Omega \cap \Omega_i$, $i = 0, 1$. OBdA $x_0 \neq x_1$. Da Ω konvex, gilt

$$\forall t \in [0, 1] = I : x_t = (1 - t)x_0 + tx_1 \in \Omega. \quad (4.4.1)$$

Setze

$$J_i = \{t \in \mathbb{R}; x_t \in \Omega \cap \Omega_i\}, \quad i = 0, 1,$$

wobei wir x_t wie in (4.4.1) definieren.

Behauptung. J_i ist offen, $i = 0, 1$.

Beweis. Sei $t \in J_0$, also $x_t \in \Omega_0$. Da Ω_0 offen, gibt es $r > 0$ mit $B_r(x_t) \subset \Omega_0$. Setze

$$\rho = r / \|x_0 - x_1\| > 0.$$

Dann gilt für $|s - t| < \rho$ auch

$$\|x_s - x_t\| = |s - t| \cdot \|x_0 - x_1\| < r;$$

das heisst, $x_s \in B_r(x_t) \subset \Omega_0$, und $B_\rho(t) \subset J_0$. Analog für J_1 . □

Weiter gilt mit (4.4.1) und $\Omega \subset (\Omega_0 \cup \Omega_1)$ auch $I \subset (J_0 \cup J_1)$. Da I gemäss Satz 4.4.1 zusammenhängend, existiert $t_0 \in J_0 \cap J_1 \cap I$, also

$$x_{t_0} \in \Omega \cap \Omega_0 \cap \Omega_1,$$

und $\Omega \cap \Omega_0 \cap \Omega_1 \neq \emptyset$. □

Beispiel 4.4.1. i) \mathbb{R}^n ist konvex, also zusammenhängend.

ii) $B_r(x)$ ist konvex für alle $x \in \mathbb{R}^n$, $r > 0$, also zusammenhängend.

Beweis. Seien $x_0, x_1 \in B_r(x)$, und sei x_t wie in (4.4.1) definiert, $0 \leq t \leq 1$. Schätze ab

$$\|x_t - x\| = \|(1-t)(x_0 - x) + t(x_1 - x)\| \leq (1-t)\|x_0 - x\| + t\|x_1 - x\| < r.$$

Also ist $B_r(x)$ konvex. □

4.5 Relativtopologie

Sei $M \subset \mathbb{R}^n$ beliebig. Setze

$$\tau_M = \{M \cap \Omega; \Omega \subset \mathbb{R}^n \text{ offen}\}.$$

Satz 4.5.1. τ_M ist eine Topologie auf M , die sogenannte **Relativtopologie**.

Beweis. i) Offenbar gilt $\emptyset, M \in \tau_M$.

ii) Seien $M_1 = M \cap \Omega_1$, $M_2 = M \cap \Omega_2 \in \tau_M$, wobei $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ offen. Nach Satz 4.1.1 ist $\Omega = \Omega_1 \cap \Omega_2$ offen, also

$$M_1 \cap M_2 = M \cap (\Omega_1 \cap \Omega_2) \in \tau_M.$$

iii) Falls $M_\iota = M \cap \Omega_\iota \in \tau_M$, wobei $\Omega_\iota \subset \mathbb{R}^n$ offen, $\iota \in I$, so folgt mit Satz 4.1.1 zunächst, dass $\bigcup_{\iota \in I} \Omega_\iota$ offen, also

$$\bigcup_{\iota \in I} M_\iota = M \cap \left(\bigcup_{\iota \in I} \Omega_\iota \right) \in \tau_M.$$

□

Definition 4.5.1. Sei $M \subset \mathbb{R}^n$ beliebig.

i) $E \subset M$ heisst **relativ offen**, falls $E \in \tau_M$.

ii) $F \subset M$ heisst **relativ abgeschlossen**, falls $M \setminus F \in \tau_M$.

Beispiel 4.5.1. i) Sei $M = [0, 1[\subset \mathbb{R}$. Dann ist $[0, \frac{1}{2}[= M \cap] - 1, \frac{1}{2}[$ relativ offen.

ii) $M =]0, 1[\subset \mathbb{R}$ ist relativ offen und abgeschlossen.

Satz 4.5.2. Für $F \subset M$ sind äquivalent:

- i) F ist relativ abgeschlossen;
- ii) $F = M \cap A$, wobei $A \subset \mathbb{R}^n$ abgeschlossen;
- iii) Für alle Folgen $(a_k)_{k \in \mathbb{N}} \subset F$ mit $a_k \rightarrow a \in M$ ($k \rightarrow \infty$) gilt $a \in F$.

Beweis. i) \Rightarrow ii) Sei F relativ abgeschlossen; das heisst, $E = M \setminus F \in \tau_M$. Dann gilt $E = M \cap \Omega$, wobei $\Omega \subset \mathbb{R}^n$ offen. Setze $A = \mathbb{R}^n \setminus \Omega$. Dann ist A abgeschlossen, und

$$F = M \setminus E = M \setminus \Omega = M \cap A.$$

ii) \Rightarrow i) Falls $F = M \cap A$ mit einer abgeschlossenen Menge $A \subset \mathbb{R}^n$, setze $\Omega = \mathbb{R}^n \setminus A$. Dann ist $\Omega \subset \mathbb{R}^n$ offen, und

$$M \setminus F = M \setminus A = M \cap \Omega \in \tau_M.$$

ii) \Rightarrow iii) Sei $F = M \cap A$, wobei $A \subset \mathbb{R}^n$ abgeschlossen, und sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in $F \subset A$ mit $a_k \rightarrow a$ ($k \rightarrow \infty$), wobei $a \in M$. Nach Satz 4.2.1 gilt $a \in A$, also $a \in M \cap A = F$.

iii) \Rightarrow ii) Es genügt zu zeigen $F = M \cap \overline{F}$, wobei $\overline{F} = F \cup \partial F$ der Abschluss von $F \subset \mathbb{R}^n$. Daher genügt es zu zeigen $M \cap \partial F \subset F$. Zu $a \in \partial F \cap M$ gibt es wie im Beweis von Satz 4.2.1 eine Folge $(a_k)_{k \in \mathbb{N}}$ in F mit $a_k \rightarrow a$ ($k \rightarrow \infty$). Da $a \in M$, folgt mit iii) auch $a \in F$. \square

Definition 4.5.2. $C \subset M$ heisst **relativ kompakt**, falls $C = M \cap K$, wobei $K \subset \mathbb{R}^n$ kompakt ist.

Bemerkung 4.5.1. i) Offenbar ist $C \subset M$ relativ kompakt genau dann, wenn C beschränkt und relativ abgeschlossen ist.

ii) Insbesondere gilt für $C = M$: Die Menge M ist relativ kompakt (in M) genau dann, wenn M beschränkt ist.

iii) Im Allgemeinen versagen jedoch das Folgenkriterium und das Heine-Borel-Kriterium.

Beispiel 4.5.2. $M =]0, 1[\subset \mathbb{R}$ ist relativ kompakt. Jedoch besitzt $(a_k)_{k \in \mathbb{N}}$ mit $a_k = \frac{1}{k} \in M$, $k \in \mathbb{N}$, keine in M konvergente Teilfolge, und die offene Überdeckung $(\frac{1}{k}, 1[)_{k \in \mathbb{N}}$ von M besitzt keine endliche Teilüberdeckung.

Bemerkung 4.5.2. Falls jedoch $M \subset \mathbb{R}^n$ abgeschlossen, so gelten Folgen- und Heine-Borel-Kriterium. Vergleiche dazu die Übungen.

Für uns ist hauptsächlich folgender Satz von Interesse.

Satz 4.5.3. Sei $M \subset \mathbb{R}^n$ zusammenhängend, $E \subset M$ relativ offen und relativ abgeschlossen in M . Dann gilt $E = M$, oder $E = \emptyset$.

Beweis. Nach Annahme gilt $E \in \tau_M$ sowie $M \setminus E \in \tau_M$, also existieren offene Mengen $\Omega_1, \Omega_2 \subset \mathbb{R}^n$ mit

$$E = M \cap \Omega_1, \quad M \setminus E = M \cap \Omega_2.$$

Es folgt:

$$M = E \cup (M \setminus E) = M \cap (\Omega_1 \cup \Omega_2) \subset (\Omega_1 \cup \Omega_2)$$

sowie

$$M \cap \Omega_1 \cap \Omega_2 = (M \cap \Omega_1) \cap (M \cap \Omega_2) = E \cap (M \setminus E) = \emptyset;$$

da M nach Annahme zusammenhängend, folgt

$$E = M \cap \Omega_1 = \emptyset \quad \text{oder} \quad M \setminus E = M \cap \Omega_2 = \emptyset, \quad E = M.$$

□

Korollar 4.5.1. \mathbb{R}^n und \emptyset sind die einzigen Teilmengen von \mathbb{R}^n , die sowohl offen als auch abgeschlossen sind.

Beweis. $M = \mathbb{R}^n$ ist zusammenhängend (gemäß Beispiel 4.4.1.i). Somit folgt die Behauptung aus Satz 4.5.3. □

4.6 Hausdorffräume

Zum Schluss ein Exkurs in die Welt abstrakter topologischer Räume. Sei M eine beliebige Menge, $\tau \subset \mathcal{P}(M)$ eine Topologie auf M , das Paar (M, τ) somit ein *topologischer Raum*.

Definition 4.6.1. Sei $a \in M$. Eine Menge $U \subset M$ heisst τ -**Umgebung** von a , falls ein $\Omega \in \tau$ existiert mit $a \in \Omega \subset U$.

Mit Hilfe des Umgebungsbegriffs können wir analog zu Abschnitt 4.2 Konvergenz in beliebigen topologischen Räumen definieren.

Definition 4.6.2. Eine Folge $(a_k)_{k \in \mathbb{N}}$ in M τ -konvergiert gegen $a \in M$ genau dann, wenn für jede τ -Umgebung U von a eine Zahl $k_0 = k_0(U) \in \mathbb{N}$ existiert, so dass $a_k \in U$ für jedes $k \geq k_0$.

Beispiel 4.6.1. i) Sei M beliebig, $\tau = \mathcal{P}(M)$ die diskrete Topologie. Da für jedes $a \in M$ gilt

$$\Omega = \{a\} \in \tau,$$

gilt für jede Folge $(a_k)_{k \in \mathbb{N}}$ in M :

$$a_k \rightarrow a \quad (k \rightarrow \infty) \Leftrightarrow \exists k_0, \forall k \geq k_0 : a_k = a.$$

ii) Sei M beliebig, $\tau = \{\emptyset, M\}$ die triviale Topologie, $(a_k)_{k \in \mathbb{N}}$ eine Folge in M . Für jedes $a \in M$ ist $U = M$ die einzige Umgebung von a in M . Also konvergiert $(a_k)_{k \in \mathbb{N}}$ gegen jedes $a \in M$.

Offenbar ist es für Konvergenzbetrachtungen wichtig, eine genügend “reichhaltige” aber auch nicht zu “feine” Topologie zugrunde zu legen. Für einen “vernünftigen” Konvergenzbegriff sollte der Limes einer konvergenten Folge eindeutig bestimmt sein.

Definition 4.6.3. Eine Topologie auf M erfüllt das **Hausdorffsche Trennungsaxiom**, falls je zwei Punkte $a \neq b$ in M disjunkte Umgebungen $U \ni a$, $V \ni b$, $U \cap V = \emptyset$ besitzen. In diesem Fall heisst (M, τ) ein **Hausdorffraum**.

Beispiel 4.6.2. \mathbb{R}^n ist Hausdorffraum mit der durch $\|\cdot\|$ induzierten Topologie der offenen Mengen.

Beweis. Zu $x \neq y \in \mathbb{R}^n$ wähle $r = \frac{1}{2}\|x - y\|$ mit $B_r(x) \cap B_r(y) = \emptyset$. \square

Satz 4.6.1. Seien (M, τ) ein Hausdorffraum, $(a_k)_{k \in \mathbb{N}}$ eine Folge in M , $a, b \in M$ mit $a_k \rightarrow a$, $a_k \rightarrow b$ ($k \rightarrow \infty$). Dann folgt $a = b$.

Beweis (indirekt). Nimm an, $a \neq b$. Wähle disjunkte Umgebungen U und V von a beziehungsweise b . Für

$$k \geq \max\{k_0(U), k_0(V)\}$$

folgt $a_k \in U \cap V$. Widerspruch. \square

4.7 Ausblick

In diesem Abschnitt haben wir gesehen, dass eine Norm eine Metrik und diese wiederum eine Topologie erzeugt. Es gibt also eine *“Hierarchie von Strukturen”*. Umgekehrt kann man sich fragen, wann eine Topologie *“metrisch”* ist, das heisst, welche Eigenschaften einer Topologie τ dafür kennzeichnend sind, dass diese von einer Metrik induziert wird. Offenbar ist die Bedingung, dass τ Hausdorffsch ist, hierfür notwendig (aber allein nicht hinreichend). Die Suche nach geeigneten *“Metrisierbarkeitskriterien”* ist ein wichtiger Teil der Forschungsrichtung *“mengentheoretische Topologie”*.

Kapitel 5

Stetigkeit

5.1 Grenzwerte von Funktionen

Bisher konnten wir mit Satz 3.2.2 Grenzwerte von Ausdrücken der Form

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

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 $x_k \rightarrow x_0 = (a, b, c, d)$ ($k \rightarrow \infty$).

Allgemein untersuchen wir nun für eine Funktion $f: \Omega \rightarrow \mathbb{R}^n$ 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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$). Der Limes x_0 der Folge $(x_k)_{k \in \mathbb{N}}$ muss dabei nicht notwendig selbst wieder in Ω liegen; nach Satz 4.2.1 gilt aber stets $x_0 \in \overline{\Omega}$.

Beispiel 5.1.1. Sei $f(x) = \frac{x^2-1}{x-1}$, $x \neq 1$. Die Funktion f hat eine ‘‘Definitions-lücke’’ 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 $1 \neq x_k \rightarrow x_0 := 1$ für $k \rightarrow \infty$ erhalten wir so

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

Sei $f: \Omega \rightarrow \mathbb{R}^n$, $x_0 \in \overline{\Omega}$.

Definition 5.1.1. f hat an der Stelle x_0 den **Grenzwert** $a \in \mathbb{R}^n$, falls für jede Folge $(x_k)_{k \in \mathbb{N}}$ in Ω mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$) gilt $f(x_k) \rightarrow a$ ($k \rightarrow \infty$).

Notation: $\lim_{x \rightarrow x_0} f(x) = a$.

Beispiel 5.1.2. Für die Funktion f in Beispiel 5.1.1 gilt $\lim_{x \rightarrow 1} f(x) = 2$.

Bemerkung 5.1.1. Falls $\lim_{x \rightarrow x_0} f(x) =: a$ existiert, und falls zusätzlich $x_0 \in \Omega$, so muss gelten $a = f(x_0)$. (Betrachte die konstante Folge $x_k = x_0 \in \Omega$, $k \in \mathbb{N}$.)

Definition 5.1.2. Sei $\Omega \subset \mathbb{R}^d$, $f: \Omega \rightarrow \mathbb{R}^n$.

i) f heisst **stetig an der Stelle** $x_0 \in \Omega$, falls $\lim_{x \rightarrow x_0} f(x) =: a$ existiert (mit $a = f(x_0)$ gemäss Bemerkung 5.1.1).

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

Beispiel 5.1.3. i) Sei $p: \mathbb{C} \rightarrow \mathbb{C}$ das Polynom

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

mit Koeffizienten $a_l \in \mathbb{C}$, $l = 0, \dots, n$. Nach Satz 3.2.2 gilt für $z_k \rightarrow z_0$ ($k \rightarrow \infty$)

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

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

ii) Eine rationale Funktion p/q mit Polynomen $p, q: \mathbb{C} \rightarrow \mathbb{C}$ 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 5.1.1 gezeigt, können wir die Funktion f an der Stelle $x = 1$ durch $f(1) = 2$ stetig ergänzen.

iv) Sei $\Omega = \mathbb{R} \setminus \{0\}$, $f(x) = 1/x$, $x \neq 0$. Dann gilt für $x_k \rightarrow x_0 \neq 0$ nach Satz 3.2.2 stets

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

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

v) Sei $\Omega = \mathbb{R}$, $\chi_{\mathbb{Q}}: \mathbb{R} \rightarrow \mathbb{R}$ 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 4.1.4.vii) und Satz 4.1.3 eine Folge $(x_k)_{k \in \mathbb{N}} \subset \mathbb{Q}$ mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$) und $\lim_{k \rightarrow \infty} \chi_{\mathbb{Q}}(x_k) = 1 \neq \chi_{\mathbb{Q}}(x_0) = 0$; analog für $x_0 \in \mathbb{Q}$.

vi) Die stückweise konstante Funktion $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ 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.

Für monotone Funktionen f auf \mathbb{R} kann man die Stetigkeitspunkte vollständig charakterisieren.

Satz 5.1.1. Sei $-\infty \leq a < b \leq \infty$, und sei $f:]a, b[\rightarrow \mathbb{R}$ monoton wachsend, das heisst,

$$\forall x, y \in]a, b[: x \leq y \Rightarrow f(x) \leq f(y).$$

Dann existieren für jedes $x_0 \in]a, b[$ die links- und rechtsseitigen Grenzwerte

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

und f ist stetig an der Stelle x_0 genau dann, wenn $f(x_0^-) = f(x_0^+) = f(x_0)$. Analog, falls $f:]a, b[\rightarrow \mathbb{R}$ monoton fallend.

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

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

so ist die Folge $(f(x_k))_{k \in \mathbb{N}}$ monoton wachsend und beschränkt. Gemäss Satz 3.3.2 existiert

$$\lim_{k \rightarrow \infty} f(x_k) =: s.$$

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

Behauptung. $s = \lim_{x \rightarrow x_0, x < x_0} f(x)$.

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

$$\forall k \geq k_0 : s - \epsilon < f(x_k) \leq s.$$

Da $x_{k_0} < x_0$, $y_k \rightarrow x_0$ ($k \rightarrow \infty$), gibt es $k_1 \in \mathbb{N}$ mit

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

Zusammen mit der Monotonie von f folgt

$$\forall k \geq k_1 : s - \epsilon < f(x_{k_0}) \leq f(y_k) \leq \lim_{k \rightarrow \infty} f(x_k) = s;$$

das heisst,

$$f(y_k) \rightarrow s \quad (k \rightarrow \infty).$$

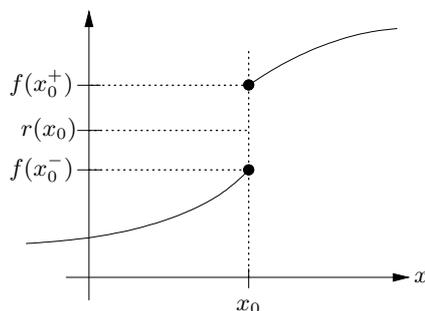
□

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

Satz 5.1.2. Sei $f:]a, b[\rightarrow \mathbb{R}$ monoton wachsend. Dann ist f in höchstens abzählbar vielen Punkten unstetig.

Beweis. Gemäss Satz 5.1.1 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. \square

5.1.1 Lipschitz stetige Funktionen

Wir untersuchen nun wieder Funktionen f auf einer Menge $\Omega \subset \mathbb{R}^n$.

Beispiel 5.1.4. Die Norm $\|\cdot\| : \mathbb{R}^d \rightarrow \mathbb{R}$ ist an jeder Stelle $x_0 \in \Omega$ stetig.

Beweis. Mit der Dreiecksungleichung aus Satz 2.5.2.iii) gilt $\|x\| - \|y\| \leq \|x - y\|$ für alle $x, y \in \mathbb{R}^d$. Nach Vertauschen von x und y folgt die Ungleichung

$$\left| \|x\| - \|y\| \right| \leq \|x - y\|. \quad (5.1.1)$$

Sei $(x_k)_{k \in \mathbb{N}} \subset \mathbb{R}^d$ mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$). Mit (5.1.1) schätzen wir ab

$$\left| \|x_k\| - \|x_0\| \right| \leq \|x_k - x_0\| \rightarrow 0 \quad (k \rightarrow \infty).$$

Es folgt $\lim_{x \rightarrow x_0} \|x\| = \|x_0\|$. \square

Die Normfunktion ist ein Beispiel einer grossen Klasse stetiger Funktionen.

Definition 5.1.3. Eine Funktion $f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ heisst **Lipschitz stetig** mit **Lipschitz-Konstante** L , falls gilt

$$\forall x, y \in \Omega : \|f(x) - f(y)\| \leq L \|x - y\|. \quad (5.1.2)$$

Beispiel 5.1.5. i) Die euklidische Norm $\|\cdot\| : \mathbb{R}^d \rightarrow \mathbb{R}$ ist gemäss (5.1.1) Lipschitz stetig mit Lipschitz-Konstante $L = 1$.

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. Identifizieren wir $\mathbb{R}^n \times \mathbb{R}^n$ mit \mathbb{R}^{2n} , so erhalten wir

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n : \|(x, y)\|_{\mathbb{R}^n \times \mathbb{R}^n}^2 = \|x\|_{\mathbb{R}^n}^2 + \|y\|_{\mathbb{R}^n}^2.$$

Für $(x, y), (x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^n$ folgt

$$\begin{aligned} \|(x+y) - (x_0+y_0)\|_{\mathbb{R}^n}^2 &\leq (\|x-x_0\|_{\mathbb{R}^n} + \|y-y_0\|_{\mathbb{R}^n})^2 \\ &\leq 2(\|x-x_0\|_{\mathbb{R}^n}^2 + \|y-y_0\|_{\mathbb{R}^n}^2) = 2\|(x,y) - (x_0,y_0)\|_{\mathbb{R}^n \times \mathbb{R}^n}^2, \end{aligned}$$

wobei wir die elementare Ungleichung $(a+b)^2 \leq 2a^2 + 2b^2$ für $a, b \in \mathbb{R}$ benutzen, welche sich unmittelbar aus der Youngschen Ungleichung, Satz 2.2.2, ergibt. \square

Satz 5.1.3. Sei $f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ Lipschitz stetig mit Lipschitz-Konstante $L \geq 0$. Dann ist f stetig (ergänzzbar) 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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$). Mit (5.1.2) folgt

$$\|f(x_k) - f(x_0)\| \leq L \|x_k - x_0\| \rightarrow 0 \quad (k \rightarrow \infty).$$

Also ist f an der Stelle $x_0 \in \Omega$ stetig.

Falls $x_0 \in \overline{\Omega} \setminus \Omega$, $(x_k)_{k \in \mathbb{N}} \subset \Omega$ mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$), so folgt mit (5.1.2) analog

$$\|f(x_k) - f(x_l)\| \leq L \|x_k - x_l\| \rightarrow 0 \quad (k, l \rightarrow \infty);$$

$(f(x_k))_{k \in \mathbb{N}}$ ist also Cauchy-Folge, und gemäss Satz 3.6.2 existiert $a = \lim_{k \rightarrow \infty} f(x_k)$.

Der Limes ist unabhängig von der Folge (x_k) , denn für jede weitere Folge $(y_k)_{k \in \mathbb{N}}$ mit $y_k \rightarrow x_0$ ($k \rightarrow \infty$) existiert ebenfalls der Limes $b = \lim_{k \rightarrow \infty} f(y_k)$, und

$$\|a - b\| = \lim_{k \rightarrow \infty} \|f(x_k) - f(y_k)\| \leq \lim_{k \rightarrow \infty} (L \|x_k - y_k\|) = 0.$$

Also ist f an jeder Stelle $x_0 \in \overline{\Omega} \setminus \Omega$ stetig ergänzzbar. \square

Beispiel 5.1.6. 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 $L = \sqrt{2}R$.

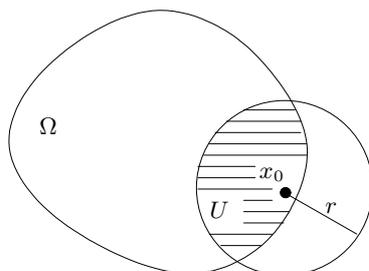
Beweis. Für $x, y, x_0, y_0 \in B_R(0)$ schätze ab mit Satz 2.5.1

$$\begin{aligned} |x \cdot y - x_0 \cdot y_0| &\leq |(x-x_0) \cdot y| + |x_0 \cdot (y-y_0)| \leq \|x-x_0\| \|y\| + \|x_0\| \|y-y_0\| \\ &\leq R(\|x-x_0\| + \|y-y_0\|) \leq \sqrt{2}R \|(x,y) - (x_0,y_0)\|, \end{aligned}$$

analog zu Beispiel 5.1.5.iii). \square

Definition 5.1.4. $U \subset \Omega$ heisst **Umgebung** von x_0 relativ zu Ω , falls $r > 0$ existiert mit

$$B_r(x_0) \cap \Omega \subset U.$$



Beispiel 5.1.6 motiviert die folgende Definition.

Definition 5.1.5. $f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ heisst **lokal Lipschitz stetig**, falls zu jedem $x_0 \in \Omega$ eine Umgebung $U \subset \Omega$ existiert, so dass die auf U eingeschränkte Funktion

$$f|_U: U \ni x \mapsto f(x) \in \mathbb{R}^n$$

auf U Lipschitz stetig ist.

Analog zu Satz 5.1.3 gilt

Satz 5.1.4. Sei $f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ lokal Lipschitz stetig. Dann ist f stetig an jeder Stelle $x_0 \in \Omega$.

Beweis. Zu $x_0 \in \Omega$ wähle eine Umgebung $U \subset \Omega$ 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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$), so gilt für $k \geq k_0(U)$ auch $x_k \in U$; also

$$\|f(x_k) - f(x_0)\| \leq L \|x_k - x_0\| \rightarrow 0 \quad (k \geq k_0, k \rightarrow \infty).$$

□

Bemerkung 5.1.2. Später werden wir sehen, dass Funktionen der Klasse C^1 auf einer offenen Menge $\Omega \subset \mathbb{R}^d$ lokal Lipschitz stetig sind, insbesondere die Funktion $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ mit $f(x) = 1/x$, $x \neq 0$, aus Beispiel 5.1.4.iv). Wie dort gezeigt, ist diese Funktion aber im Gegensatz zur Aussage aus Satz 5.1.3 an der Stelle $x_0 = 0$ nicht stetig ergänzbar.

5.2 Stetigkeitskriterien

Wir können verschiedene Strukturen des \mathbb{R}^d benutzen, um die Stetigkeit einer Funktion f zu charakterisieren.

Satz 5.2.1. Sei $f: \Omega \rightarrow \mathbb{R}^n$, $x_0 \in \Omega$. Es sind äquivalent

- i) (**Folgenkriterium**) f ist stetig an der Stelle x_0 gemäss Definition 5.1.2;
- ii) (**Weierstrass'sches $\epsilon - \delta$ Kriterium**);

$$\forall \epsilon > 0 \exists \delta > 0 \forall x \in \Omega: \|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \epsilon;$$

- iii) (**Topologisches Kriterium**) 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 iii) (indirekt): Nimm an, für eine Umgebung V von $f(x_0)$ in \mathbb{R}^n ist die Menge $U := f^{-1}(V)$ keine Umgebung von x_0 , das heisst

$$\forall r > 0: (B_r(x_0) \cap \Omega) \setminus U \neq \emptyset.$$

Für $r_k = \frac{1}{k}$ wähle $x_k \in (B_{1/k}(x_0) \cap \Omega) \setminus U$, $k \in \mathbb{N}$. Dann gilt $x_k \rightarrow x_0$ ($k \rightarrow \infty$). Mit $x_k \notin U$ erhalten wir jedoch andererseits $f(x_k) \notin V$ für jedes $k \in \mathbb{N}$; also $f(x_k) \not\rightarrow f(x_0)$ ($k \rightarrow \infty$) im Widerspruch zu Annahme i).

iii) \Rightarrow ii): Sei $\epsilon > 0$ vorgegeben. Zu $V = B_\epsilon(f(x_0))$ gibt es nach iii) ein $\delta > 0$ mit $B_\delta(x_0) \cap \Omega \subset f^{-1}(V)$; das heisst,

$$\forall x \in \Omega : \|x - x_0\| < \delta \Rightarrow \|f(x) - f(x_0)\| < \epsilon.$$

ii) \Rightarrow i): Sei $(x_k)_{k \in \mathbb{N}} \subset \Omega$ mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$). Zu $\epsilon > 0$ wähle $\delta > 0$ gemäss ii), dazu k_0 mit

$$\forall k \geq k_0 : \|x_k - x_0\| < \delta.$$

Mit Annahme ii) folgt

$$\forall k \geq k_0 : \|f(x_k) - f(x_0)\| < \epsilon;$$

das heisst, $f(x_k) \rightarrow f(x_0)$ ($k \rightarrow \infty$). \square

Beispiel 5.2.1. Mit Satz 5.2.1 erkennt man nun sofort, dass $f = \chi_{\mathbb{Q}}: \mathbb{R} \rightarrow \mathbb{R}$ 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

$$f^{-1}\left(\left]\frac{1}{2}, \frac{3}{2}\right[\right) = \mathbb{Q}$$

ist wegen Beispiel 4.1.4.v) keine Umgebung von x_0 . Analog ist f in keiner Stelle $x_0 \notin \mathbb{Q}$ stetig, denn

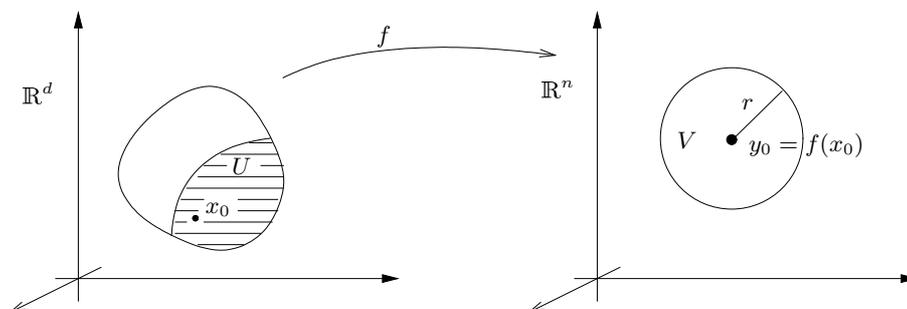
$$f^{-1}\left(\left]-\frac{1}{2}, \frac{1}{2}\right[\right) = \mathbb{R} \setminus \mathbb{Q}$$

hat leeres Inneres; vergleiche Beispiel 4.1.4.iv).

Aus Satz 5.2.1 ergibt sich nun das folgende, äusserst elegante topologische Kriterium für die Stetigkeit einer Funktion $f: \Omega \rightarrow \mathbb{R}^n$ in *allen* Punkten $x_0 \in \Omega$.

Satz 5.2.2. Für $f: \Omega \rightarrow \mathbb{R}^n$ 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 $A = f^{-1}(F)$ jeder abgeschlossenen Menge $F \subset \mathbb{R}^n$ ist relativ abgeschlossen.



Beweis. i) \Rightarrow ii): Sei $V \subset \mathbb{R}^n$ offen, $x_0 \in U = f^{-1}(V)$. Da f stetig an der Stelle x_0 , ist U nach Satz 5.2.1 eine Umgebung von x_0 relativ zu Ω . Da $x_0 \in U$ beliebig, ist U relativ offen.

ii) \Rightarrow iii): Sei $F \subset \mathbb{R}^n$ abgeschlossen. Dann ist $V = \mathbb{R}^n \setminus F$ offen, mit ii) also $U = f^{-1}(V)$ relativ offen, und mit Definition 4.5.1.ii) folgt

$$A = f^{-1}(F) = f^{-1}(\mathbb{R}^n \setminus V) = \Omega \setminus f^{-1}(V) = \Omega \setminus U$$

ist relativ abgeschlossen.

iii) \Rightarrow ii): analog.

ii) \Rightarrow 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 5.2.1 ist f also stetig an der Stelle x_0 . \square

Mit Satz 5.2.1 erkennen wir, dass Stetigkeit eine lokale Eigenschaft ist.

Satz 5.2.3. Seien $f, g : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ und sei f stetig an der Stelle $x_0 \in \Omega$. Weiter sei $f = g$ auf einer Umgebung U_0 von $x_0 \in \Omega$. Dann ist g stetig in x_0 .

Beweis. Sei V Umgebung von $f(x_0) = g(x_0)$. Da f stetig in x_0 , ist $U = f^{-1}(V)$ nach Satz 5.2.1 Umgebung von x_0 , also auch $U \cap U_0 \subset g^{-1}(V)$. \square

5.3 Stetige Funktionen

Sei $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$.

Definition 5.3.1. f heisst stetig auf Ω , falls f in jedem Punkt $x_0 \in \Omega$ stetig ist.

Beispiel 5.3.1. 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 $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ lokal Lipschitz stetig, so ist f stetig auf Ω .

iv) Ist $f : \Omega \rightarrow \mathbb{R}^n$ stetig, $U \subset \Omega$, so ist auch die eingeschränkte Funktion $f|_U : U \rightarrow \mathbb{R}^n$ stetig.

v) Nach Beispiel 5.1.3.vi) ist die stückweise konstante Funktion

$$g = a\chi_{]-\infty, 0[} + b\chi_{]0, \infty[} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$$

stetig (auf $\mathbb{R} \setminus \{0\}$; es kommt auf den Definitionsbereich an!)

Bemerkung 5.3.1. i) Insbesondere das letzte Beispiel illustriert einen wichtigen Aspekt unseres Stetigkeitsbegriffs. Die Funktion g ist stetig, obwohl der Graph von g an der Stelle $x_0 = 0$ einen Sprung macht; die ‘‘Sprungstelle’’ $x_0 = 0$ gehört *nicht* zum Definitionsbereich von g .

ii) Der heutige, auf Cauchy-Weierstrass zurückgehende Stetigkeitsbegriff unterscheidet sich grundlegend von dem Eulerschen, der ‘‘kontinuierliche’’ Funktionen mit Abbildungen gleichsetzt, die durch einen geschlossenen ‘‘analytischen’’

Funktionsausdruck definiert sind. Auf der Mittelschule ist eine präzise Definition in der Regel nicht möglich, und die verschiedenen Vorstellungen verschwimmen zu einem diffusen Bild. Der tiefere Grund hierfür ist die Tatsache, dass Stetigkeit im Kern ein *topologischer* Begriff ist.

Satz 5.3.1. *Sind $f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ und $g: \mathbb{R}^n \rightarrow \mathbb{R}^l$ stetig, so ist auch deren Komposition $g \circ f: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^l$ stetig.*

Beweis. Sei $W \subset \mathbb{R}^l$ offen. Da g stetig, ist $V = g^{-1}(W)$ offen; da f stetig, ist analog auch $U = f^{-1}(V)$ offen. Also ist

$$U = f^{-1}(V) = f^{-1}(g^{-1}(W)) = (g \circ f)^{-1}(W)$$

offen, und die Behauptung folgt mit Satz 5.2.2. \square

Mit Beispiel 5.1.5.iii), Beispiel 5.1.6 und Satz 5.3.1 folgt

Satz 5.3.2. *Sind $f, g: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig, so sind auch die Funktionen $f + g$ und αf stetig, wobei $\alpha \in \mathbb{R}$ beliebig. Die stetigen Funktionen $f: \Omega \rightarrow \mathbb{R}^n$ bilden also einen \mathbb{R} -Vektorraum, den Raum*

$$C^0(\Omega; \mathbb{R}^n) = \{f: \Omega \rightarrow \mathbb{R}^n; f \text{ ist stetig}\}.$$

5.3.1 Stetigkeit und Kompaktheit.

Stetige Funktionen auf kompakten Mengen haben eine Reihe von besonderen Qualitäten.

Satz 5.3.3. *Sei $K \subset \mathbb{R}^d$ kompakt, $f: K \rightarrow \mathbb{R}^n$ stetig. Dann ist $f(K)$ kompakt.*

Beweis. Sei $(V_\iota)_{\iota \in I}$ offene Überdeckung von $f(K)$. Setze

$$U_\iota = f^{-1}(V_\iota), \quad \iota \in I.$$

Da V_ι offen und f stetig, ist U_ι offen, $\iota \in I$. Weiter gilt

$$K = f^{-1}(f(K)) \subset f^{-1}\left(\bigcup_{\iota \in I} V_\iota\right) = \bigcup_{\iota \in I} U_\iota.$$

Da K kompakt, genügen nach Satz 4.3.1 endlich viele $(U_{\iota_\ell})_{1 \leq \ell \leq L}$, um K zu überdecken. Es folgt

$$f(K) \subset \bigcup_{\ell=1}^L f(U_{\iota_\ell}) = \bigcup_{\ell=1}^L V_{\iota_\ell},$$

und $(V_{\iota_\ell})_{1 \leq \ell \leq L}$ ist endliche Teilüberdeckung von $(V_\iota)_{\iota \in I}$. Somit ist $f(K)$ kompakt nach Satz 4.3.1. \square

Übung: Beweise Satz 5.3.3 mit dem Folgenkriterium.

Lemma 5.3.1. 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 $a_k \rightarrow \inf K \geq -\infty$ ($k \rightarrow \infty$). Da K kompakt, gibt es eine Teilfolge $\Lambda \subset \mathbb{N}$, $a \in K$ mit $a_k \rightarrow a$ ($k \rightarrow \infty$, $k \in \Lambda$), und

$$a = \lim_{k \rightarrow \infty, k \in \Lambda} a_k = \inf K = \min K.$$

Insbesondere ist $\inf K > -\infty$. Analog für b . □

Mit Satz 5.3.3 folgt:

Korollar 5.3.1. Sei $K \subset \mathbb{R}^d$ kompakt, $f: K \rightarrow \mathbb{R}$ stetig. Dann nimmt die Funktion f ihr Supremum und ihr Infimum an.

Beispiel 5.3.2. i) Sei $f: [a, b] \rightarrow \mathbb{R}$ 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 $f: K \rightarrow \mathbb{R}^n$ ein $x_0 \in K$ mit

$$\|f(x_0)\| = \max_{x \in K} \|f(x)\|.$$

Beweis. Die Funktion $F = \|\cdot\| \circ f: K \rightarrow \mathbb{R}$ ist stetig wegen Satz 5.3.1 und Beispiel 5.1.4. □

5.3.2 Eine Norm für $C^0(K; \mathbb{R}^n)$.

Die folgende Definition orientiert sich an den Eigenschaften der euklidischen Norm aus Satz 2.5.2.

Definition 5.3.2. Sei X ein \mathbb{R} -Vektorraum. Eine **Norm** auf X ist eine Abbildung $\|\cdot\|: X \rightarrow \mathbb{R}$ mit den Eigenschaften

- i) *Definitheit:* $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$, für beliebiges $x \in X$;
- ii) *Positive Homogenität:* $\|\alpha x\| = |\alpha| \|x\|$ für alle $x \in X$, $\alpha \in \mathbb{R}$;
- iii) *Dreiecks-Ungleichung:* $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in X$.

Sei $K \subset \mathbb{R}^d$ kompakt, $f: K \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig. Nach Beispiel 5.3.2.ii) existiert

$$\|f\|_{C^0(K; \mathbb{R}^n)} = \sup_{x \in K} \|f(x)\| < \infty. \quad (5.3.1)$$

Satz 5.3.4. Sei $K \subset \mathbb{R}^d$ kompakt. Dann definiert (5.3.1) eine Norm auf $C^0(K; \mathbb{R}^d)$, die Supremumsnorm.

Notation: Oft schreiben wir $\|f\|_{C^0}$ statt $\|f\|_{C^0(K; \mathbb{R}^n)}$, wenn Definitionsbereich und Bildbereich von f aus dem Kontext klar sind.

Beweis. i) Offenbar gilt $\|f\|_{C^0} \geq 0$ für jedes $f \in C^0$, und $\|f\|_{C^0} = 0$ genau dann, wenn $f \equiv 0$.

ii) Für $\lambda \in \mathbb{R}$, $f \in C^0(K; \mathbb{R}^n)$ gilt:

$$\|\lambda f\|_{C^0} = \sup_{x \in K} \|\lambda f(x)\| = |\lambda| \sup_{x \in K} \|f(x)\| = |\lambda| \|f\|_{C^0}.$$

iii) Für $f, g \in C^0(K; \mathbb{R}^n)$ und beliebiges $x \in K$ gilt

$$\|(f + g)(x)\| = \|f(x) + g(x)\| \leq \|f(x)\| + \|g(x)\| \leq \|f\|_{C^0} + \|g\|_{C^0}.$$

Es folgt:

$$\|f + g\|_{C^0} = \sup_{x \in K} \|f(x) + g(x)\| \leq \|f\|_{C^0} + \|g\|_{C^0}.$$

□

5.3.3 Homöomorphismen

Satz 5.3.5. Sei $K \subset \mathbb{R}^d$ kompakt, $f : K \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig und injektiv. Setze

$$\tilde{K} = f(K) \subset \mathbb{R}^n.$$

Dann ist $f^{-1} : \tilde{K} \subset \mathbb{R}^n \rightarrow K \subset \mathbb{R}^d$ stetig.

Beweis. Betrachte $g = f^{-1} : \tilde{K} \rightarrow K$. Sei $A \subset \mathbb{R}^d$ abgeschlossen. Dann ist $K \cap A$ kompakt, das Urbild

$$g^{-1}(A) = g^{-1}(K \cap A) = f(K \cap A)$$

von A unter g nach Satz 5.3.3 also ebenfalls kompakt, insbesondere abgeschlossen. Nach Satz 5.2.2 ist f^{-1} stetig. □

Alternativ kann man Satz 5.3.5 auch bequem mit dem Folgenkriterium beweisen.

Definition 5.3.3. Eine bijektive Abbildung $f : \Omega \subset \mathbb{R}^d \rightarrow \tilde{\Omega} \subset \mathbb{R}^n$, für die sowohl f als auch $\tilde{f} = f^{-1}$ stetig sind, heißt ein **Homöomorphismus**.

Bemerkung 5.3.2. Nach Satz 5.3.5 induziert eine stetige injektive Abbildung $f : [0, 1] \rightarrow \mathbb{R}^2$ einen Homöomorphismus von $I = [0, 1]$ auf das Bild $f(I) \subset \mathbb{R}^2$; die Mengen I und $f(I)$ sind “topologisch” nicht unterscheidbar. Derartige “Jordankurven” können jedoch höchst kompliziert sein und eine Menge positiven zweidimensionalen Masses überdecken. (“Peano-Kurven”; vergleiche dazu die Übungen.) Es stellt sich dann die Frage, ob der *Dimensionsbegriff* eine “topologische Invariante” ist. Dazu gilt der folgende Satz.

Satz 5.3.6. (Brouwer: “invariance of domain”) Sei $f : A \rightarrow B$ ein Homöomorphismus der Mengen $A \subset \mathbb{R}^d$, $B \subset \mathbb{R}^n$, wobei $A^\circ \neq \emptyset \neq B^\circ$. Dann gilt $d = n$.

5.4 Äquivalenz von Normen auf \mathbb{R}^d .

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 2.5.2 kann man auf \mathbb{R}^d jedoch auch andere Normen mit den Eigenschaften gemäss Definition 5.3.2 erklären.

Beispiel 5.4.1. Für $1 \leq p < \infty$ definiert

$$\|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{1/p}, \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d,$$

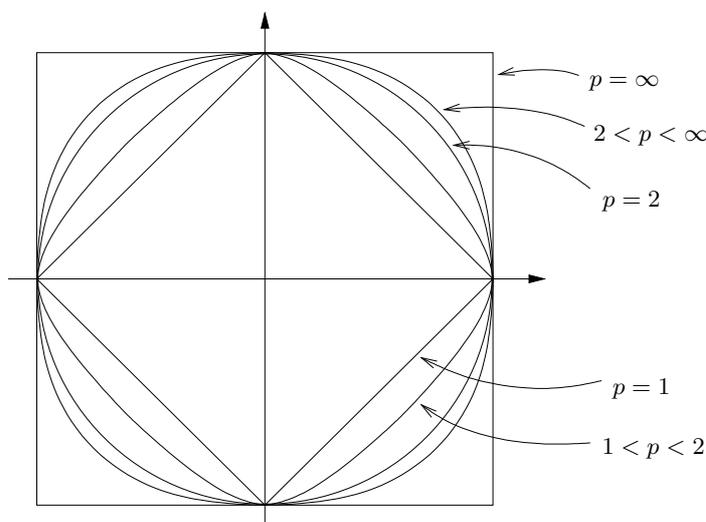
eine Norm auf \mathbb{R}^d , ebenso für $p = \infty$ der Ausdruck

$$\|x\|_\infty = \max_{1 \leq i \leq d} |x_i|, \quad x = (x_1, \dots, x_d) \in \mathbb{R}^d.$$

Für $p = 1$ oder $p = \infty$ ist die Dreiecksungleichung offensichtlich; in den Fällen $1 < p < 2$ oder $2 < p < \infty$ werden wir die Dreiecksungleichung später beweisen.

Offenbar gilt für alle $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, $1 \leq p \leq \infty$:

$$\|x\|_\infty = \max_i |x_i| \leq \|x\|_p \leq \sum_i |x_i| = \|x\|_1 \leq d \|x\|_\infty. \quad (5.4.1)$$



Definition 5.4.1. Zwei Normen $\|\cdot\|^{(1)}, \|\cdot\|^{(2)} : \mathbb{R}^d \rightarrow \mathbb{R}$ heissen **äquivalent**, falls $C > 0$ existiert mit

$$\forall x \in \mathbb{R}^d : \frac{1}{C} \|x\|^{(1)} \leq \|x\|^{(2)} \leq C \|x\|^{(1)}. \quad (5.4.2)$$

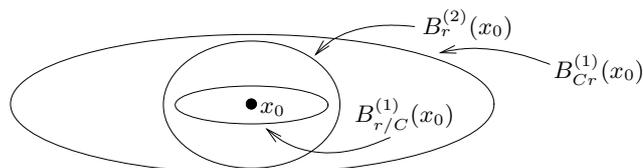
Beispiel 5.4.2. Je zwei der in Beispiel 5.4.1 definierten Normen $\|\cdot\|_p, \|\cdot\|_q$, $1 \leq p, q \leq \infty$ sind wegen (5.4.1) äquivalent.

Bemerkung 5.4.1. 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 (5.4.2)

$$B_{r/C}^{(1)}(x_0) \subset B_r^{(2)}(x_0) \subset B_{Cr}^{(1)}(x_0);$$

also ist $x_0 \in \Omega$ innerer Punkt von Ω bzgl. $\|\cdot\|^{(2)}$ genau dann, wenn $x_0 \in \Omega$ innerer Punkt ist bzgl. $\|\cdot\|^{(1)}$. \square



Satz 5.4.1. Je zwei Normen $\|\cdot\|^{(1)}, \|\cdot\|^{(2)} : \mathbb{R}^d \rightarrow \mathbb{R}$ sind äquivalent.

Beweis. Es genügt, den Fall $\|\cdot\|^{(2)} = \|\cdot\|_2$ zu betrachten, wobei $\|\cdot\|_2$ wie in Beispiel 5.4.1 die euklidische Norm bezeichnet. Der Kürze halber schreiben wir auch $\|\cdot\|$ anstelle von $\|\cdot\|^{(1)}$.

Behauptung. $\|\cdot\| : \mathbb{R}^d \rightarrow \mathbb{R}$ ist Lipschitz stetig bzgl. $\|\cdot\|_2$.

Beweis. Für $x, y \in \mathbb{R}^d$ mit $y - x = z = (z_1, \dots, z_d) = \sum_{i=1}^d z_i e_i$ schätze ab

$$\begin{aligned} \left| \|x\| - \|y\| \right| &\leq \|x - y\| = \|z\| \leq \sum_{i=1}^d \|z_i e_i\| = \sum_{i=1}^d |z_i| \|e_i\| \\ &\leq C \sum_{i=1}^d |z_i| \leq Cd \max_i |z_i| \leq Cd \|z\|_2 = Cd \|x - y\|_2. \end{aligned}$$

\square

Nach Satz 4.1.4.i) ist $S^{d-1} := \partial B_1(0) = \{x \in \mathbb{R}^d; \|x\|_2 = 1\} \subset \mathbb{R}^d$ abgeschlossen. Da S^{d-1} offenbar auch beschränkt, ist S^{d-1} kompakt. Gemäss Korollar 5.3.1 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 der Definitheit von $\|\cdot\|$ gemäss Definition 5.3.2. Wähle $C = \max\{\Lambda, \frac{1}{\lambda}\}$. Es folgt

$$\forall x \in \mathbb{R}^d \setminus \{0\} : \frac{1}{C} \leq \lambda \leq \left\| \frac{x}{\|x\|_2} \right\| = \frac{\|x\|}{\|x\|_2} \leq \Lambda \leq C,$$

wie gewünscht. \square

Im folgenden schreiben wir zur besseren Übersichtlichkeit stets $|x|$ anstelle von $\|x\|$ für die euklidische Norm von $x \in \mathbb{R}^d$.

5.5 Stetige Fortsetzung

Sei $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig.

Definition 5.5.1. $F : \overline{\Omega} \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ heisst eine **stetige Fortsetzung** von f , falls F stetig auf $\overline{\Omega}$ und $F|_{\Omega} = f$.

Unter welchen Bedingungen kann man f zu einer stetigen Funktion auf $\overline{\Omega}$ fortsetzen? Offenbar ist die Existenz des Grenzwertes

$$\lim_{x \rightarrow x_0, x \in \Omega} f(x) =: F(x_0)$$

notwendig und hinreichend für die stetige Ergänzung von f im Punkt x_0 .

Beispiel 5.5.1. i) Falls $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ auf Ω Lipschitz stetig ist, so ist f gemäss Satz 5.1.3 auf $\overline{\Omega}$ stetig ergänzbar.

ii) Die Funktion $f(x) = \frac{1}{x} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ ist für $x \downarrow 0$ unbeschränkt, daher nicht stetig auf \mathbb{R} fortsetzbar; vergleiche Beispiel 5.1.3.iv).

Eine allgemeine hinreichende Bedingung für stetige Fortsetzbarkeit liefert der folgende Begriff.

Definition 5.5.2. $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ heisst **gleichmässig stetig** auf Ω , falls gilt:

$$\forall \epsilon > 0 \exists \delta = \delta(\epsilon) > 0 \forall x, y \in \Omega : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon. \quad (5.5.1)$$

Vergleiche dies mit dem Weierstrass-Kriterium für Stetigkeit nach Satz 5.2.1:

$$\forall \epsilon > 0 \forall x \in \Omega \exists \delta = \delta(\epsilon, x) > 0 \forall y \in \Omega : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Beispiel 5.5.2. Lipschitz stetige Funktionen (mit Lipschitz-Konstante L) sind gleichmässig stetig. (Zu $\epsilon > 0$ wähle $\delta = \epsilon/L$.)

Satz 5.5.1. Sei $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ gleichmässig stetig. Dann existiert genau eine stetige Funktion $F : \overline{\Omega} \rightarrow \mathbb{R}^n$ mit $F|_{\Omega} = f$.

Beweis. Sei $x_0 \in \overline{\Omega} \setminus \Omega$, $(x_k)_{k \in \mathbb{N}}$ eine Folge in Ω mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$).

Behauptung 1. $(f(x_k))_{k \in \mathbb{N}}$ ist eine Cauchy-Folge.

Beweis. Sei $\epsilon > 0$ vorgegeben. Wähle $\delta = \delta(\epsilon) > 0$ gemäss (5.5.1) und dazu $k_0 = k_0(\delta)$ gemäss Satz 3.5.1 mit

$$\forall k, \ell \geq k_0 : |x_k - x_\ell| < \delta.$$

Mit (5.5.1) folgt dann

$$\forall k, \ell \geq k_0 : |f(x_k) - f(x_\ell)| < \epsilon,$$

wie gewünscht. □

Also existiert $a := \lim_{k \rightarrow \infty} f(x_k)$.

Behauptung 2. $a = \lim_{x \rightarrow x_0, x \in \Omega} f(x) =: F(x_0)$.

Beweis. Sei $(y_k)_{k \in \mathbb{N}} \subset \Omega$ mit $y_k \rightarrow x_0$ ($k \rightarrow \infty$), und sei $\epsilon > 0$ vorgegeben. Wähle dazu $\delta = \delta(\epsilon) > 0$ wie in (5.5.1), $k_1 = k_1(\delta) \in \mathbb{N}$ so, dass für jedes $k \geq k_1$ gilt $|x_k - x_0|, |y_k - x_0| < \delta/2$, also auch

$$\forall k, \ell \geq k_1 : |y_k - x_\ell| \leq |y_k - x_0| + |x_\ell - x_0| < \delta.$$

Dann folgt mit (5.5.1) für $k \geq k_1$ und beliebiges $\ell \geq k_1$:

$$|f(y_k) - a| \leq |f(y_k) - f(x_\ell)| + |f(x_\ell) - F(x_0)| < \epsilon + |f(x_\ell) - F(x_0)|.$$

Nach Grenzübergang $\ell \rightarrow \infty$ erhalten wir $|f(y_k) - a| \leq \epsilon$, wie gewünscht. \square

Setze $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ fort zu $F : \overline{\Omega} \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ gemäss Behauptung 2.

Behauptung 3. $F : \overline{\Omega} \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ ist (gleichmässig) stetig.

Beweis. Zu $\epsilon > 0$ wähle $\delta = \delta(\epsilon) > 0$ wie in (5.5.1) für f . Seien $x_0, y_0 \in \overline{\Omega}$ mit $|x_0 - y_0| < \delta$. Wähle $x, y \in \Omega$ mit $|x - y| < \delta$ und

$$|F(x_0) - f(x)| + |F(y_0) - f(y)| < \epsilon.$$

Dies ist möglich gemäss Behauptung 2. Mit (5.5.1) folgt:

$$|F(x_0) - F(y_0)| \leq |F(x_0) - f(x)| + |f(x) - f(y)| + |F(y_0) - f(y)| < 2\epsilon.$$

\square

Schliesslich sei $F_1 : \overline{\Omega} \rightarrow \mathbb{R}^n$ stetig mit $F_1|_{\Omega} = f$. Dann erhalten wir mit Stetigkeit von F_1 und Behauptung 2 für alle $x_0 \in \overline{\Omega}$ die Identität

$$F_1(x_0) = \lim_{x \rightarrow x_0, x \in \Omega} F_1(x) = \lim_{x \rightarrow x_0, x \in \Omega} f(x) = F(x_0),$$

und $F_1 = F$. Die stetige Fortsetzung ist also eindeutig bestimmt. \square

Definition 5.5.3. Seien $\Omega \subset M \subset \mathbb{R}^d$. Ω heisst **dicht** in M , falls $\overline{\Omega} = M$.

Beispiel 5.5.3. i) Sei $F : \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig und es gelte $F|_{\Omega} \equiv 0$, wobei $\Omega \subset \mathbb{R}^d$ dicht in \mathbb{R}^d . Dann gilt $F \equiv 0$ auf \mathbb{R}^d . Insbesondere verschwindet jede stetige Funktion $F : \mathbb{R} \rightarrow \mathbb{R}^n$ mit $F|_{\mathbb{Q}} \equiv 0$ identisch.

ii) Gemäss Satz 3.9.2 verschwindet die Funktion $F(x) = \text{Exp}(x) - e^x : \mathbb{R} \rightarrow \mathbb{R}$ auf \mathbb{Q} . Somit ist Exp die eindeutig bestimmte stetige Fortsetzung der elementar definierten Exponentialfunktion.

Beweis. Setze $f = F|_{\Omega} : \Omega \rightarrow \mathbb{R}^n$. Dann ist $f \equiv 0$ gleichmässig stetig. Offenbar sind sowohl F als auch die Funktion $F_0 \equiv 0 : \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetige Fortsetzungen von f . Mit Satz 5.5.1 folgt $F = F_0 \equiv 0$. \square

Satz 5.5.2. Sei $f : K \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig, wobei K kompakt. Dann ist f gleichmässig stetig.

Folgerung. Falls $\Omega \subset \mathbb{R}^d$ beschränkt, also $\overline{\Omega}$ kompakt, so ist gemäss den Sätzen 5.5.1 und 5.5.2 die gleichmässige Stetigkeit von $f : \Omega \rightarrow \mathbb{R}^n$ nicht nur hinreichend sondern auch notwendig für die Existenz einer stetigen Fortsetzung von f .

Beweis von Satz 5.5.2. Sei $\epsilon > 0$ vorgegeben. Da f stetig, gibt es zu jedem $x_0 \in K$ ein $\delta = \delta(\epsilon, x_0) > 0$ mit

$$\forall x \in K : |x - x_0| < 2\delta \Rightarrow |f(x) - f(x_0)| < \frac{\epsilon}{2}. \quad (5.5.2)$$

Die Bälle $(B_{\delta(\epsilon, x_0)}(x_0))_{x_0 \in K}$ bilden eine offene Überdeckung von K . Da K nach Annahme kompakt, gibt es $x_1, \dots, x_L \in K$ mit

$$K \subset \bigcup_{\ell=1}^L B_{\delta_\ell}(x_\ell), \quad \delta_\ell = \delta(\epsilon, x_\ell) > 0, \quad 1 \leq \ell \leq L.$$

Wähle

$$\delta = \min_{1 \leq \ell \leq L} \delta_\ell > 0.$$

Seien $x, y \in K$ mit $|x - y| < \delta$. Wähle $\ell \in \{1, \dots, L\}$ mit $x \in B_{\delta_\ell}(x_\ell)$. Da $|x - y| < \delta \leq \delta_\ell$, folgt $x, y \in B_{2\delta_\ell}(x_\ell)$. Mit (5.5.2) erhalten wir

$$|f(x) - f(x_\ell)|, |f(y) - f(x_\ell)| < \frac{\epsilon}{2},$$

also

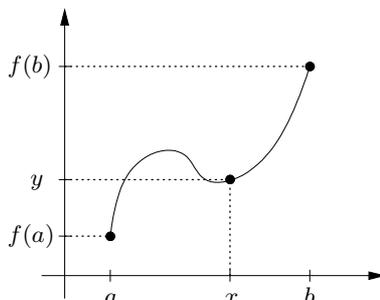
$$|f(x) - f(y)| \leq |f(x) - f(x_\ell)| + |f(y) - f(x_\ell)| < \epsilon.$$

□

5.6 Zwischenwertsatz und Folgerungen

Die Stetigkeit einer Funktion hat noch weitere Konsequenzen. Wir betrachten zunächst reelle Funktionen auf Intervallen.

Satz 5.6.1. Seien $-\infty < a < b < \infty$, und sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $f(a) \leq f(b)$. Dann gibt es zu jedem $y \in [f(a), f(b)]$ ein $x \in [a, b]$ mit $f(x) = y$.



Nach Satz 4.4.1 ist Satz 5.6.1 ein Spezialfall des folgenden Satzes.

Satz 5.6.2. Sei $f : \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig, Ω zusammenhängend. Dann ist $f(\Omega)$ zusammenhängend.

Beweis. Seien $V_1, V_2 \subset \mathbb{R}^n$ offen mit

$$f(\Omega) \subset V_1 \cup V_2, \quad f(\Omega) \cap V_i \neq \emptyset, \quad i = 1, 2.$$

Setze $U_i = f^{-1}(V_i)$, $i = 1, 2$. Da f stetig ist, sind U_1, U_2 relativ offen und

$$\Omega \subset U_1 \cup U_2, \quad \Omega \cap U_i \neq \emptyset, \quad i = 1, 2.$$

Da Ω nach Annahme zusammenhängend ist, gibt es $x_0 \in \Omega \cap U_1 \cap U_2$; also

$$f(x_0) \in f(\Omega) \cap V_1 \cap V_2,$$

und $f(\Omega) \cap V_1 \cap V_2 \neq \emptyset$, wie gewünscht. \square

Beweis von Satz 5.6.1. Gemäss Satz 4.4.1 ist $I = [a, b]$ zusammenhängend, nach Satz 5.6.2 also auch das Bild $f(I) \subset \mathbb{R}$. Wieder nach Satz 4.4.1 ist dann $f(I)$ ein Intervall; insbesondere $[f(a), f(b)] \subset f(I)$. \square

Bemerkung 5.6.1. Alternativ kann man zum Beweis von Satz 5.6.1 das “Bisektionsverfahren” benutzen. Definiere $a_1 = a$, $b_1 = b$. Setze

$$a_2 = a = a_1, \quad b_2 = \frac{a+b}{2}, \quad \text{falls } f\left(\frac{a+b}{2}\right) \geq y,$$

bzw. setze

$$a_2 = \frac{a+b}{2}, \quad b_2 = b = b_1, \quad \text{falls } f\left(\frac{a+b}{2}\right) < y,$$

so dass $f(a_2) < y \leq f(b_2)$ und $|a_2 - b_2| = 2^{-1}|a - b|$. Allgemein seien a_1, \dots, a_k sowie b_1, \dots, b_k bereits definiert mit

$$a_1 \leq \dots \leq a_k \leq b_k \leq \dots \leq b_1$$

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 3.3.2 gibt es

$$\bar{a} = \lim_{k \rightarrow \infty} a_k \leq \bar{b} = \lim_{k \rightarrow \infty} b_k,$$

und mit Satz 3.2.2 folgt

$$|\bar{a} - \bar{b}| = \lim_{k \rightarrow \infty} |a_k - b_k| = 0.$$

Das heisst, $\bar{a} = \bar{b} =: x \in [a, b]$. Da f stetig, folgt mit Satz 3.2.2 schliesslich

$$y \leq \lim_{k \rightarrow \infty} f(b_k) = f(x) = \lim_{k \rightarrow \infty} f(a_k) \leq y;$$

also $f(x) = y$, wie gewünscht.

Beispiel 5.6.1. i) Sei $p: \mathbb{R} \rightarrow \mathbb{R}$ ein Polynom von ungeradem Grad. Dann hat p eine Nullstelle.

Beweis. Beachte $|p(x)| \rightarrow \infty$ für $|x| \rightarrow \infty$. OBdA $p(x) \rightarrow \infty$ für $x \rightarrow \infty$. (Sonst betrachte $\tilde{p} = -p$.) Da der Grad von p ungerade ist, folgt $p(x) \rightarrow -\infty$ für $x \rightarrow -\infty$, und die Behauptung folgt mit Satz 5.6.1 und Beispiel 5.3.1.i). \square

ii) Jede 3×3 Matrix A mit Koeffizienten in \mathbb{R} hat mindestens einen reellen Eigenwert.

Beweis. Das charakteristische Polynom p von A hat Grad 3; die Nullstellen von p sind genau die Eigenwerte von A . (Siehe: "Lineare Algebra"). \square

Eine weitere Anwendung von Satz 5.6.1 ist der folgende Fixpunktsatz.

Satz 5.6.3. Sei $f: [a, b] \rightarrow [a, b]$ stetig. Dann gibt es $x_0 \in [a, b]$ mit $f(x_0) = x_0$.

Beweis. Definiere die Funktion $g(x) = x - f(x)$, $x \in [a, b]$. Dann ist $g: [a, b] \rightarrow \mathbb{R}$ stetig mit

$$g(a) = a - f(a) \leq 0 \leq b - f(b) = g(b).$$

Nach Satz 5.6.1 existiert $x_0 \in [a, b]$ mit $g(x_0) = 0$; das heisst, $f(x_0) = x_0$. \square

5.6.1 Monotone Funktionen

Definition 5.6.1. $f: [a, b] \rightarrow \mathbb{R}$ heisst **streng monoton wachsend**, falls gilt

$$a \leq x < y \leq b \Rightarrow f(x) < f(y). \quad (5.6.1)$$

Satz 5.6.4. Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und streng monoton wachsend. Setze $f(a) = c$, $f(b) = d$. Dann ist $f: [a, b] \rightarrow [c, d]$ bijektiv, und f^{-1} ist stetig.

Beweis. f ist injektiv mit $f([a, b]) \subset [c, d]$ wegen (5.6.1) und surjektiv gemäss Satz 5.6.1, also bijektiv. Die Stetigkeit von f^{-1} folgt mit Satz 5.3.5. \square

Beispiel 5.6.2. Sei $f: [0, 1] \cup [2, 3] \rightarrow \mathbb{R}$ gegeben mit

$$f(x) = \begin{cases} x, & 0 \leq x < 1, \\ x - 1, & 2 \leq x \leq 3. \end{cases}$$

f ist stetig und streng monoton, aber $f^{-1}: [0, 2] \rightarrow \mathbb{R}$ ist unstetig in $y = 1$.

Die Annahme, dass der Definitionsbereich ein Intervall ist, scheint somit unentbehrlich. Hingegen gilt ein zu Satz 5.6.4 analoger Satz auf offenen Intervallen.

Satz 5.6.5. Sei $f:]a, b[\rightarrow \mathbb{R}$ stetig und streng monoton wachsend mit monotonen Limites

$$-\infty \leq c := \lim_{x \downarrow a} f(x) < \lim_{x \uparrow b} f(x) =: d \leq \infty.$$

Dann ist $f:]a, b[\rightarrow]c, d[$ bijektiv, und f^{-1} ist stetig.

Beweis. Wegen (5.6.1) ist $f:]a, b[\rightarrow]c, d[$ injektiv. Weiter gibt es zu jedem $y \in]c, d[$ Punkte $\bar{a}, \bar{b} \in]a, b[$ mit $\bar{a} < \bar{b}$ und $f(\bar{a}) < y < f(\bar{b})$. Gemäss Satz 5.6.1 existiert $x \in]\bar{a}, \bar{b}[$ mit $f(x) = y$; also ist f surjektiv und somit bijektiv.

Zum Nachweis der Stetigkeit von f^{-1} an einer beliebigen Stelle $y \in]c, d[$ argumentieren wir mit dem Folgenkriterium. Seien $\bar{a}, \bar{b} \in]a, b[$ mit $f(\bar{a}) < y < f(\bar{b})$ wie oben. Zu $y_k = f(x_k) \in]c, d[$ mit $y_k \rightarrow y$ ($k \rightarrow \infty$) wähle $k_0 \in \mathbb{N}$ mit

$$\forall k \geq k_0 : f(\bar{a}) \leq y_k \leq f(\bar{b}), \text{ also auch } \bar{a} \leq x_k \leq \bar{b}.$$

Sei \bar{f} die auf das abgeschlossene Intervall $[\bar{a}, \bar{b}]$ eingeschränkte Funktion f . Auf \bar{f} ist Satz 5.6.4 anwendbar, und wir erhalten

$$x_k = f^{-1}(y_k) = (\bar{f})^{-1}(y_k) \xrightarrow{(k \rightarrow \infty, k \geq k_0)} x := (\bar{f})^{-1}(y) = f^{-1}(y);$$

also ist f^{-1} stetig. □

Beispiel 5.6.3. i) Sei $n \in \mathbb{N}$. Die Potenzfunktion $\mathbb{R} \ni x \mapsto x^n \in \mathbb{R}$ ist stetig gemäss Beispiel 5.3.1.i), und sie ist streng monoton wachsend auf $\mathbb{R}_+ =]0, \infty[$ gemäss (3.2.1). Mit Satz 5.6.5 folgt, dass die n -te Wurzelfunktion

$$\mathbb{R}_+ \ni y \mapsto \sqrt[n]{y} \in \mathbb{R}_+$$

stetig ist.

ii) Betrachte die Funktion $Exp: \mathbb{R} \rightarrow \mathbb{R}$ aus Beispiel 3.7.2.i).

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

$$\forall x \in \mathbb{R} : Exp(x) = \left(Exp\left(\frac{x}{2}\right)\right)^2 \geq 0,$$

wegen

$$Exp(x) Exp(-x) = 1 \tag{5.6.2}$$

also

$$\forall x \in \mathbb{R} : Exp(x) > 0.$$

Weiter gilt für $|h| < 1$

$$|Exp(h) - 1| = \left| \sum_{k=1}^{\infty} \frac{h^k}{k!} \right| \leq \sum_{k=1}^{\infty} |h|^k = \frac{|h|}{1 - |h|} \rightarrow 0 \quad (h \rightarrow 0),$$

also für $x = x_0 + h \rightarrow x_0$ mit Korollar 3.8.1

$$\text{Exp}(x) - \text{Exp}(x_0) = \text{Exp}(x_0)(\text{Exp}(h) - 1) \rightarrow 0, \quad (5.6.3)$$

und die Funktion Exp ist stetig. Da

$$\text{Exp}(h) - 1 = \sum_{k=1}^{\infty} \frac{h^k}{k!} > 0 \text{ für } h > 0,$$

ergibt (5.6.3) zudem die gewünschte Monotonie

$$\text{Exp}(x_0) < \text{Exp}(x) \text{ für } x_0 < x = x_0 + h.$$

Schliesslich gilt offenbar $\text{Exp}(x) \rightarrow \infty$ ($x \rightarrow \infty$); mit (5.6.2) folgt dann auch $\text{Exp}(x) = 1/\text{Exp}(-x) \rightarrow 0$ ($x \rightarrow -\infty$). \square

Gemäss Satz 5.6.5 besitzt $\text{Exp}: \mathbb{R} \rightarrow]0, \infty[$ die stetige Umkehrfunktion

$$\text{Log} = (\text{Exp}|_{\mathbb{R}})^{-1} :]0, \infty[\rightarrow \mathbb{R}.$$

Wegen der Identität $\text{Exp}(x) = e^x$ für alle $x \in \mathbb{Q}$ gemäss Satz 3.9.2 stimmt Log überein mit dem *natürlichen Logarithmus* $\text{Log} = \log$. (Die vielfach gebräuchliche Notation \ln verwenden wir nicht.) Mit

$$\text{Exp}(\text{Log}(x) + \text{Log}(y)) = \text{Exp}(\text{Log}(x)) \cdot \text{Exp}(\text{Log}(y)) = xy$$

folgt zudem das *Additionstheorem*

$$\forall x, y > 0 : \text{Log}(xy) = \text{Log}(x) + \text{Log}(y). \quad (5.6.4)$$

Bemerkung 5.6.2. Die Eigenschaft (5.6.4) ermöglicht das vereinfachte Multiplizieren mittels Rechenschieber oder Logarithmentafel (Jost Bürgi (1552-1632)).

5.6.2 Topologische Anwendungen

Der Zwischenwertsatz hat umgekehrt auch topologische Konsequenzen. Insbesondere erhalten wir nun einen eleganten Beweis für Satz 4.4.1.

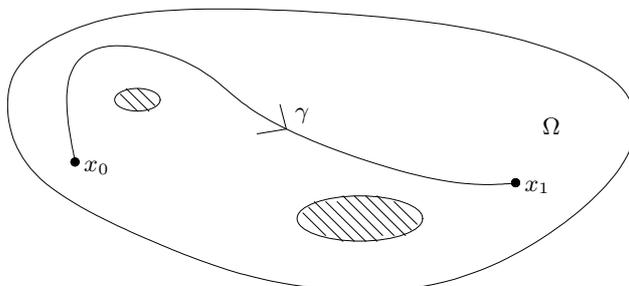
Satz 5.6.6. Sei $I = [0, 1] = \Omega_0 \cup \Omega_1$, wobei $\Omega_0, \Omega_1 \subset I$ relativ offen mit $\Omega_0 \neq \emptyset \neq \Omega_1$. Dann ist $\Omega_0 \cap \Omega_1 \neq \emptyset$.

Beweis (indirekt). Sei $I = \Omega_0 \cup \Omega_1$, wobei $\Omega_0 \neq \emptyset \neq \Omega_1$ relativ offen mit $\Omega_0 \cap \Omega_1 = \emptyset$. Definiere $f: [0, 1] \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & x \in \Omega_0 \\ 1, & x \in \Omega_1. \end{cases}$$

Dann ist f stetig gemäss Satz 5.2.2, denn alle möglichen Urbilder $\emptyset, \Omega_0, \Omega_1, I$ sind relativ offen. Für $a \in \Omega_0, b \in \Omega_1$ mit $f(a) = 0 < f(b) = 1$ liefert Satz 5.6.1 die Existenz eines Punktes $x \in [a, b]$ mit $f(x) = 1/2$, was der Definition von f widerspricht. \square

Definition 5.6.2. Eine Menge $\Omega \subset \mathbb{R}^d$ heisst (stetig) wegzusammenhängend, falls zu je zwei Punkten $x_0, x_1 \in \Omega$ ein stetiger “Weg” $\gamma: [0, 1] \rightarrow \Omega$ existiert mit $\gamma(0) = x_0, \gamma(1) = x_1$.



Beispiel 5.6.4. i) \mathbb{R}^d ist wegzusammenhängend.

ii) $B_R(0) \subset \mathbb{R}^d$ ist wegzusammenhängend, und allgemein jede konvexe Menge.

Satz 5.6.7. Sei $\Omega \subset \mathbb{R}^d$ stetig wegzusammenhängend. Dann ist Ω zusammenhängend.

Beweis (indirekt). Sei $\Omega = \Omega_0 \cup \Omega_1$, wobei $\Omega_0 \neq \emptyset \neq \Omega_1$ relativ offen in Ω mit $\Omega_0 \cap \Omega_1 = \emptyset$, und sei $\gamma: [0, 1] \rightarrow \Omega$ stetig mit $\gamma(0) = x_0, \gamma(1) = x_1$ für zwei Punkte $x_0 \in \Omega_0, x_1 \in \Omega_1$. Dann sind $E_i = \gamma^{-1}(\Omega_i) \subset [0, 1]$ relativ offen, disjunkt und nichtleer, $i = 0, 1$, und $[0, 1] = E_0 \cup E_1$ im Widerspruch zu Satz 5.6.6. \square

Falls $\Omega \subset \mathbb{R}^d$ offen, so gilt auch die Umkehrung von Satz 5.6.7.

Satz 5.6.8. Sei $\Omega \subset \mathbb{R}^d$ zusammenhängend und offen. Dann ist Ω wegzusammenhängend.

Beweis. Fixiere einen beliebigen Punkt $x_0 \in \Omega$. Setze

$$\Omega_0 = \{x_1 \in \Omega; \exists \gamma \in C^0([0, 1]; \Omega) : \gamma(0) = x_0, \gamma(1) = x_1\},$$

und definiere $\Omega_1 = \Omega \setminus \Omega_0$. Da Ω offen, gibt es zu $x_1 \in \Omega_0 \subset \Omega$ ein $r > 0$ mit $B_r(x_1) \subset \Omega$. Mit Beispiel 5.6.4.ii) erhalten wir für jedes $x_2 \in B_r(x_1)$ durch Aneinanderhängen eines Weges γ_1 von x_0 nach x_1 und eines Weges γ_2 von x_1 nach x_2 einen Weg γ von x_0 nach x_2 , mit

$$\gamma(t) = \begin{cases} \gamma_1(2t), & 0 \leq t < 1/2, \\ \gamma_2(2t - 1), & 1/2 \leq t \leq 1. \end{cases}$$

Es folgt $B_r(x_1) \subset \Omega_0$; die Menge Ω_0 ist also offen.

Analog ist Ω_1 offen: Sei $x_2 \in \Omega_1 = \Omega \setminus \Omega_0$ und dazu $r > 0$ mit $B_r(x_2) \subset \Omega$. Falls ein $x_1 \in B_r(x_2)$ existiert mit $x_1 \in \Omega_0$, so liefert das obige Argument $x_2 \in B_r(x_2) \subset \Omega_0$ im Widerspruch zur Annahme $x_2 \notin \Omega_0$.

Da $x_0 \in \Omega_0 \neq \emptyset$, und da Ω nach Annahme zusammenhängend ist, folgt $\Omega_1 = \emptyset$, wie gewünscht. \square

5.7 Folgen stetiger Funktionen

5.7.1 Punktweise und gleichmässige Konvergenz

Sei $\Omega \subset \mathbb{R}^d$, und seien $f, f_k: \Omega \rightarrow \mathbb{R}^n$, $k \in \mathbb{N}$.

Definition 5.7.1. Die Folge $(f_k)_{k \in \mathbb{N}}$ **konvergiert punktweise** gegen f , falls gilt

$$\forall x \in \Omega : f_k(x) \rightarrow f(x) \quad (k \rightarrow \infty).$$

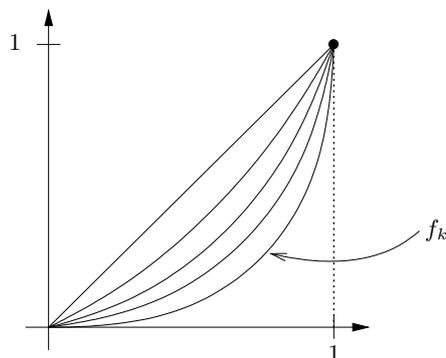
Beispiel 5.7.1. Sei $f_k(x) = x^k$, $k \in \mathbb{N}$, $0 \leq x \leq 1$. Offenbar gilt

$$f_k(x) \xrightarrow{(k \rightarrow \infty)} \begin{cases} 0, & x < 1, \\ 1, & x = 1; \end{cases}$$

die Folge $(f_k)_{k \in \mathbb{N}}$ ist also punktweise konvergent. Beachte, dass $f_k: [0, 1] \rightarrow \mathbb{R}$ für jedes $k \in \mathbb{N}$ stetig ist, die Limesfunktion $f: [0, 1] \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & x < 1 \\ 1, & x = 1 \end{cases}$$

hingegen nicht.



Definition 5.7.2. Die Folge $(f_k)_{k \in \mathbb{N}}$ **konvergiert gleichmässig** gegen f , falls $f_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$), falls

$$\sup_{x \in \Omega} |f_k(x) - f(x)| \rightarrow 0 \quad (k \rightarrow \infty).$$

Bemerkung 5.7.1. i) Gleichmässige Konvergenz impliziert punktweise Konvergenz.

ii) Falls $\overline{\Omega} \subset \mathbb{R}^d$ kompakt ist und falls f, f_k stetig (oder stetig auf $\overline{\Omega}$ ergänzbar), dann ist gleichmässige Konvergenz von (f_k) gegen f äquivalent zur Bedingung

$$\|f_k - f\|_{C^0(\overline{\Omega}; \mathbb{R}^n)} \rightarrow 0 \quad (k \rightarrow \infty).$$

Daher sagt man anstelle von gleichmässiger Konvergenz auch *normale Konvergenz* oder *Normkonvergenz*

Beispiel 5.7.2. Seien $(a_k)_{k \in \mathbb{N}_0} \subset \mathbb{C}$, $p(z) = \sum_{k=0}^{\infty} a_k z^k$ die Potenzreihe in $z \in \mathbb{C}$ mit Konvergenzradius

$$0 \leq \rho = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}} \leq \infty.$$

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

$$\begin{aligned} |p(z) - p_n(z)| &\leq \sum_{k=n}^{\infty} |a_k| |z|^k \leq \sum_{k=n}^{\infty} |a_k| r^k \\ &= \sum_{k=n}^{\infty} |a_k| \left(\frac{r}{s}\right)^k s^k \leq \left(\frac{r}{s}\right)^n \underbrace{\sum_{k=0}^{\infty} |a_k| s^k}_{< \infty} \rightarrow 0 \quad (n \rightarrow \infty), \end{aligned}$$

gleichmässig in z . □

Welche Konsequenzen hat die gleichmässige Konvergenz? Ist p insbesondere stetig in $B_\rho(0)$?

Satz 5.7.1. Seien $f_k: \Omega \subset \mathbb{R}^d \rightarrow \mathbb{R}^n$ stetig, $k \in \mathbb{N}$, und sei $f: \Omega \rightarrow \mathbb{R}^n$. Weiter gelte $f_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$). Dann ist f stetig.

Beweis. Fixiere $x_0 \in \Omega$. Zu $\epsilon > 0$ wähle $k_0 \in \mathbb{N}$ mit

$$\forall k \geq k_0 : \sup_{x \in \Omega} |f_k(x) - f(x)| < \epsilon.$$

Fixiere $k = k_0$. Da f_{k_0} stetig, gibt es $\delta > 0$, so dass gilt

$$\forall x \in \Omega : |x - x_0| < \delta \Rightarrow |f_{k_0}(x) - f_{k_0}(x_0)| < \epsilon.$$

Für alle $x \in \Omega$ mit $|x - x_0| < \delta$ folgt

$$|f(x) - f(x_0)| \leq |f(x) - f_{k_0}(x)| + |f_{k_0}(x) - f_{k_0}(x_0)| + |f_{k_0}(x_0) - f(x_0)| < 3\epsilon.$$

Nach Satz 5.2.1 ist f stetig in x_0 ; da x_0 beliebig, folgt die Behauptung. □

Korollar 5.7.1. Potenzreihen sind stetig im Innern ihres Konvergenzkreises.

Beweis. Unmittelbar aus Satz 5.7.1 mit Beispiel 5.7.2. □

5.7.2 Monotonie und Konvergenz

Mit Hilfe der Ordnungsstruktur von \mathbb{R} kann man die Aussage von Satz 5.7.1 in gewissen Fällen auch umkehren.

Satz 5.7.2. (Dini) Sei $K \subset \mathbb{R}^d$ kompakt, $f_k : K \subset \mathbb{R}^d \rightarrow \mathbb{R}$ stetig, $k \in \mathbb{N}$. Weiter sei $f : K \rightarrow \mathbb{R}$, und es gelte

$$\forall x \in K : f_k(x) \leq f_{k+1}(x) \rightarrow f(x) \quad (k \rightarrow \infty).$$

Falls dann f stetig ist, so konvergiert $(f_k)_{k \in \mathbb{N}}$ sogar gleichmässig gegen f .

Bemerkung 5.7.2. i) Das Beispiel 5.7.1 mit $f_k(x) = x^k$, $k \in \mathbb{N}$, zeigt, dass die Stetigkeit von f und die Kompaktheit von K notwendig sind.

ii) Auch die Monotonie der Konvergenz ist notwendig, wie das Beispiel der Funktionen $f_k : [0, 1] \rightarrow \mathbb{R}$ zeigt mit

$$f_k(x) = \max\{\min\{kx, 2 - kx\}, 0\}, \quad k \in \mathbb{N}.$$

Beweis von Satz 5.7.2. Sei $\epsilon > 0$. Zu $x_0 \in K$ wähle $k_0 = k_0(\epsilon, x_0) \in \mathbb{N}$ mit

$$\forall k \geq k_0 : f(x_0) - \epsilon < f_k(x_0) \leq f(x_0).$$

Fixiere $k = k_0$ und wähle $\delta = \delta(\epsilon, x_0, k_0) > 0$ mit

$$\forall x \in K \cap B_\delta(x_0) : |f(x) - f(x_0)| + |f_{k_0}(x) - f_{k_0}(x_0)| < \epsilon.$$

Für $x \in K \cap B_\delta(x_0)$ und $k \geq k_0 = k_0(\epsilon, x_0)$ folgt mit der punktweisen monotonen Konvergenz

$$f(x) \geq f_k(x) \geq f_{k_0}(x) \geq f_{k_0}(x_0) - \epsilon \geq f(x_0) - 2\epsilon \geq f(x) - 3\epsilon; \quad (5.7.1)$$

das heisst,

$$\forall k \geq k_0(\epsilon, x_0) : \sup_{x \in K \cap B_\delta(x_0)} |f_k(x) - f(x)| \leq 3\epsilon.$$

Die Familie $(B_{\delta(x_0)}(x_0))_{x_0 \in K}$, wobei wir für jedes $x_0 \in K$ zur Abkürzung $\delta(x_0) = \delta(\epsilon, x_0, k_0(\epsilon, x_0))$ schreiben, ist eine offene Überdeckung von K . Da K nach Annahme kompakt ist, gibt es $x_1, \dots, x_L \in K$ mit

$$K \subset \bigcup_{\ell=1}^L B_{\delta_\ell}(x_\ell),$$

wo $\delta_\ell = \delta(x_\ell) > 0$, $1 \leq \ell \leq L$. Dann gilt für $k \geq \max_{1 \leq \ell \leq L} k_0(\epsilon, x_\ell)$ mit (5.7.1) gleichmässig für alle $x \in K$ die Abschätzung

$$|f_k(x) - f(x)| \leq \max_{1 \leq \ell \leq L} \sup_{x \in K \cap B_{\delta_\ell}(x_\ell)} |f_k(x) - f(x)| \leq 3\epsilon;$$

also $f_k \xrightarrow{\text{glm.}} f$ ($k \rightarrow \infty$). □

Eine weitere Variante einer Monotoniebedingung erscheint im nächsten Satz.

Satz 5.7.3. (Pólya-Szegő) Sei $I = [a, b]$, $f_k: I \rightarrow \mathbb{R}$ monoton wachsend, $k \in \mathbb{N}$. Weiter sei $(f_k)_{k \in \mathbb{N}}$ punktweise konvergent gegen $f: I \rightarrow \mathbb{R}$. Dann ist f ebenfalls monoton wachsend. Falls f zudem stetig ist, so konvergiert $(f_k)_{k \in \mathbb{N}}$ gleichmässig gegen f .

Bemerkung 5.7.3. Die f_k müssen nicht stetig sein.

Beweis von Satz 5.7.3. Monotonie von f folgt mit Satz 3.2.2.iv). Sei f stetig, nach Satz 5.5.2 also gleichmässig stetig. Zu $\epsilon > 0$ wähle $\delta = \delta(\epsilon) > 0$ mit

$$\forall x, y \in I: |x - y| < 2\delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Fixiere $x_0 \in I$, und seien $x_1, x_2 \in I$ mit

$$x_0 - \delta < x_1 \leq x_0 \leq x_2 < x_0 + \delta.$$

Wähle $k_0 = k_0(\epsilon, x_1, x_2) \in \mathbb{N}$ mit

$$\forall k \geq k_0: |f_k(x_i) - f(x_i)| < \epsilon, i = 1, 2.$$

Für $x \in [x_1, x_2] \subset I$ und beliebiges $k \geq k_0$ folgt mit Monotonie von f_k und f die Abschätzung

$$f_k(x) - f(x) \leq f_k(x_2) - f(x_1) = f_k(x_2) - f(x_2) + f(x_2) - f(x_1) < 2\epsilon;$$

analog

$$f_k(x) - f(x) \geq f_k(x_1) - f(x_2) = f_k(x_1) - f(x_1) + f(x_1) - f(x_2) > -2\epsilon.$$

Das heisst,

$$\forall k \geq k_0: \sup_{x_1 \leq x \leq x_2} |f_k(x) - f(x)| \leq 2\epsilon.$$

Endlich viele Intervalle $B_\delta(x_\ell)$, $1 \leq \ell \leq L$, und damit endlich viele Intervalle $[x_1^{(\ell)}, x_2^{(\ell)}]$ mit geeigneten $x_1^{(\ell)}, x_2^{(\ell)} \in B_\delta(x_\ell)$, $1 \leq \ell \leq L$, überdecken I . Es folgt

$$\forall k \geq \max_{1 \leq \ell \leq L} k_\ell: \sup_I |f_k(x) - f(x)| \leq 2\epsilon,$$

wobei $k_\ell = k(\epsilon, x_1^{(\ell)}, x_2^{(\ell)})$, $1 \leq \ell \leq L$, wie gewünscht. \square

5.7.3 Der Raum $C^0(\overline{\Omega})$

Sei $\Omega \subset \mathbb{R}$ beschränkt, und sei

$$C^0(\overline{\Omega}) = C^0(\overline{\Omega}; \mathbb{R}^n) = \{f: \Omega \rightarrow \mathbb{R}^n; f \text{ ist stetig auf } \overline{\Omega} \text{ ergänzbar}\}$$

mit der Norm

$$\|f\|_{C^0(\overline{\Omega}; \mathbb{R}^n)} = \|f\|_{C^0} = \sup_{x \in \Omega} |f(x)|, f \in C^0(\overline{\Omega}),$$

gemäss Abschnitt 5.3.2. 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 5.7.4. Sei $(f_k)_{k \in \mathbb{N}}$ Folge in $C^0(\overline{\Omega}; \mathbb{R}^n)$ mit

$$\|f_k - f_l\|_{C^0} \rightarrow 0 \quad (k, l \rightarrow \infty).$$

Dann gibt es $f \in C^0(\overline{\Omega}; \mathbb{R}^n)$ mit $f_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$).

Beweis. Für jedes $x_0 \in \overline{\Omega}$ ist wegen

$$|f_k(x_0) - f_l(x_0)| \leq \sup_{x \in \overline{\Omega}} |f_k(x) - f_l(x)| = \|f_k - f_l\|_{C^0} \rightarrow 0 \quad (k, l \rightarrow \infty)$$

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

$$f(x_0) := \lim_{k \rightarrow \infty} f_k(x_0).$$

Weiter gilt

$$\begin{aligned} |f_k(x) - f(x)| &= \lim_{l \rightarrow \infty} |f_k(x) - f_l(x)| \\ &\leq \limsup_{l \rightarrow \infty} \|f_k - f_l\|_{C^0} \rightarrow 0 \quad (k \rightarrow \infty), \end{aligned}$$

gleichmässig in x ; das heisst, $f_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$), und mit Satz 5.7.1 ist f zudem stetig auf $\overline{\Omega}$. \square

Bemerkung 5.7.4. i) Gemäss Satz 5.7.4 ist der Raum $C^0(\overline{\Omega}; \mathbb{R}^n)$ metrisch vollständig bezüglich der Supremumsnorm; er ist ein *Banachraum*.

ii) Im Unterschied zu den uns bisher bekannten Räumen ist der Raum $C^0(\overline{\Omega}; \mathbb{R}^n)$ unendlich-dimensional. Betrachte zum Beispiel die Funktionen $f_k \in C^0([0, 1])$ mit

$$f_k(x) = \max\{\min\{2^k x, 2 - 2^k x\}, 0\}, \quad k \in \mathbb{N}.$$

Behauptung. Die Funktionen $(f_k)_{k \in \mathbb{N}}$ sind linear unabhängig.

Beweis (indirekt). Seien $a_1, \dots, a_n \in \mathbb{R}^n \setminus \{0\}$ mit

$$a_1 f_1 + a_2 f_2 + \dots + a_n f_n = 0$$

für ein $n \geq 2$. Sei $k_0 = \min\{k; a_k \neq 0\}$, und fixiere $x = 2^{-k_0}$. Für $k > k_0$ gilt dann $f_k(x) = 0$, während $f_{k_0}(x) = 1$; also

$$a_{k_0} = \sum_{k=1}^n a_k f_k(x) = 0$$

im Widerspruch zur Wahl von k_0 . \square

iii) Abgeschlossene und beschränkte Teilmengen von $C^0(\overline{\Omega})$ sind im Allgemeinen nicht folgenkompakt (und auch nicht überdeckungskompakt), wie das Beispiel 5.7.1 zeigt. Die Folge der Funktionen $f_k \in C^0([0, 1])$ mit $f_k(x) = x^k$, $k \in \mathbb{N}$, konvergiert punktweise gegen die unstetige Grenzfunktion

$$f(x) = \begin{cases} 0, & 0 \leq x < 1, \\ 1, & x = 1, \end{cases}$$

kann also keine C^0 -konvergente Teilfolge haben. Die Menge

$$A = \{f_k; k \in \mathbb{N}\} \subset C^0([0, 1])$$

ist somit (folgen-)abgeschlossen und beschränkt aber nicht folgenkompakt.

Ebenso wie \mathbb{R} die rationalen Zahlen \mathbb{Q} als abzählbare dichte Teilmenge enthält, so gibt es überraschenderweise auch in $C^0(\overline{\Omega}; \mathbb{R}^n)$ stets abzählbare dichte Teilmengen; das heisst, $C^0(\overline{\Omega})$ ist *separabel*.

Falls insbesondere Ω ein Intervall in \mathbb{R} ist, so gilt beispielsweise der folgende Satz.

Satz 5.7.5. (Weierstrass) Sei $I = [0, 1]$, und sei $f \in C^0(I)$ gegeben. Dann gibt es Polynome p_k mit rationalen Koeffizienten, $k \in \mathbb{N}$, so dass $p_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$).

Beweis. Sei $\epsilon > 0$ vorgegeben. Da $f \in C^0(I)$ nach Satz 5.5.2 auf I gleichmässig stetig ist, gibt es ein $\delta = \delta(\epsilon) > 0$ mit

$$\forall x, y \in I : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon. \quad (5.7.2)$$

Wähle $0 = x_0 < x_1 < \dots < x_N = 1$ mit

$$\delta/2 < |x_n - x_{n-1}| < \delta, \quad 1 \leq n \leq N = N(\delta) \leq 2/\delta.$$

Für $H : \mathbb{R} \rightarrow [0, 1]$ mit

$$H(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0, \end{cases}$$

betrachte die stückweise konstante Funktion

$$g(x) = f(x_0) + \sum_{n=1}^N (f(x_n) - f(x_{n-1}))H(x - x_n)$$

mit

$$g(x) = f(x_n) \quad \text{für } x_n \leq x < x_{n+1}, \quad 0 \leq n \leq N.$$

Mit (5.7.2) folgt

$$\sup_{x \in I} |f(x) - g(x)| = \max_{0 \leq n \leq N} \sup_{x_n \leq x < x_{n+1}} |f(x) - f(x_n)| < \epsilon. \quad (5.7.3)$$

Behauptung 1. Es gibt ein Polynom $q : \mathbb{R} \rightarrow \mathbb{R}$ mit $0 \leq q(x) \leq 1$ für $-1 \leq x \leq 1$, so dass

$$\sup_{\delta \leq |x| \leq 1} |H(x) - q(x)| \leq \delta.$$

Beweis. Betrachte zunächst die Folge von Polynomen

$$q_k(x) = (1 - x^k)^{2^k}, \quad 0 \leq x \leq 1.$$

Mit Bernoulli folgt für $0 \leq x \leq \alpha := \frac{1-\delta}{2} < \frac{1}{2}$ und genügend grosses $k_0 = k_0(\delta)$:

$$\forall k \geq k_0 : 1 \geq q_k(x) \geq q_k(\alpha) = (1 - \alpha^k)^{2^k} \geq 1 - 2^k \alpha^k = 1 - (1 - \delta)^k > 1 - \delta.$$

Analog gilt für $1 \geq x \geq \beta := \frac{1+\delta}{2} > \frac{1}{2}$ und genügend grosses $k_0 = k_0(\delta)$:

$$\begin{aligned} \forall k \geq k_0 : \frac{1}{q_k(x)} &\geq \frac{1}{q_k(\beta)} = \left(\frac{1}{1-\beta^k} \right)^{2^k} = \left(1 + \frac{\beta^k}{1-\beta^k} \right)^{2^k} \\ &\geq 1 + \frac{2^k \beta^k}{1-\beta^k} \geq 1 + (1+\delta)^k \geq 1/\delta; \end{aligned}$$

also

$$\forall k \geq k_0, 1 \geq x \geq \beta : 0 \leq q_k(x) \leq \delta.$$

Das Polynom q mit

$$q(x) = q_{k_0} \left(\frac{1-x}{2} \right)$$

erfüllt dann unsere Behauptung. \square

Sei q gemäss Behauptung 1 gewählt. Setze

$$p(x) = f(x_0) + \sum_{n=1}^N (f(x_n) - f(x_{n-1})) q(x - x_n).$$

Da $\delta/2 < x_n - x_{n-1} < \delta$, $1 \leq n \leq N$, gibt es zu $x \in [0, 1]$ ein $n_0 \in \{0, \dots, N\}$ mit

$$x_{n_0-2} < x - \delta < x_{n_0} \leq x < x_{n_0+1} < x + \delta < x_{n_0+3};$$

insbesondere folgt $x - x_n > \delta$ für $n \leq n_0 - 2$ sowie $x_n - x > \delta$ für $n \geq n_0 + 3$.

Mit (5.7.3) erhalten wir für jedes $x \in I$ die Abschätzung

$$\begin{aligned} |f(x) - p(x)| &\leq |f(x) - g(x)| + |g(x) - p(x)| \\ &\leq \epsilon + \sum_{n=1}^N |f(x_n) - f(x_{n-1})| |(H - q)(x - x_n)|. \end{aligned}$$

Wähle n_0 mit $x_{n_0} \leq x < x_{n_0+1}$ und schätze ab mit (5.7.2) und Behauptung 1:

$$\begin{aligned} &\sum_{n=1}^N |f(x_n) - f(x_{n-1})| |(H - q)(x - x_n)| \\ &\leq N\epsilon \sup_{\delta \leq |x| \leq 1} |H(x) - q(x)| + \epsilon \sum_{n=n_0-1}^{n_0+2} |(H - q)(x - x_n)| \leq N\epsilon\delta + 4\epsilon \leq 6\epsilon, \end{aligned}$$

wobei wir benutzen, dass gemäss Behauptung 1 stets gilt $|H(x) - q(x)| \leq 1$.

Es folgt die Abschätzung $|f(x) - p(x)| \leq 7\epsilon$, gleichmässig in x . Schliesslich kann man offenbar p bis auf den Fehler ϵ auf I durch ein Polynom mit rationalen Koeffizienten annähern. \square

Bemerkung 5.7.5. Die Menge der Polynome mit rationalen Koeffizienten ist analog zu Satz 2.4.1 und Bemerkung 2.4.1 abzählbar.

Kapitel 6

Differentialrechnung auf \mathbb{R}

6.1 Differential und Differentiationsregeln

Sei $\Omega \subset \mathbb{R}$ offen, $f : \Omega \rightarrow \mathbb{R}$, $x_0 \in \Omega$.

Definition 6.1.1. i) f heisst **differenzierbar** an der Stelle x_0 , falls der Grenzwert

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - f(x_0)}{x - x_0} =: f'(x_0) =: \frac{df}{dx}(x_0)$$

existiert. In diesem Fall heisst $f'(x_0)$ die **Ableitung** oder das **Differential** von f an der Stelle x_0 .

ii) Analog heisst $f = (f_1, \dots, f_n) : \Omega \rightarrow \mathbb{R}^n$ an der Stelle x_0 differenzierbar, falls jede der Komponentenfunktionen f_i an der Stelle x_0 differenzierbar ist, und $f'(x_0) = (f'_1(x_0), \dots, f'_n(x_0))$.

Bemerkung 6.1.1. 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 6.1.2. $f : \Omega \rightarrow \mathbb{R}^n$ heisst **auf Ω differenzierbar**, falls f an jeder Stelle $x_0 \in \Omega$ differenzierbar ist.

Beispiel 6.1.1. i) Sei $f(x) = mx + b$, $x \in \mathbb{R}$, mit Konstanten m , $b \in \mathbb{R}$. Es gilt

$$\forall x \neq x_0 : \frac{f(x) - f(x_0)}{x - x_0} = m;$$

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 $f(x) = \text{Exp}(x)$, $x \in \mathbb{R}$. Mit Korollar 3.8.1 folgt für $x_0 \neq x = x_0 + h \in \mathbb{R}$

$$\begin{aligned} \frac{\text{Exp}(x_0 + h) - \text{Exp}(x_0)}{h} &= \frac{\text{Exp}(x_0)(\text{Exp}(h) - 1)}{h} \\ &= \text{Exp}(x_0) \sum_{k=1}^{\infty} \frac{h^{k-1}}{k!} \rightarrow \text{Exp}(x_0) \quad (h \rightarrow 0); \end{aligned}$$

das heisst, die Funktion $\text{Exp}: \mathbb{R} \rightarrow \mathbb{R}$ ist an jeder Stelle $x_0 \in \mathbb{R}$ differenzierbar mit $\text{Exp}'(x_0) = \text{Exp}(x_0)$, oder

$$\text{Exp}' = \text{Exp}.$$

iv) Ebenso wie vektorwertige Funktionen werden Funktionen $f: \mathbb{R} \rightarrow \mathbb{C} \cong \mathbb{R}^2$ komponentenweise differenziert. Betrachte die in Abschnitt 3.9. definierte Funktion

$$f(t) = \text{Exp}(it) = \text{Cos}(t) + i \text{Sin}(t), \quad t \in \mathbb{R}.$$

Analog zu iii) gilt für $t_0 \neq t = t_0 + h \in \mathbb{R}$

$$\frac{f(t) - f(t_0)}{t - t_0} = \text{Exp}(it_0) \frac{\text{Exp}(ih) - 1}{h}$$

mit

$$\frac{\text{Exp}(ih) - 1}{h} \rightarrow i \quad (h \rightarrow 0).$$

Es folgt, dass f an jeder Stelle $t_0 \in \mathbb{R}$ differenzierbar ist mit

$$\begin{aligned} f'(t_0) &= \text{Cos}'(t_0) + i \text{Sin}'(t_0) \\ &= i \text{Exp}(it_0) = -\text{Sin}(t_0) + i \text{Cos}(t_0). \end{aligned}$$

Das heisst,

$$\text{Cos}' = -\text{Sin}, \quad \text{Sin}' = \text{Cos}.$$

“Differenzierbarkeit” ist mehr als “Stetigkeit”; genauer gilt:

Satz 6.1.1. *Ist $f: \Omega \rightarrow \mathbb{R}$ 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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$) gilt gemäss Satz 3.2.2

$$f(x_k) - f(x_0) = \underbrace{\frac{f(x_k) - f(x_0)}{x_k - x_0}}_{\rightarrow f'(x_0)} \cdot \underbrace{(x_k - x_0)}_{\rightarrow 0} \rightarrow 0 \quad (k \rightarrow \infty, x_k \neq x_0),$$

bzw.

$$f(x_k) - f(x_0) = 0, \quad \text{falls } x_k = x_0.$$

Also folgt in jedem Fall $f(x_k) \rightarrow f(x_0)$ ($k \rightarrow \infty$), wie gewünscht. \square

Bemerkung 6.1.2. i) Das Beispiel 6.1.1.ii) der Funktion $f(x) = |x|$, $x \in \mathbb{R}$, zeigt jedoch, dass stetige Funktionen nicht unbedingt überall differenzierbar sein müssen; vergleiche auch Beispiel 6.4.3.iv).

ii) Es gibt sogar stetige Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$, die an **keiner** Stelle $x_0 \in \mathbb{R}$ differenzierbar sind (“Koch-Kurven”); vergleiche dazu die Übungen.

6.1.1 Differentiationsregeln

Satz 6.1.2. Seien $f, g: \Omega \rightarrow \mathbb{R}$ 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 3.2.2

$$\begin{aligned} \frac{(f + g)(x) - (f + g)(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \\ &\stackrel{(x \rightarrow x_0)}{\rightarrow} f'(x_0) + g'(x_0); \end{aligned}$$

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 6.1.1 und

$$\begin{aligned} \frac{(fg)(x) - (fg)(x_0)}{x - x_0} &= \frac{(f(x) - f(x_0))g(x) + f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0}g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0} \\ &\stackrel{(x \rightarrow x_0)}{\rightarrow} f'(x_0)g(x_0) + f(x_0)g'(x_0) \end{aligned}$$

die gewünschte Aussage für $f \cdot g$.

iii) Wegen ii) genügt es, den Fall $f \equiv 1$ zu betrachten. Mit Satz 6.1.1 folgt aus $g(x_0) \neq 0$, dass $g(x) \neq 0$ für alle x in einer Umgebung von x_0 , und $g(x) \rightarrow g(x_0)$ ($x \rightarrow x_0$, $x \in \Omega$). Mit Satz 3.2.2 erhalten wir

$$\frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} = \frac{g(x_0) - g(x)}{x - x_0} \cdot \frac{1}{g(x)g(x_0)} \stackrel{(x \rightarrow x_0, x \neq x_0)}{\rightarrow} -\frac{g'(x_0)}{g^2(x_0)}.$$

□

Beispiel 6.1.2. 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 6.1.1.i).

$n \rightarrow n + 1$: Setze $f(x) = x^n$, $g(x) = x$. Nach Induktionsvoraussetzung sind f und g differenzierbar mit

$$f'(x) = nx^{n-1}, \quad g'(x) = 1.$$

Mit Satz 6.1.2.ii) folgt

$$\frac{dx^{n+1}}{dx}(x) = (fg)'(x) = f'(x)g(x) + f(x)g'(x) = (n+1)x^n.$$

□

ii) Polynome $p(x) = a_n x^n + \dots + a_1 x + a_0$ sind auf \mathbb{R} differenzierbar mit $p'(x) = na_n x^{n-1} + \dots + a_1$.

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 6.1.3. (Kettenregel) Sei $f: \Omega \rightarrow \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar, und sei $g: \mathbb{R} \rightarrow \mathbb{R}$ an der Stelle $y_0 = f(x_0)$ differenzierbar. Dann ist die Funktion $g \circ f: \Omega \rightarrow \mathbb{R}$ an der Stelle x_0 differenzierbar, und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beispiel 6.1.3. Für affine Funktionen

$$f(x) = mx + c, \quad g(y) = ly + d$$

gilt

$$(g \circ f)(x) = l(mx + c) + d = (lm)x + (lc + d),$$

und daher

$$(g \circ f)'(x_0) = lm = g'(f(x_0)) \cdot f'(x_0).$$

Beweis von Satz 6.1.3. 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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$), und sei $f(x_k) \neq f(x_0)$, $k \in \mathbb{N}$. Da nach Satz 6.1.1 mit $x_k \rightarrow x_0$ auch $f(x_k) \rightarrow f(x_0)$ ($k \rightarrow \infty$), erhalten wir

$$\lim_{k \rightarrow \infty} \frac{(g \circ f)(x_k) - (g \circ f)(x_0)}{x_k - x_0} = g'(f(x_0))f'(x_0). \quad (6.1.1)$$

Falls für eine Folge $x_k \rightarrow x_0$ ($k \rightarrow \infty$) gilt

$$x_k \neq x_0, f(x_k) = f(x_0), k \in \mathbb{N},$$

so erhalten wir

$$f'(x_0) = \lim_{k \rightarrow \infty} \frac{f(x_k) - f(x_0)}{x_k - x_0} = 0,$$

und

$$\lim_{k \rightarrow \infty} \frac{g(f(x_k)) - g(f(x_0))}{x_k - x_0} = 0 = g'(f(x_0))f'(x_0).$$

Zusammen mit (6.1.1) liefert dies die gewünschte Konvergenz nun auch für *jede* Folge $(x_k) \subset \Omega$ mit $x_k \rightarrow x_0$ ($k \rightarrow \infty$). \square

Bemerkung 6.1.3. 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.

Beispiel 6.1.4. 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 6.1.2.i) und Satz 6.1.3 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 6.1.1.i) und iii) folgt

$$\left. \frac{d}{dt}(e^{\lambda t}) \right|_{t=t_0} = \underbrace{e^{\lambda t_0}}_{=g'(f(t_0))} \cdot \underbrace{\lambda}_{=f'(t_0)}.$$

iii) Analog gilt mit $g = f = \text{Exp}$ gemäss Beispiel 6.1.1.iii)

$$\left. \frac{d}{dx}(e^{e^x}) \right|_{x=x_0} = \underbrace{e^{e^{x_0}}}_{=g'(f(x_0))=e^{f(x_0)}} \cdot \underbrace{e^{x_0}}_{=f'(x_0)}.$$

6.2 Der Mittelwertsatz und Folgerungen

Im folgenden betrachten wir stets differenzierbare Funktionen auf einem Intervall $\Omega =]a, b[\subset \mathbb{R}$.

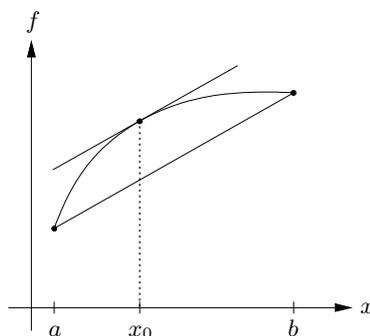
Satz 6.2.1. Seien $-\infty < a < b < \infty$. Sei $f: [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar in $]a, b[$. Dann existiert $x_0 \in]a, b[$ mit

$$f(b) = f(a) + f'(x_0)(b - a);$$

das heisst,

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}$$

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 Korollar 5.3.1 und Beispiel 5.3.2.i) gibt es $\underline{x}, \bar{x} \in [a, b]$ mit

$$f(\underline{x}) = \min_{a \leq x \leq b} f(x) \leq 0 \leq \max_{a \leq x \leq b} f(x) = f(\bar{x}).$$

Falls $f(\underline{x}) = 0 = f(\bar{x})$, so ist $f \equiv 0$; also gilt $f'(x) = 0$ für alle $x \in]a, b[$.

Andernfalls sei oBdA $f(\bar{x}) > 0$. (Sonst betrachte die Funktion $\tilde{f} = -f$.) Dann gilt offenbar $a < \bar{x} < b$, und es folgt

$$0 \geq \lim_{x \downarrow \bar{x}} \frac{f(x) - f(\bar{x})}{x - \bar{x}} = f'(\bar{x}) = \lim_{x \uparrow \bar{x}} \frac{f(x) - f(\bar{x})}{x - \bar{x}} \geq 0,$$

also $f'(\bar{x}) = 0$.

ii) Für allgemeines f betrachte die Funktion $g: [a, b] \rightarrow \mathbb{R}$ mit

$$g(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right).$$

Offenbar ist g stetig auf $[a, b]$ und in $]a, b[$ differenzierbar mit

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}, \quad x \in]a, b[.$$

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 6.2.1. Sei f wie in Satz 6.2.1.

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 6.2.1. i) Für eine differenzierbare Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ und ein $\lambda \in \mathbb{R}$ gelte die Beziehung $f' = \lambda f$, also

$$\forall t \in \mathbb{R} : f'(t) = \lambda f(t).$$

Dann gilt

$$\forall t \in \mathbb{R} : f(t) = f(0)e^{\lambda t}.$$

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 6.1.2 und Beispiel 6.1.4.ii) folgt

$$g'(t) = \frac{d}{dt}(e^{-\lambda t}) \cdot f(t) + e^{-\lambda t} f'(t) = e^{-\lambda t}(-\lambda f(t) + f'(t)) = 0$$

für alle $t \in \mathbb{R}$; also

$$g(t) \equiv g(0) = f(0)$$

gemäss Korollar 6.2.1.i), und

$$\forall t \in \mathbb{R} : f(t) = e^{\lambda t} g(t) = f(0)e^{\lambda t}.$$

□

ii) Die Funktion $f: x \mapsto \frac{2x}{1+x^2}$ erfüllt gemäss Satz 6.1.2

$$f'(x) = \frac{2(1+x^2) - 4x^2}{(1+x^2)^2} = \frac{2(1-x^2)}{(1+x^2)^2} < 0$$

für $x > 1$; also ist $f:]1, \infty[\rightarrow]0, 1[$ streng monoton fallend.

Korollar 6.2.2. (Bernoulli-de l'Hospital) Seien $f, g: [a, b] \rightarrow \mathbb{R}$ stetig und differenzierbar in $]a, b[$ mit $g'(x) \neq 0$ für jedes $x \in]a, b[$. Weiter gelte $f(a) = 0 = g(a)$, 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 6.2.1 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 $h: [a, s] \rightarrow \mathbb{R}$ ist stetig und differenzierbar in $]a, s[$ mit $h(a) = 0 = h(s)$. Nach Satz 6.2.1 gibt es $x = x(s) \in]a, s[$ mit

$$0 = h'(x) = \frac{f(s)}{g(s)}g'(x) - f'(x);$$

das heisst,

$$\frac{f(s)}{g(s)} = \frac{f'(x)}{g'(x)}.$$

Mit $s \rightarrow a$ folgt dann $x(s) \rightarrow a$, und

$$\frac{f(s)}{g(s)} \rightarrow \lim_{x \downarrow a} \frac{f'(x)}{g'(x)} = A.$$

□

Beispiel 6.2.2. i) Es gilt

$$\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = \lim_{x \rightarrow 1} \frac{2x}{1} = 2;$$

vergleiche Beispiel 5.1.1.

ii) Mit Beispiel 6.1.1.iv) erhalten wir

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

iii) Man kann die Bernoullische Regel auch mehrmals anwenden. Mit ii) folgt so

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

Es gibt viele Varianten der Bernoulli-de l'Hospitalschen Regel. Oft kann man jedoch auch ohne Gebrauch dieser Regeln durch geschicktes Umformen zum Ziel gelangen.

Beispiel 6.2.3. i) Beachte, dass mit

$$\frac{\text{Exp}(x)}{x^k} \geq \frac{x^{k+1}}{(k+1)!} \cdot \frac{1}{x^k} = \frac{x}{(k+1)!} \rightarrow \infty \quad (0 < x \rightarrow \infty)$$

auch gilt

$$\lim_{x \rightarrow \infty} (x^k e^{-x}) = \lim_{x \rightarrow \infty} \frac{1}{\left(\frac{\text{Exp}(x)}{x^k}\right)} = 0.$$

ii) Für beliebiges $\alpha > 0$ und $x > 0$ definieren wir

$$x^\alpha := \text{Exp}(\alpha \text{Log}(x)).$$

Mit Korollar 3.8.1 erhalten wir im Falle $\alpha = 1/n$, $n \in \mathbb{N}$, für jedes $x > 0$ die Gleichung

$$\left(\text{Exp}\left(\frac{1}{n} \text{Log}(x)\right)\right)^n = \text{Exp}(\text{Log}(x)) = x.$$

Die abstrakte definierte Potenzfunktion stimmt also für $\alpha = 1/n$ überein mit der in Beispiel 5.6.3.i) definierten Wurzelfunktion; wie in Satz 3.9.2 erhalten wir dasselbe Resultat für jedes $\alpha \in \mathbb{Q}$.

Analog zu i) wollen wir nun den Limes des Ausdrucks $x^\alpha \text{Log}(x)$ für $x \downarrow 0$ bestimmen. Nach Substitution $y = -\alpha \text{Log}(x)$ mit $y \rightarrow \infty$ für $x \downarrow 0$ erhalten wir $x^\alpha = e^{-y}$ und

$$\lim_{x \downarrow 0} (x^\alpha \text{Log}(x)) = -\frac{1}{\alpha} \lim_{y \rightarrow \infty} (e^{-y} y) = 0.$$

iii) Da Exp stetig, folgt mit ii) nun auch

$$\lim_{x \downarrow 0} x^x = \lim_{x \downarrow 0} \left(\text{Exp}(x \text{Log}(x))\right) = \text{Exp}(0) = 1;$$

vergleiche Beispiel 3.2.1.iii).

Eine weitere Anwendung des Mittelwertsatzes erhalten wir durch Koppelung von Korollar 6.2.1.ii) mit Satz 5.6.5.

Satz 6.2.2. (Umkehrsatz) Sei $f:]a, b[\rightarrow \mathbb{R}$ 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 $f:]a, b[\rightarrow]c, d[$ bijektiv, und die Umkehrfunktion $f^{-1}:]c, d[\rightarrow \mathbb{R}$ ist differenzierbar mit

$$\forall x \in]a, b[: (f^{-1})'(f(x)) = (f'(x))^{-1},$$

oder

$$\forall y \in]c, d[: (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Beweis. Gemäss Korollar 6.2.1.ii) ist f streng monoton wachsend und zudem stetig nach Satz 6.1.1. Nach Satz 5.6.5 ist $f:]a, b[\rightarrow]c, d[$ bijektiv, und f^{-1} ist stetig.

Behauptung. f^{-1} ist differenzierbar mit $(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}$ für jedes x_0 .

Beweis. Fixiere $y_0 = f(x_0)$. Sei $(y_k)_{k \in \mathbb{N}} \subset]c, d[$ mit

$$y_k = f(x_k) \rightarrow y_0 \quad (k \rightarrow \infty), \quad y_k \neq y_0, \quad k \in \mathbb{N}.$$

Es folgt $x_k \neq x_0$ für alle k . Da f^{-1} stetig, gilt zudem

$$x_k = f^{-1}(y_k) \rightarrow x_0 = f^{-1}(y_0) \quad (k \rightarrow \infty),$$

also

$$\frac{f^{-1}(y_k) - f^{-1}(y_0)}{y_k - y_0} = \frac{x_k - x_0}{f(x_k) - f(x_0)} = \frac{1}{\frac{f(x_k) - f(x_0)}{x_k - x_0}} \rightarrow \frac{1}{f'(x_0)}$$

für $k \rightarrow \infty$, wie gewünscht. \square

Nach der Substitution $f(x) = y$ erhält man auch die zweite Darstellung von $(f^{-1})'$. \square

Beispiel 6.2.4. i) $Exp: \mathbb{R} \rightarrow]0, \infty[$ ist differenzierbar mit $Exp' = Exp > 0$. Also ist $Log = Exp^{-1}:]0, \infty[\rightarrow \mathbb{R}$ differenzierbar mit

$$Log'(Exp(x_0)) = \frac{1}{Exp'(x_0)} = \frac{1}{Exp(x_0)},$$

oder, nach Substitution von $y = Exp(x_0)$:

$$\forall y > 0: Log'(y) = \frac{1}{y}.$$

ii) Wir können nun auch die in Beispiel 6.2.3.ii) 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 6.1.3 erhalten wir

$$\frac{dx^\alpha}{dx} \Big|_{x=x_0} = \underbrace{Exp'(\alpha Log(x_0))}_{=x_0^\alpha} \alpha \underbrace{Log'(x_0)}_{=\frac{1}{x_0}} = \alpha x_0^{\alpha-1}.$$

Für $\alpha = n \in \mathbb{N}$ stimmt dies überein mit der Regel für die Ableitung der Potenzfunktion $f(x) = x^n$ gemäss Beispiel 6.1.2.i); auch für deren Umkehrfunktion $f^{-1}(y) = \sqrt[n]{y}:]0, \infty[\rightarrow \mathbb{R}$ liefert Satz 6.2.2 mit

$$(\sqrt[n]{y})' \Big|_{y=y_0} = (f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))} = \frac{1}{n \sqrt[n]{y_0}^{1-n}} = \frac{1}{n} y_0^{\frac{1}{n}-1};$$

dasselbe Resultat.

iii) Mit Ketten- und Produktregel sowie i) folgt nun auch:

$$\forall x > 0: \frac{d}{dx}(x^x) = \frac{d}{dx}(\exp(x \log(x))) = x^x \frac{d}{dx}(x \log(x)) = x^x(1 + \log(x)).$$

6.3 Die trigonometrischen Funktionen

Wir können nun endlich auch die lange vermutete Verbindung zwischen den über ihre Potenzreihe definierten Funktionen Sin und Cos gemäss Abschnitt 3.9. und den elementargeometrisch erklärten Funktionen sin und cos herstellen.

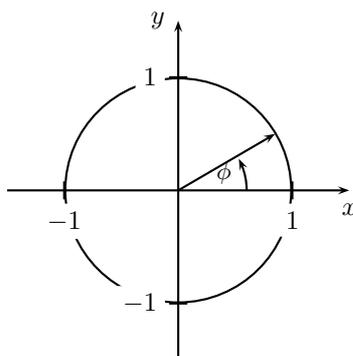
Satz 6.3.1. (Euler) Für alle $\phi \in \mathbb{R}$ gilt

$$|Exp(i\phi)|^2 = Cos^2(\phi) + Sin^2(\phi) = 1,$$

und

$$Exp(i\phi) = Cos(\phi) + iSin(\phi) = \cos \phi + i \sin \phi = e^{i\phi},$$

wobei $\cos \phi$, $\sin \phi$ Real-, bzw. Imaginärteil der Zahl $z = e^{i\phi} \in \mathbb{C}$ mit $|z| = 1$ und Polarwinkel ϕ bezeichnen (vergleiche Skizze).



Beweis. i) Da $Cos(-\phi) = Cos(\phi)$, $Sin(-\phi) = -Sin(\phi)$ für alle $\phi \in \mathbb{R}$, gilt

$$\overline{Exp(i\phi)} = Cos(\phi) - iSin(\phi) = Exp(-i\phi),$$

also

$$\forall \phi \in \mathbb{R} : |Exp(i\phi)|^2 = Exp(i\phi) \cdot \overline{Exp(i\phi)} = Exp(i\phi) \cdot Exp(-i\phi) = 1.$$

ii) Gemäss Beispiel 6.1.1.iv) gilt weiter

$$\frac{d}{d\phi} Exp(i\phi) = iExp(i\phi)$$

mit

$$\left| \frac{d}{d\phi} Exp(i\phi) \right| = |Exp(i\phi)| = 1;$$

die Kurve $\phi \mapsto Exp(i\phi)$ durchläuft also den Einheitskreis im Gegenuhrzeigersinn mit Geschwindigkeit 1. Da $Exp(0) = 1$, stimmt das Argument ϕ des Punktes $Exp(i\phi) \in \mathbb{C}$ überein mit der Bogenlänge am Einheitskreis.

□

Korollar 6.3.1. $\forall z \in \mathbb{C} : \text{Exp}(z + 2\pi i) = \text{Exp}(z)$.

Beweis. Gemäss Satz 6.3.1 gilt $\text{Exp}(2\pi i) = e^{2\pi i} = 1$, und die Behauptung folgt mit Korollar 3.8.1. \square

Polarform: Jede Zahl $z \in \mathbb{C}$ besitzt nach Satz 6.3.1 eine Darstellung

$$z = re^{i\phi} = r(\cos \phi + i \sin \phi),$$

wobei $r = |z|$, $\phi \in \mathbb{R}$. Falls $z \neq 0$, so ist das *Argument* ϕ bis auf Vielfache von 2π eindeutig bestimmt. Die Zahlen (r, ϕ) heissen *Polarkoordinaten* für z .

Das Additionstheorem Korollar 3.8.1 für die Exponentialfunktion ergibt für alle $\phi, \psi \in \mathbb{R}$ die Beziehung $e^{i\phi}e^{i\psi} = e^{i(\phi+\psi)}$. Somit folgt für

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

die einfache Darstellung

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

Beispiel 6.3.1. i) $i + 1 = \sqrt{2}e^{i\frac{\pi}{4}}$; 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) $\sqrt{3} - i = 2e^{-i\frac{\pi}{6}}$;

iii) Mit i) und ii) erhalten wir nun sehr leicht:

$$\frac{(i + 1)^6}{(\sqrt{3} - i)^3} = \frac{2^3 e^{i\frac{3\pi}{2}}}{2^3 e^{-\frac{\pi}{2}}} = e^{2\pi i} = 1.$$

Bemerkung 6.3.1. i) Mit Hilfe der Polarform kann man in \mathbb{C} die Gleichung $z^2 = c$ für jede Zahl $c = se^{i\psi}$ leicht lösen. Der Ansatz $z = re^{i\phi}$ führt auf

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

das heisst,

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

oder

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

ii) Allgemein gilt für jede Zahl

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

dass

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

die n verschiedenen Lösungen der Gleichung

$$z^n = c$$

beschreibt. Insbesondere sind für jedes $n \in \mathbb{N}$ die Zahlen $z_k = e^{2\pi ik/n}$, $0 \leq k < n$, die n -ten Einheitswurzeln.

iii) Korollar 6.3.1 zeigt, 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. Ausser für reelle $z > 0$ oder für $\alpha \in \mathbb{Z}$ sind daher Ausdrücke der Form z^α im Allgemeinen nicht eindeutig definiert.

Beispielsweise könnte man mit $i = e^{\pi i/2}$ für i^i den Ausdruck setzen

$$i^i = (e^{\pi i/2})^i = e^{-\frac{\pi}{2}};$$

mit $i = e^{2\pi ik + \pi i/2}$, $k \in \mathbb{Z}$, wäre jedoch auch

$$i^i = e^{-2\pi k - \pi/2}$$

für beliebiges $k \in \mathbb{Z}$ gleichermassen plausibel. Diese und ähnliche Fragen werden in der Vorlesung *Funktionentheorie/Komplexe Analysis* vertieft; sie führen hin zur Idee der *Riemannschen Fläche*.

6.3.1 Zyklometrische Funktionen (Arcus-Funktionen).

Im folgenden schreiben wir *cos* statt *Cos*, etc. Mit Satz 6.2.2 können wir diese Funktionen auf geeigneten Intervallen auch umkehren.

i) Da $\sin' = \cos > 0$ in $] -\frac{\pi}{2}, \frac{\pi}{2}[$, besitzt $\sin:] -\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow] -1, 1[$ gemäss Satz 6.2.2 eine differenzierbare Umkehrfunktion $\arcsin = \sin^{-1}:] -1, 1[\rightarrow] -\frac{\pi}{2}, \frac{\pi}{2}[$, 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 6.3.1.

ii) Analog besitzt $\cos:]0, \pi[\rightarrow] -1, 1[$ mit $\cos' = -\sin < 0$ in $]0, \pi[$ die differenzierbare Umkehrfunktion $\arccos:] -1, 1[\rightarrow]0, \pi[$, und

$$\arccos'(x) = \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1.$$

iii) Die Tangensfunktion $\tan = \frac{\sin}{\cos}:] -\frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathbb{R}$ 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 $\arctan: \mathbb{R} \rightarrow] -\frac{\pi}{2}, \frac{\pi}{2}[$ mit

$$\arctan'(x) = \frac{1}{\tan'(\arctan x)} = \frac{1}{1+x^2}, \quad x \in \mathbb{R}.$$

6.3.2 Hyperbel- und Areafunktionen.

Die Hyperbelfunktionen

$$\begin{aligned}\cosh x &= \frac{e^x + e^{-x}}{2} : \mathbb{R} \rightarrow [1, \infty[, \\ \sinh x &= \frac{e^x - e^{-x}}{2} : \mathbb{R} \rightarrow \mathbb{R}, \\ \tanh x &= \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} : \mathbb{R} \rightarrow]-1, 1[\end{aligned}$$

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

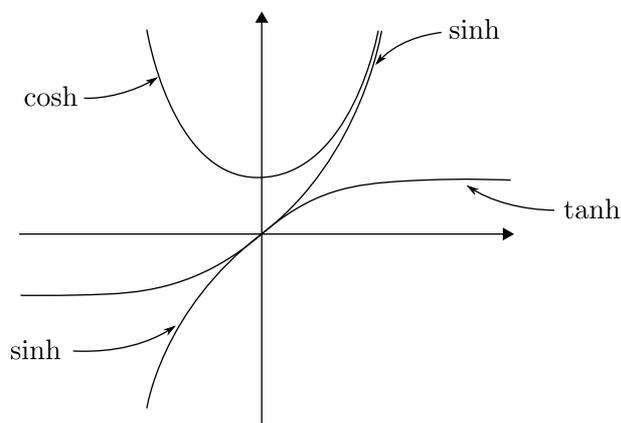
$$\cosh^2(x) - \sinh^2(x) = e^x \cdot e^{-x} = 1.$$

Insbesondere existieren die Areafunktionen

$$\begin{aligned} \operatorname{arsinh} &= \sinh^{-1} : \mathbb{R} \rightarrow \mathbb{R}, \\ \operatorname{arcosh} &= \cosh^{-1} : [1, \infty[\rightarrow \mathbb{R}, \\ \operatorname{artanh} &= \tanh^{-1} :]-1, 1[\rightarrow \mathbb{R} \end{aligned}$$

mit

$$\begin{aligned} \operatorname{arsinh}'(x) &= \frac{1}{\cosh(\operatorname{arsinh} x)} = \frac{1}{\sqrt{1+x^2}}, \quad x \in \mathbb{R}, \\ \operatorname{artanh}'(x) &= \frac{1}{(1-\tanh^2)(\operatorname{artanh}(x))} = \frac{1}{1-x^2}, \quad -1 < x < 1. \end{aligned}$$



6.4 Funktionen der Klasse C^1

Sei $\Omega \subset \mathbb{R}$ offen, $f: \Omega \rightarrow \mathbb{R}$ differenzierbar.

Definition 6.4.1. f heißt von der Klasse C^1 , falls die Ableitungsfunktion $x \mapsto f'(x)$ stetig ist.

Notation:

$$C^1(\Omega) = \{f: \Omega \rightarrow \mathbb{R}; f, f' \text{ stetig}\}.$$

Beispiel 6.4.1. i) Die Funktionen \exp , \cos , \sin und Polynome sind in C^1 .

ii) Sei die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ 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

$$f'(0) = \lim_{x \rightarrow 0, x \neq 0} \frac{f(x)}{x} = \lim_{x \rightarrow 0} \left(x \sin \frac{1}{x}\right) = 0;$$

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 6.4.1. Sei $(f_k)_{k \in \mathbb{N}}$ eine Folge in $C^1(\Omega)$ mit

$$f_k \xrightarrow{glm} f, \quad f'_k \xrightarrow{glm} g \quad (k \rightarrow \infty),$$

wobei $f, g: \Omega \rightarrow \mathbb{R}$. Dann gilt $f \in C^1(\Omega)$ und $f' = g$.

Beweis. Nach Satz 5.7.1 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 6.2.1

$$\frac{f(x) - f(x_0)}{x - x_0} - g(x_0) = \lim_{k \rightarrow \infty} \frac{f_k(x) - f_k(x_0)}{x - x_0} - g(x_0) = \lim_{k \rightarrow \infty} f'_k(x_k) - g(x_0),$$

wobei $x_k = x_0 + t_k(x - x_0)$ mit geeignetem $0 < t_k < 1$, $k \in \mathbb{N}$. Mit $f'_k \xrightarrow{glm} g$ ($k \rightarrow \infty$), und da für jedes $k \in \mathbb{N}$ gilt $|x_k - x_0| < |x - x_0|$, folgt mit der Stetigkeit

von g :

$$\begin{aligned} \left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| &\leq \limsup_{k \rightarrow \infty} (|f'_k(x_k) - g(x_k)| + |g(x_k) - g(x_0)|) \\ &\leq \lim_{k \rightarrow \infty} \sup_{y \in \Omega} |f'_k(y) - g(y)| + \sup_{|y - x_0| < |x - x_0|} |g(y) - g(x_0)| \rightarrow 0 \quad (x \rightarrow x_0). \end{aligned}$$

Also ist f an der Stelle x_0 differenzierbar mit $f'(x_0) = g(x_0)$. \square

Offenbar gilt Satz 6.4.1 auch für vektorwertige Funktionen.

Beispiel 6.4.2. Die gleichmässige Konvergenz $f'_k \rightarrow g$ ($k \rightarrow \infty$) ist notwendig für die Aussage von Satz 6.4.1, wie das folgende Beispiel zeigt. Sei

$$f_k(x) = \sqrt{x^2 + (1/k)^2}, \quad |x| < 1, \quad k \in \mathbb{N}.$$

Mit der Abschätzung

$$\begin{aligned} 0 \leq f_k(x) - |x| &= (\sqrt{x^2 + (1/k)^2} - |x|) \cdot \frac{\sqrt{x^2 + (1/k)^2} + |x|}{\sqrt{x^2 + (1/k)^2} + |x|} \\ &= \frac{(1/k)^2}{\sqrt{x^2 + (1/k)^2} + |x|} \leq \frac{(1/k)^2}{1/k} = \frac{1}{k} \rightarrow 0 \quad (k \rightarrow \infty) \end{aligned}$$

erhalten wir gleichmässige Konvergenz $f_k \xrightarrow{\text{glm}} f$ für $|x| < 1$, wobei $f(x) = |x|$. Zudem konvergiert

$$f'_k(x) = \frac{x}{\sqrt{x^2 + (1/k)^2}} \rightarrow \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

punktweise für $k \rightarrow \infty$; 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

$$f(x) = \sum_{k=0}^{\infty} a_k x^k, \quad |x| < \rho = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}} \leq \infty$$

die zugehörige Potenzreihe $f = \lim_{n \rightarrow \infty} f_n$, wobei

$$f_n(x) = \sum_{k=0}^n a_k x^k, \quad x \in \mathbb{R}.$$

Nach Beispiel 6.1.2.ii) ist jedes f_n differenzierbar mit

$$f'_n(x) = \sum_{k=0}^n k a_k x^{k-1}.$$

Da $\sqrt[k]{k} \rightarrow 1$ ($k \rightarrow \infty$), hat die Potenzreihe

$$g(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}$$

denselben Konvergenzradius ρ wie f , und wie in Beispiel 5.7.2 folgt für $r < \rho$ gleichmässige Konvergenz

$$f_n \rightarrow f, \quad f'_n \rightarrow g \quad \text{in } B_r(0) \quad (n \rightarrow \infty).$$

Satz 6.4.1 liefert somit das folgende Resultat.

Satz 6.4.2. Eine Potenzreihe $f(x) = \sum_{k=0}^{\infty} a_k x^k$ ist im Innern ihres Konvergenzkreises differenzierbar, und

$$f'(x) = \sum_{k=0}^{\infty} k a_k x^{k-1}.$$

Beispiel 6.4.3. i) Es gilt

$$\text{Exp}'(x) = \sum_{k=0}^{\infty} k \frac{x^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{x^{k-1}}{(k-1)!} = \text{Exp}(x).$$

ii) Sei $f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$, $|x| < 1$. Dann folgt mit der Quotientenregel, bzw. mit Satz 6.4.2

$$f'(x) = \frac{1}{(1-x)^2} = \sum_{k=0}^{\infty} k x^{k-1}, \quad |x| < 1;$$

vgl. Beispiel 3.8.2.

iii) Analog zu Satz 6.4.2 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 $\zeta = \lim_{l \rightarrow \infty} \zeta_l$, 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 $x \geq s > r > 1$:

$$\begin{aligned} |\zeta(x) - \zeta_l(x)| &\leq \sum_{n>l} \frac{1}{n^s} \rightarrow 0 \quad (l \rightarrow \infty) \\ |\eta(x) - \zeta'_l(x)| &\leq \sum_{n>l} \frac{\log n}{n^s} \leq \sup_{n>l} \frac{\log n}{n^{s-r}} \cdot \sum_{n>l} \frac{1}{n^r} \rightarrow 0 \quad (l \rightarrow \infty), \end{aligned}$$

wobei wir Beispiel 6.2.3.ii) benutzen. Die Behauptung folgt mit Satz 6.4.1. \square

Riemann-Funktion Sei $(q_k)_{k \in \mathbb{N}}$ eine Abzählung von $\mathbb{Q} \cap]0, 1[$.

Behauptung. Die Funktion $f :]0, 1[\rightarrow \mathbb{R}$ mit

$$f(x) = \sum_{k=1}^{\infty} 2^{-k} |x - q_k|$$

ist stetig; f ist differenzierbar an jeder Stelle $x_0 \in]0, 1[\setminus \mathbb{Q}$ mit Ableitung

$$f'(x) = g(x) := \sum_{k=1}^{\infty} 2^{-k} \text{sign}(x - q_k), \quad x \in]0, 1[\setminus \mathbb{Q},$$

und f ist nicht differenzierbar an jeder Stelle $x_0 = q_k$, $k \in \mathbb{N}$.

Beweis. Sei

$$f_n(x) = \sum_{k=1}^n 2^{-k} |x - q_k|, \quad n \in \mathbb{N}.$$

Dann ist f_n stetig und an jeder Stelle $x_0 \in]0, 1[\setminus \mathbb{Q}$ differenzierbar mit

$$f'_n(x) = \sum_{k=1}^n 2^{-k} \text{sign}(x - q_k).$$

Offenbar gilt

$$0 \leq f(x) - f_n(x) = \sum_{k>n} 2^{-k} |x - q_k| \leq \sum_{k>n} 2^{-k} = 2^{-n} \rightarrow 0 \quad (n \rightarrow \infty),$$

gleichmässig in $x \in]0, 1[$; also $f_n \xrightarrow{\text{glm.}} f$ in $]0, 1[$, und nach Satz 5.7.1 ist f stetig.

Zum Beweis der Differenzierbarkeit an jeder Stelle $x_0 \notin \mathbb{Q}$, fixiere $x_0 \in]0, 1[\setminus \mathbb{Q}$. Zu vorgegebenem $\epsilon > 0$ wähle $n_0 = n_0(\epsilon) \in \mathbb{N}$ mit

$$\sum_{k>n_0} 2^{-k} = 2^{-n_0} < \epsilon. \quad (6.4.1)$$

Setze

$$\delta = \min_{1 \leq k \leq n_0} |x_0 - q_k| > 0.$$

Die Funktion f_{n_0} ist affin in $B_\delta(x_0) \cap]0, 1[$ mit

$$\frac{f_{n_0}(x) - f_{n_0}(x_0)}{x - x_0} = f'_{n_0}(x_0) = \sum_{k=1}^{n_0} 2^{-k} \text{sgn}(x_0 - q_k) \quad (6.4.2)$$

für alle $x \in B_\delta(x_0) \cap]0, 1[\setminus \{x_0\}$. Beachte weiter, dass die Funktionen

$$\ell_k(x) = 2^{-k} |x - q_k|$$

Lipschitz stetig sind mit Lipschitzkonstante $L_k = 2^{-k}$, $k \in \mathbb{N}$. Für $x \neq x_0$ können wir daher abschätzen

$$\left| \frac{f(x) - f(x_0)}{x - x_0} - \frac{f_{n_0}(x) - f_{n_0}(x_0)}{x - x_0} \right| \leq \sum_{k>n_0} L_k = \sum_{k>n_0} 2^{-k} = 2^{-n_0} < \epsilon.$$

Mit (6.4.1) und (6.4.2) folgt

$$\limsup_{x \rightarrow x_0, x \neq x_0} \left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| \leq |f'_{n_0}(x_0) - g(x_0)| + \epsilon < 2\epsilon.$$

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

Sei abschliessend $x_0 = q_{k_0} \in \mathbb{Q} \cap]0, 1[$. Wie oben gezeigt, ist die Funktion

$$\tilde{f}(x) = \sum_{k \neq k_0} 2^{-k} |x - q_k|$$

an jeder Stelle $x_0 \in]0, 1[\setminus \mathbb{Q}$ und zusätzlich an der Stelle q_{k_0} differenzierbar. Wäre f an der Stelle q_{k_0} differenzierbar, so auch $f - \tilde{f} = \ell_{k_0}$, im Widerspruch zu Beispiel 6.1.1.ii). \square

Sei nun $\Omega \subset \mathbb{R}$ offen und beschränkt. Analog zu Abschnitt 5.7.3 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

$$\|f\|_{C^1(\overline{\Omega})} = \sup_{x \in \overline{\Omega}} |f(x)| + \sup_{x \in \overline{\Omega}} |f'(x)| = \|f\|_{C^0(\overline{\Omega})} + \|f'\|_{C^0(\overline{\Omega})} < \infty.$$

Offenbar ist $\|\cdot\|_{C^1(\overline{\Omega})}$ eine Norm auf $C^1(\overline{\Omega}; \mathbb{R}^n)$. Weiter liefern Satz 5.7.4 und Satz 6.4.1 das folgende Resultat analog zu Satz 5.7.4.

Satz 6.4.3. *Der Raum $C^1(\overline{\Omega}; \mathbb{R}^n)$ ist vollständig bzgl. $\|\cdot\|_{C^1(\overline{\Omega})}$; er ist ein Banachraum.*

Beweis. Sei $(f_k)_{k \in \mathbb{N}} \subset C^1(\overline{\Omega}; \mathbb{R}^n)$ Cauchy-Folge. Dann sind $(f_k)_{k \in \mathbb{N}}, (f'_k)_{k \in \mathbb{N}}$ Cauchy-Folgen in $C^0(\overline{\Omega}; \mathbb{R}^n)$ mit Limites $f = \lim_{k \rightarrow \infty} f_k, g = \lim_{k \rightarrow \infty} f'_k \in C^0(\overline{\Omega}; \mathbb{R}^n)$ gemäss Satz 5.7.4, und $g = f'$ gemäss Satz 6.4.1; also $f \in C^1(\overline{\Omega}; \mathbb{R}^n)$, und $f_k \xrightarrow{C^1\text{-glm}} f (k \rightarrow \infty)$. \square

Bemerkung 6.4.1. Offenbar gilt $C^1(\overline{\Omega}) \subset C^0(\overline{\Omega})$, und $\|f\|_{C^0} \leq \|f\|_{C^1}$ für jedes $f \in C^1(\overline{\Omega})$. Gilt die Aussage von Satz 6.4.3 auch, wenn wir den Raum $C^1(\overline{\Omega})$ mit der C^0 -Norm ausstatten? – Beispielsweise im Falle $\Omega =]0, 1[$ gibt es nach Satz 5.7.5 für jedes $f \in C^0([0, 1])$ Polynome $p_k \in C^1([0, 1])$ mit $\|f - p_k\|_{C^0} \rightarrow 0 (k \rightarrow \infty)$. Falls $f \notin C^1([0, 1])$, so hat die C^0 -Cauchy-Folge $(p_k)_{k \in \mathbb{N}}$ keinen Grenzwert im Raum $C^1([0, 1])$. Versehen mit der C^0 -Norm ist der Raum $C^1([0, 1])$ also nicht vollständig. Die C^0 -Norm und die C^1 -Norm sind daher auch nicht äquivalent.

6.5 Höhere Ableitungen

6.5.1 Der Raum $C^m(\Omega)$

Iterativ können wir auch höhere Ableitungen bilden. Sei $m \in \mathbb{N}$.

Definition 6.5.1. 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

$$f^{(m)} = \frac{df^{(m-1)}}{dx} = \frac{d^m f}{dx^m} : \Omega \rightarrow \mathbb{R}$$

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 $f = f^{(0)}$, $f' = f^{(1)}$, \dots , $f^{(m)}$ stetig sind.

Notation:

$$C^m(\Omega, \mathbb{R}^n) = \{f : \Omega \rightarrow \mathbb{R}^n; f \text{ ist } m\text{-mal diffbar, } f, \dots, f^{(m)} \text{ stetig}\}.$$

Beispiel 6.5.1. 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; vergleiche Satz 6.4.2.

Falls $\Omega \subset \mathbb{R}$ offen und beschränkt, setzen wir

$$C^m(\overline{\Omega}; \mathbb{R}^n) = \{f \in C^m(\Omega; \mathbb{R}^n); \\ f = f^{(0)}, f' = f^{(1)}, \dots, f^{(m)} \text{ sind stetig auf } \overline{\Omega} \text{ ergänzbar}\}$$

und definieren

$$\|f\|_{C^m} = \sum_{j=0}^m \|f^{(j)}\|_{C^0}, \quad f \in C^m(\overline{\Omega}; \mathbb{R}^n).$$

Dann gilt offenbar

$$C^m(\overline{\Omega}; \mathbb{R}^n) \subsetneq C^{m-1}(\overline{\Omega}; \mathbb{R}^n) \subsetneq \dots \subsetneq C^0(\overline{\Omega}; \mathbb{R}^n)$$

und

$$\forall f \in C^m(\overline{\Omega}; \mathbb{R}^n) : \|f\|_{C^0} \leq \|f\|_{C^1} \leq \dots \leq \|f\|_{C^m}.$$

Analog zu Satz 6.4.3 gilt schliesslich

Satz 6.5.1. Der Raum $C^m(\overline{\Omega}; \mathbb{R}^n)$ ist vollständig bzgl. $\|\cdot\|_{C^m(\overline{\Omega}; \mathbb{R}^n)}$.

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 $\lim_{k \rightarrow \infty} f_k = f =: g_0 \in C^0(\overline{\Omega})$. Analog existieren $\lim_{k \rightarrow \infty} f_k^{(j)} =: g_j \in C^0(\overline{\Omega})$, $0 < j \leq m$. Da $f_k^{(j+1)} = df_k^{(j)}/dx$, folgt mit Satz 6.4.1 die Beziehung $g_{j+1} = g_j'$, $0 \leq j < m$; also

$$g_1 = f', \quad g_2 = g_1' = f'', \quad \dots, \quad g_m = \dots = f^{(m)},$$

und $f \in C^m(\overline{\Omega})$ mit $\|f_k - f\|_{C^m(\overline{\Omega})} \rightarrow 0$ ($k \rightarrow \infty$), wie gewünscht. \square

6.5.2 Taylor-Formel

Sei $\Omega =]a, b[$, $-\infty < a < b < \infty$, und $m \in \mathbb{N}$.

Satz 6.5.2. Sei $f \in C^{m-1}([a, b])$ auf $]a, b[$ m -mal differenzierbar. Dann gibt es $\xi \in]a, b[$ mit

$$f(b) = f(a) + f'(a)(b-a) + f''(a)\frac{(b-a)^2}{2} + \dots \\ \dots + f^{(m-1)}(a)\frac{(b-a)^{m-1}}{(m-1)!} + f^{(m)}(\xi)\frac{(b-a)^m}{m!}.$$

Beweis. Wir führen den Satz zurück auf den Mittelwertsatz, Satz 6.2.1. Betrachte die Funktion

$$g(x) = f(x) + f'(x)(b-x) + \dots \\ + f^{(m-1)}(x)\frac{(b-x)^{m-1}}{(m-1)!} + K\frac{(b-x)^m}{m!} - f(b), \quad (6.5.1)$$

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 6.2.1 existiert ein $\xi \in]a, b[$ mit $g'(\xi) = 0$; das heisst,

$$0 = f'(\xi) + \left(f''(\xi)(b-\xi) - f'(\xi) \right) + \left(f'''(\xi)\frac{(b-\xi)^2}{2} - f''(\xi)(b-\xi) \right) + \dots \\ + \left(f^{(m)}(\xi)\frac{(b-\xi)^{m-1}}{(m-1)!} - f^{(m-1)}(\xi)\frac{(b-\xi)^{m-2}}{(m-2)!} \right) - K\frac{(b-\xi)^{m-1}}{(m-1)!} \\ = (f^{(m)}(\xi) - K)\frac{(b-\xi)^{m-1}}{(m-1)!},$$

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 (6.5.1) die Behauptung. \square

Bemerkung 6.5.1. Die *Tangente* an $f \in C^1([a, b])$ im Punkt $x_0 = a$,

$$T_1 f(x; a) = f(a) + f'(a)(x-a),$$

die *Schmiegeparabel* für $f \in C^2([a, b])$,

$$T_2 f(x; a) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2},$$

und allgemein das *Taylor-Polynom m -ter Ordnung* für $f \in C^m([a, b])$,

$$T_m f(x; a) = f(a) + f'(a)(x-a) + \dots + f^{(m)}(a)\frac{(x-a)^m}{m!},$$

hat nach Satz 6.5.2 die folgende Approximationseigenschaft. Für $a < x < b$ gilt

$$f(x) - T_m f(x; a) = (f^{(m)}(\xi) - f^{(m)}(a)) \cdot \frac{(x-a)^m}{m!} =: r_m f(x; a)$$

für ein $\xi \in]a, x[$. Für den Restterm $r_m f$ erhalten wir die Abschätzung

$$|r_m f(x; a)| \leq \sup_{a < \xi < x} |f^{(m)}(\xi) - f^{(m)}(a)| \frac{(x-a)^m}{m!}.$$

Falls $f \in C^{m+1}$, können wir dies mit Satz 6.2.1 noch verbessern zu

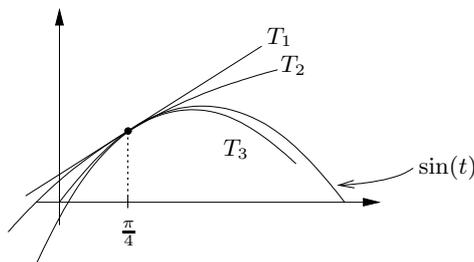
$$|r_m f(x; a)| \leq \sup_{a < \xi < x} |f^{(m+1)}(\xi)| \frac{(x-a)^{m+1}}{m!}.$$

Beispiel 6.5.2. Was ist der Sinus von $47^\circ \doteq \frac{\pi}{4} + \frac{2\pi}{180}$? – Mit $\sin' = \cos$, $\sin'' = \cos' = -\sin$, usw., folgt aus Bemerkung 6.5.1 bei Wahl von $m = 2$ die Näherung

$$\begin{aligned} \sin\left(\frac{\pi}{4} + \frac{\pi}{90}\right) &= \sin\left(\frac{\pi}{4}\right) + \cos\left(\frac{\pi}{4}\right) \cdot \frac{\pi}{90} - \sin\left(\frac{\pi}{4}\right) \frac{\pi^2}{2 \cdot 90^2} + r_2 \\ &= \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \cdot \frac{\pi}{90} - \frac{\sqrt{2}}{2} \cdot \frac{\pi^2}{2 \cdot 90^2} + r_2, \end{aligned}$$

wobei

$$|r_2| \leq \frac{\pi^3}{2 \cdot 90^3} \approx 10^{-5}.$$



6.5.3 Lokale Extrema.

Sei $\Omega \subset \mathbb{R}$ offen, $f: \Omega \rightarrow \mathbb{R}$.

Definition 6.5.2. Ein $x_0 \in \Omega$ heißt (strikte) lokale Minimalstelle von f , falls in einer Umgebung U von x_0 gilt

$$\forall x \in U : f(x) \geq f(x_0), \quad (\text{bzw. } \forall x \in U \setminus \{x_0\} : f(x) > f(x_0)).$$

Falls f an einer lokalen Minimalstelle x_0 differenzierbar ist, so folgt wie im Beweis von Satz 6.2.1

$$0 \leq \lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) = \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \leq 0;$$

also $f'(x_0) = 0$. Allgemein gilt der folgende Satz.

Korollar 6.5.1. Sei $f \in C^m(\Omega)$, $x_0 \in \Omega$ mit $f'(x_0) = \dots = f^{(m-1)}(x_0) = 0$.

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 6.5.2 existiert für $x \in \Omega$ nahe x_0 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

$$f^{(m)}(x_0) = \lim_{\xi \rightarrow x_0} f^{(m)}(\xi) = \begin{cases} m! \lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{(x - x_0)^m} \geq 0, \\ m! \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{(x - x_0)^m} \leq 0; \end{cases}$$

also $f^{(m)}(x_0) = 0$.

ii) Für $m = 2k$, $x \neq x_0$ gilt $(x - x_0)^m > 0$. Falls $f^{(m)}(x_0) = \lim_{\xi \rightarrow x_0} f^{(m)}(\xi) > 0$, folgt $f(x) > f(x_0)$ für $x \neq x_0$ nahe x_0 ; also ist x_0 striktes lokales Minimum. \square

Beispiel 6.5.3. i) Sei $f(x) = x^4 - x^2 + 1$, $x \in \mathbb{R}$. Nach Korollar 6.5.1.i) ist notwendig für das Vorliegen einer Extremalstelle im Punkt x_0 die Bedingung

$$f'(x_0) = 4x_0^3 - 2x_0 = 2(2x_0^2 - 1)x_0 = 0;$$

das heisst,

$$x_0 \in \left\{ -\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right\}.$$

Nach Korollar 6.5.1.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 $a_1, \dots, a_n \in \mathbb{R}$. Gesucht ist die "least square"-Näherung $x_0 \in \mathbb{R}$ von $(a_i)_{1 \leq i \leq n}$ mit

$$f(x_0) = \sum_{k=1}^n (x_0 - a_k)^2 = \min_x f(x).$$

Beachte $f(x) \rightarrow \infty$ ($|x| \rightarrow \infty$); also existiert $x_0 \in \mathbb{R}$ mit $f(x_0) = \min_x f(x)$.

Korollar 6.5.1.i) liefert die notwendige Bedingung:

$$f'(x_0) = 2 \sum_{k=1}^n (x_0 - a_k) = 2nx_0 - 2 \sum_{k=1}^n a_k = 0;$$

das arithmetische Mittel

$$x_0 = \frac{1}{n} \sum_{k=1}^n a_k$$

ist also die einzig mögliche Minimalstelle. Zur Probe bestimmen wir noch

$$f'' \equiv 2n > 0.$$

Der Punkt x_0 ist also tatsächlich die gesuchte Minimalstelle.

6.5.4 Konvexe Funktionen.

Sei $-\infty < a < b < \infty$.

Satz 6.5.3. Sei $f \in C^2(]a, b[)$ mit $f'' \geq 0$. Dann gilt für alle $x_0 \neq x_1 \in]a, b[$, $0 < t < 1$ die Ungleichung:

$$f(tx_1 + (1-t)x_0) \leq tf(x_1) + (1-t)f(x_0), \quad (6.5.2)$$

und die Ungleichung ist strikt, falls $f'' > 0$.

Beweis. i) Fixiere $x_0 \neq x_1 \in]a, b[$. 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

$$g(0) = g(1) = 0, \quad g''(t) = f''(tx_1 + (1-t)x_0)(x_1 - x_0)^2 \geq 0, \quad 0 \leq t \leq 1.$$

Nimm widerspruchswise an,

$$\max_{0 \leq t \leq 1} g(t) = g(t_{max}) > 0, \quad \text{wobei } 0 < t_{max} < 1.$$

Gemäss Korollar 6.5.1.i) folgt $g'(t_{max}) = 0$. Nach Satz 6.5.2 gibt es $\tau \in]t_{max}, 1[$ mit

$$0 = g(1) = g(t_{max}) + g'(t_{max})(1 - t_{max}) + g''(\tau) \frac{(1 - t_{max})^2}{2} \geq g(t_{max}) > 0,$$

und es folgt der gewünschte Widerspruch.

ii) Analog erhält man einen Widerspruch im Falle $f'' > 0$, wenn man annimmt, dass $g(t_0) = 0$ für ein $0 < t_0 < 1$. \square

Definition 6.5.3. Eine Funktion $f:]a, b[\rightarrow \mathbb{R}$ mit der Eigenschaft (6.5.2) heisst **konvex**; sie heisst **strikt konvex**, wenn gilt

$$\forall x_0 \neq x_1 \in]a, b[, \quad 0 < t < 1: \quad f(tx_1 + (1-t)x_0) < tf(x_1) + (1-t)f(x_0).$$

Bemerkung 6.5.2. i) Offenbar ist f (strikt) konvex genau dann, wenn der Epigraph von f ,

$$\text{epi}(f) := \{(x, y); x \in]a, b[, y \geq f(x)\} \subset \mathbb{R}^2,$$

eine (strikt) konvexe Menge ist.

ii) Satz 6.5.3 liefert ein bequemes Kriterium für (strikte) Konvexität von f .

Beispiel 6.5.4. i) Die Funktion $f(x) = |x|$, $x \in \mathbb{R}$, ist konvex.

ii) Mit $f, g:]a, b[\rightarrow \mathbb{R}$ ist auch $f + g:]a, b[\rightarrow \mathbb{R}$ konvex, mit $f_k:]a, b[\rightarrow \mathbb{R}$, $k \in \mathbb{N}$, auch $f = \sum_{k=1}^{\infty} f_k$. Insbesondere ist die Riemann-Funktion aus Beispiel 6.4.3.iv) konvex.

iii) $\exp'' = \exp > 0$; also ist die Funktion \exp strikt konvex.

iv) Sei $f(x) = x \log x$, $x > 0$. Es gilt

$$f'(x) = \log x + 1, \quad f''(x) = \frac{1}{x} > 0;$$

also ist f strikt konvex gemäss Satz 6.5.3.

v) Für eine feste Zahl $\alpha > 1$ sei $f(x) = x^\alpha = \exp(\alpha \log x)$, $x > 0$. Da

$$f'(x) = \alpha x^{\alpha-1}, \quad f''(x) = \alpha(\alpha-1)x^{\alpha-2} > 0,$$

ist f strikt konvex gemäss Satz 6.5.3.

Die Eigenschaft (6.5.2) gilt analog auch für mehr als zwei Punkte.

Satz 6.5.4. (Jensen) Sei $f:]a, b[\rightarrow \mathbb{R}$ konvex. Dann gilt für beliebige Punkte $x_1, \dots, x_N \in]a, b[$ und Zahlen $0 \leq t_1, \dots, t_N \leq 1$ mit $\sum_{i=1}^N t_i = 1$ die Ungleichung

$$f\left(\sum_{i=1}^N t_i x_i\right) \leq \sum_{i=1}^N t_i f(x_i).$$

Beweis (Induktion nach N). $N = 1$ ist klar; ebenso $N = 2$ nach Definition.

$N \rightarrow N + 1$: 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

$$f\left(\sum_{i=1}^{N+1} t_i x_i\right) = f(t_1 x_1 + (1-t_1)x_0) \leq t_1 f(x_1) + (1-t_1)f(x_0)$$

Da nach Induktions-Annahme gilt

$$f(x_0) = f\left(\sum_{i=2}^{N+1} \frac{t_i}{1-t_1} x_i\right) \leq \sum_{i=2}^{N+1} \frac{t_i}{1-t_1} f(x_i),$$

folgt die Behauptung. □

Beispiel 6.5.5. (Vergleich von arithmetischem und geometrischem Mittel) Für alle $0 < x_1, \dots, x_n < \infty$, $0 \leq \alpha_1, \dots, \alpha_n \leq 1$ mit $\sum_{i=1}^n \alpha_i = 1$ gilt

$$\prod_{i=1}^n x_i^{\alpha_i} \leq \sum_{i=1}^n \alpha_i x_i.$$

Beweis. Da die Funktion \exp nach Beispiel 6.5.4.i) konvex ist, folgt die Aussage mit Satz 6.5.4 aus der Darstellung

$$\begin{aligned} \prod_{i=1}^n x_i^{\alpha_i} &= \prod_{i=1}^n \exp(\alpha_i \log x_i) = \exp\left(\sum_{i=1}^n \alpha_i \log x_i\right) \\ &\stackrel{(\text{Satz 6.5.4})}{\leq} \sum_{i=1}^n \alpha_i \exp(\log x_i) = \sum_{i=1}^n \alpha_i x_i. \end{aligned}$$

□

Insbesondere erhalten wir für $\alpha_i = \frac{1}{n}$, $1 \leq i \leq n$,

$$\sqrt[n]{\prod_{i=1}^n x_i} \leq \frac{1}{n} \sum_{i=1}^n x_i.$$

Konjugiert konvexe Funktion. Sei $\Omega =]a, b[$, $f : \Omega \rightarrow \mathbb{R}$ konvex. Für $y \in \mathbb{R}$ definiere:

$$f^*(y) = \sup_{x \in \Omega} \underbrace{(xy - f(x))}_{=: \ell_x(y)} \leq \infty. \quad (6.5.3)$$

mit

$$\text{epi}(f^*) = \bigcap_{x \in \Omega} \text{epi}(\ell_x).$$

Da der Durchschnitt konvexer Mengen konvex ist, ist $\text{epi}(f^*)$ konvex. Also ist $f^* : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ konvex und

$$\Omega^* = \{y \in \mathbb{R}; f^*(y) < \infty\} \subset \mathbb{R}$$

ist konvex, ein Intervall.

Definition 6.5.4. Die Funktion $f^* : \Omega^* \rightarrow \mathbb{R}$ heisst die zu f **konjugiert konvexe Funktion**.

Beispiel 6.5.6. i) Sei $1 < p < \infty$. Für $x > 0$ sei

$$f(x) = \frac{x^p}{p}$$

mit $f''(x) = (p-1)x^{p-2} > 0$ für $x > 0$. Also ist f konvex mit konjugierter Funktion

$$f^*(y) = \sup_{x>0} (xy - f(x)) = \sup_{x>0} \left(xy - \frac{x^p}{p}\right).$$

Das Supremum wird für jedes vorgegebene $y > 0$ angenommen in $x_0 > 0$ mit

$$0 = \left(xy - \frac{x^p}{p}\right)' \Big|_{x=x_0} = y - x_0^{p-1}.$$

Das heisst, $x_0 = y^{\frac{1}{p-1}}$, und mit $q := \frac{p}{p-1}$ erhalten wir

$$f^*(y) = x_0 y - \frac{x_0^p}{p} = y^{\frac{p}{p-1}} \left(1 - \frac{1}{p}\right) = \frac{p-1}{p} y^{\frac{p}{p-1}} = \frac{y^q}{q}, \quad y > 0.$$

Beachte

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{p} + \frac{p-1}{p} = 1.$$

Derartige Exponenten $1 \leq p, q \leq \infty$ heissen *konjugiert*. Insbesondere gilt

$$(f^*)^* = f.$$

ii) Zu $f(x) = e^x$, $x \in \mathbb{R}$ ist

$$f^*(y) = \begin{cases} y(\log(y) - 1), & y > 0, \\ 0, & y = 0, \\ \infty, & y < 0. \end{cases}$$

Sei $f : \Omega \rightarrow \mathbb{R}$ konvex mit konjugierter Funktion $f^* : \Omega^* \rightarrow \mathbb{R}$. Aus (6.5.3) folgt unmittelbar die Abschätzung:

$$\forall x \in \Omega, y \in \Omega^* : xy \leq f(x) + f^*(y).$$

Mit Beispiel 6.5.6 folgt:

Satz 6.5.5. i) Seien $1 < p, q < \infty$ konjugiert mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann folgt:

$$\forall x, y > 0 : xy \leq \frac{|x|^p}{p} + \frac{|y|^q}{q} \quad (\text{Youngsche Ungleichung}).$$

ii) Weiter gilt:

$$\forall x \in \mathbb{R}, y > 0 : xy \leq e^x + y(\log(y) - 1).$$

Anwendung. Sei $1 < p < \infty$ und $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}$ die Norm

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p}, \quad x \in \mathbb{R}^n.$$

Mit dem zu p konjugierten Exponenten $q = \frac{p}{p-1}$ erhalten wir die folgende “duale” Charakterisierung der Norm.

Behauptung. $\|x\|_p = \sup_{y \in \mathbb{R}^n; \|y\|_q \leq 1} \langle x, y \rangle$, $\frac{1}{p} + \frac{1}{q} = 1$.

Beweis. Für $x = 0$ gilt die Behauptung trivialerweise. Betrachten wir für $x \neq 0$ den Vektor $\frac{x}{\|x\|_p}$, dürfen wir oBdA annehmen $\|x\|_p = 1$. Für $y \in \mathbb{R}^n$ mit $\|y\|_q \leq 1$ folgt aus Satz 6.5.5:

$$\begin{aligned} \langle x, y \rangle &= \sum_{i=1}^n x_i y_i \leq \sum_{i=1}^n \left(\frac{|x_i|^p}{p} + \frac{|y_i|^q}{q} \right) \\ &= \frac{\|x\|_p^p}{p} + \frac{\|y\|_q^q}{q} \leq \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p. \end{aligned}$$

mit Gleichheit für $y \in \mathbb{R}^n$ mit $y_i = x_i |x_i|^{p-2}$, $1 \leq i \leq n$. □

Es folgt für alle $x, y \in \mathbb{R}^n$ die *Minkowski-Ungleichung*:

$$\|x + y\|_p = \sup_{\substack{z \in \mathbb{R}^n \\ \|z\|_q \leq 1}} \langle x + y, z \rangle = \sup_{\substack{z \in \mathbb{R}^n \\ \|z\|_q \leq 1}} (\langle x, z \rangle + \langle y, z \rangle) \leq \|x\|_p + \|y\|_p;$$

das heisst, die Dreiecks-Ungleichung für die p -Norm aus Abschnitt 5.4.

Weiter folgt für alle $x, y \in \mathbb{R}^n \setminus \{0\}$ aus der Abschätzung

$$\frac{\langle x, y \rangle}{\|y\|_q} = \langle x, y/\|y\|_q \rangle \leq \|x\|_p$$

die allgemeine *Cauchy-Schwarzsche Ungleichung*

$$\forall x, y \in \mathbb{R}^n : \langle x, y \rangle \leq \|x\|_p \|y\|_q.$$

6.5.5 “Glatte” Funktionen, analytische Funktionen

Sei $\Omega =]a, b[$, $f : \Omega \rightarrow \mathbb{R}$

Definition 6.5.5. *f* heisst **glatt**, falls *f* beliebig oft differenzierbar ist.

Notation:

$$C^\infty(\Omega) = \{f : \Omega \rightarrow \mathbb{R}; f \text{ glatt}\} = \bigcap_{m=0}^{\infty} C^m(\Omega).$$

Beispiel 6.5.7. i) Polynome sowie rationale Funktionen sind glatt auf ihrem Definitionsbereich.

ii) Potenzreihen sind glatt im Innern des Konvergenzkreises.

iii) Die Funktionen *exp*, *sin*, *cos* sind glatt auf \mathbb{R} .

Gibt es eine Taylor-Approximation der “Ordnung $m = \infty$ ”? - Kann man jede glatte Funktion *exakt* durch eine Potenzreihe darstellen?

Definition 6.5.6. Sei $f \in C^\infty(]a, b[)$, $x_0 \in]a, b[$. Die formale Potenzreihe

$$p_f(x; x_0) = \sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x - x_0)^k}{k!}$$

heisst die **Taylor-Reihe** von *f* um x_0 .

Beispiel 6.5.8. i) Sei $f(x) = 1 + x^2$, $x_0 \in \mathbb{R}$. Zerlegen wir $x = x_0 + (x - x_0)$, so erhalten wir die Darstellung

$$\begin{aligned} f(x) &= 1 + (x_0 + (x - x_0))^2 = 1 + x_0^2 + 2x_0(x - x_0) + (x - x_0)^2 \\ &= f(x_0) + f'(x_0)(x - x_0) + f''(x_0) \frac{(x - x_0)^2}{2} = T_2 f(x; x_0) = p_f(x; x_0). \end{aligned}$$

Analog kann man zeigen, dass jedes Polynom durch seine Taylorreihe um einen beliebigen Punkt $x_0 \in \mathbb{R}$ dargestellt wird.

ii) Offenbar hat $f(x) = e^x$ um $x_0 = 0$ die Taylor-Reihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Für $x_0 \neq 0$ folgt mit dem Additionstheorem

$$\begin{aligned} f(x) &= \exp((x - x_0) + x_0) = \exp(x_0) \exp(x - x_0) \\ &= e^{x_0} \sum_{k=0}^{\infty} \frac{(x - x_0)^k}{k!} = \sum_{k=0}^{\infty} e^{x_0} \frac{(x - x_0)^k}{k!} = p_f(x; x_0), \end{aligned}$$

da $f^{(k)}(x_0) = e^{x_0}$ für jedes $k \in \mathbb{N}_0$.

Allgemein gilt:

Satz 6.5.6. Sei

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

mit Konvergenzradius

$$0 < \rho = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}} \leq \infty,$$

und sei $x_0 \in B_\rho(0)$. Dann ist die Taylor-Reihe $p_f(x; x_0)$ von f für $x \in B_r(x_0)$ mit $r = \rho - |x_0|$ absolut konvergent, und

$$\forall x \in B_r(x_0) : f(x) = p_f(x; x_0) \quad .$$

Beweis. Nach Satz 6.5.2 hat

$$p_f(x; x_0) = \sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x - x_0)^k}{k!} = \sum_{k=0}^{\infty} b_k (x - x_0)^k$$

die Koeffizienten $b_k \in \mathbb{R}$ mit

$$b_k = \frac{f^{(k)}(x_0)}{k!} = \sum_{\ell=k}^{\infty} a_\ell \frac{\ell!}{(\ell - k)! k!} x_0^{\ell - k}, \quad k \in \mathbb{N}_0.$$

Falls Umordnen erlaubt, so folgt nach Einsetzen:

$$\begin{aligned} p_f(x; x_0) &= \sum_{k=0}^{\infty} \underbrace{\sum_{\ell=k}^{\infty} a_\ell \frac{\ell!}{(\ell - k)! k!} x_0^{\ell - k}}_{=b_k} (x - x_0)^k \\ &= \sum_{\ell=0}^{\infty} a_\ell \underbrace{\sum_{k=0}^{\ell} \frac{\ell!}{(\ell - k)! k!} x_0^{\ell - k} (x - x_0)^k}_{=\binom{\ell}{k}} = f(x). \\ &= \underbrace{\sum_{k=0}^{\ell} \binom{\ell}{k} x_0^{\ell - k} (x - x_0)^k}_{=((x-x_0)+x_0)^\ell = x^\ell} \end{aligned}$$

Umordnen ist nach Satz 3.8.1 zulässig, falls $p_f(x; x_0)$ absolut konvergiert.

Betrachte dazu die Reihe

$$\tilde{f}(x) = \sum_{k=0}^{\infty} |a_k| x^k = \lim_{L \rightarrow \infty} \tilde{f}_L(x), \quad \tilde{f}_L(x) = \sum_{k=0}^L |a_k| x^k, \quad L \in \mathbb{N}.$$

Der Konvergenzradius von \tilde{f} ist

$$\tilde{\rho} = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|a_k|}} = \rho.$$

Die obige Rechnung zeigt die Gleichheit

$$T_L \tilde{f}_L(x; x_0) = \sum_{k=0}^L \sum_{\ell=k}^L |a_k| \frac{\ell!}{(\ell-k)!k!} x_0^{\ell-k} (x-x_0)^k = \tilde{f}_L(x)$$

für jedes L , analog zu Beispiel 6.5.8.i). Da $\tilde{f}(x)$ für $|x| < \rho$ konvergiert, erhalten wir für $0 \leq x_0 < x < \rho$ absolute Konvergenz der Reihe

$$p_{\tilde{f}}(x; x_0) = \lim_{L \rightarrow \infty} \sum_{k=0}^L \sum_{\ell=k}^L |a_k| \frac{\ell!}{(\ell-k)!k!} \underbrace{x_0^{\ell-k}}_{\geq 0} \underbrace{(x-x_0)^k}_{> 0} = \tilde{f}(x) < \infty.$$

Also konvergiert auch $p_f(x; x_0)$ absolut für $|x - x_0| < r = \rho - |x_0|$. \square

Beispiel 6.5.9. Sei

$$f(x) = \frac{1}{1-x}, \quad x < 1.$$

Für $x_0, x \in \mathbb{R}$ mit $x_0, x < 1$ schreibe $x = (x - x_0) + x_0$. Es folgt:

$$\begin{aligned} f(x) &= \frac{1}{1 - (x - x_0) - x_0} = \frac{1}{1 - x_0} \cdot \frac{1}{1 - \frac{x-x_0}{1-x_0}} = \frac{1}{1-x_0} \sum_{k=0}^{\infty} \left(\frac{x-x_0}{1-x_0} \right)^k \\ &= \sum_{k=0}^{\infty} \frac{k!}{(1-x_0)^{k+1}} \frac{(x-x_0)^k}{k!} = \sum_{k=0}^{\infty} f^{(k)}(x_0) \frac{(x-x_0)^k}{k!} = p_f(x; x_0), \end{aligned}$$

mit Konvergenzradius

$$\rho_{p_f(\cdot, x_0)} = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{\frac{1}{(1-x_0)^{k+1}}}} = 1 - x_0 \geq \rho - |x_0|.$$

Insbesondere für $x_0 = 0$ erhalten wir die bekannte Summenformel

$$f(x) = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \quad |x| < 1.$$

Definition 6.5.7. Eine glatte Funktion $f \in C^\infty(\Omega)$ heisst **analytisch**, falls f in einer Umgebung eines jeden Punktes $x_0 \in \Omega$ durch die Taylor-Reihe $p_f(\cdot, x_0)$ dargestellt wird.

Notation:

$$C^\omega(\Omega) = \{f \in C^\infty(\Omega); f \text{ ist analytisch}\}.$$

Bemerkung 6.5.3. Beispiel 6.5.9 zeigt, dass die durch die Potenzreihe $f(x) = \sum_{k=0}^{\infty} x^k$ zunächst nur auf $B_1(0)$ definierte Funktion durch $f(x) = \frac{1}{1-x}$ auf $\Omega =]-\infty, 1[$ analytisch fortgesetzt werden kann.

Nicht alle glatten Funktionen sind analytisch, wie das folgende Beispiel zeigt.

Beispiel 6.5.10. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} e^{-1/x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

f ist glatt auf $\mathbb{R} \setminus \{0\}$ mit

$$f^{(k)}(x) = \frac{p_k(x)}{x^{2k}} e^{-1/x}, \quad x > 0, k \in \mathbb{N}_0,$$

wobei p_k ein Polynom vom Grad $\leq k$ ist. Mit Beispiel 6.2.3.i) erhalten wir nach der Substitution $y = 1/x$ für alle $\ell \in \mathbb{N}_0$ die Identität

$$\lim_{x \downarrow 0} (x^{-\ell} e^{-1/x}) = \lim_{y \uparrow \infty} (y^\ell e^{-y}) = 0.$$

Damit folgt rekursiv für alle $k \in \mathbb{N}_0$, dass $f^{(k)}$ differenzierbar ist an der Stelle $x_0 = 0$ mit $f^{(k+1)}(0) = 0$; das heisst, $f \in C^\infty(\mathbb{R})$. Jedoch ist f nicht analytisch, da für $x_0 = 0 < x$ gilt

$$0 = p_f(x; 0) < f(x).$$

Schliesslich können wir mit den Ideen aus Satz 6.5.2 einen Beweis des Fundamentalsatzes der Algebra gemäss Abschnitt 2.6 gewinnen.

Satz 6.5.7. Jedes Polynom $p : \mathbb{C} \rightarrow \mathbb{C}$ vom Grad ≥ 1 hat (mindestens) eine Nullstelle.

Beweis(indirekt). Sei $p(z) = a_n z^n + \dots + a_0$ mit $a_n \neq 0$ ein Polynom vom Grad $n \geq 1$ ohne Nullstelle. OBdA sei $a_n = 1$. (Sonst betrachte $\tilde{p} = \frac{p}{a_n}$.) Sei

$$\mu = \inf_{z \in \mathbb{C}} |p(z)| \geq 0.$$

Da wir für genügend grosses $r_0 > 0$ für $|z| \geq r_0$ abschätzen können

$$|p(z)| = |z|^n (1 + a_{n-1} z^{-1} + \dots + a_0 z^{-n}) \geq \frac{1}{2} |z|^n,$$

können wir einen Radius $r_0 > 0$ wählen mit

$$|p(z)| \geq 1 + \mu$$

für $|z| \geq r_0$. Nach Korollar 5.3.1 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 Beispiel 6.5.8.i). 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

$$k_0 = \min\{k \in \{1, \dots, n\}; b_k \neq 0\}.$$

Da $p \neq \text{const}$, 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

$$r_{k_0}(z; z_0) = \sum_{k=k_0+1}^n b_k (z - z_0)^k.$$

Beachte, dass

$$r_{k_0}(z; z_0)/(z - z_0)^{k_0} \rightarrow 0 \quad (z \rightarrow z_0).$$

Schliesslich seien $s_0 > 0$, $0 \leq \phi_0 < 2\pi$ mit

$$-\frac{p(z_0)}{b_{k_0}} = s_0 e^{i\phi_0}.$$

Für $z = z_0 + s e^{i\phi_0/k_0}$, $s > 0$ folgt

$$|p(z)| = |p(z_0) + b_{k_0} s^{k_0} e^{i\phi_0} + r_{k_0}(z; z_0)| = |p(z_0)| \left| 1 - s^{k_0} \left(\frac{1}{s_0} - \frac{r_{k_0}(z; z_0)}{p(z_0) s^{k_0}} \right) \right|.$$

Da $r_{k_0}(z; z_0)/s^{k_0} \rightarrow 0$ ($s \rightarrow 0$), folgt der gewünschte Widerspruch

$$|p(z)| < |p(z_0)| = \mu, \quad \text{falls } 0 < s \ll 1.$$

□

6.6 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 6.6.1. i) Die Funktionen \exp , \sin , \cos , \tan stehen mit ihrer Ableitung in Beziehung:

$$\exp' = \exp, \quad \sin'' = -\sin, \quad \cos'' = -\cos, \quad \tan' = 1 + \tan^2.$$

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. Beispielsweise ist beim *radioaktiven Zerfall* die pro Zeiteinheit zerfallende Masse proportional zur noch vorhandenen Masse $f(t)$ eines Stoffes; das heisst, mit einer Zahl $\alpha > 0$ gilt

$$\dot{f} = \frac{df}{dt} = -\alpha f, \quad f(0) = f_0. \quad (6.6.1)$$

Gemäss Beispiel 6.2.1.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 ungedämpften *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; das heisst, wir haben die Gleichung

$$\ddot{f} + \omega_0^2 f = 0, \quad \omega_0^2 = \frac{K}{m} > 0. \quad (6.6.2)$$

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 (6.6.2). Sind alle Lösungen von (6.6.2) von dieser Form?

iv) Beim mehrstufigen radioaktiven Zerfall einer Substanz s_1 in die stabile Substanz s_n über Zwischenstufen s_2, \dots, s_{n-1} mit Massen $f_i(t)$ und Zerfallsraten $\alpha_i > 0$ erhalten wir das **System** von Differentialgleichungen

$$\begin{aligned} \dot{f}_1 &= -\alpha_1 f_1, \\ \dot{f}_2 &= \alpha_1 f_1 - \alpha_2 f_2, \\ &\vdots \\ \dot{f}_n &= \alpha_{n-1} f_{n-1}. \end{aligned} \quad (6.6.3)$$

Mit der Notation

$$F = F(t) = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} : \mathbb{R} \rightarrow \mathbb{R}^n, \quad A = \begin{pmatrix} -\alpha_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & -\alpha_{n-1} \\ & & & & \alpha_{n-1} & \\ & & & & & 0 \end{pmatrix}$$

können wir (6.6.3) in der Form schreiben

$$\dot{F} = AF. \quad (6.6.4)$$

analog zu (6.6.1).

v) Führen wir im Fall (6.6.2) die Funktion

$$F = \begin{pmatrix} f \\ \dot{f} \end{pmatrix} : \mathbb{R} \rightarrow \mathbb{R}^2$$

ein, so lässt sich auch diese Gleichung in der Form (6.6.4) 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 6.6.1. Die Gleichung (6.6.4) ist die Standardform eines **homogenen Systems linearer Differentialgleichungen mit konstanten Koeffizienten**.

Bemerkung 6.6.1. Mit F , F_1 , F_2 sind für jedes $a, a_1, a_2 \in \mathbb{R}$ offenbar auch die Funktionen aF und $a_1F_1 + a_2F_2$ Lösungen von (6.6.4) (*“Superpositionsprinzip”*).

Natürlich können wir die Gleichung (6.6.4) auch mit einer komplexen Matrix A und einer Funktion $F: \mathbb{R} \rightarrow \mathbb{C}^n$ betrachten.

Satz 6.6.1. (Existenz- und Eindeutigkeitsatz für (6.6.4)) 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**

$$\frac{dF}{dt} = AF, \quad F(0) = F_0 \tag{6.6.5}$$

genau eine Lösung $F \in C^1(\mathbb{R}; \mathbb{R}^n)$ (bzw. $C^1(\mathbb{R}; \mathbb{C}^n)$). Die Komponenten von F sind analytische Funktionen auf \mathbb{R} ; insbesondere $F \in C^\infty(\mathbb{R}; \mathbb{R}^n)$ (bzw. $F \in C^\infty(\mathbb{R}; \mathbb{C}^n)$).

Beispiel 6.6.2. Insbesondere ist die Funktion $F(t) = \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}$ die einzige Lösung des Anfangswertproblems

$$\dot{F} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} F, \quad F(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{6.6.6}$$

Somit ist auch $f(t) = \cos t$ die einzige Lösung des Anfangswertproblems

$$\ddot{f} + f = 0, \quad f(0) = 1, \quad \dot{f}(0) = 0, \tag{6.6.7}$$

denn jede Lösung f von (6.6.7) induziert eine Lösung $F = \begin{pmatrix} f \\ \dot{f} \end{pmatrix}$ von (6.6.6).

Beweis von Satz 6.6.1. Nach Identifikation $\mathbb{C} \ni z = x + iy \mapsto (x, y) \in \mathbb{R}^2$ genügt es, den Fall reeller Koeffizienten zu behandeln.

i) (Eindeutigkeit.) Es seien $F, G \in C^1(\mathbb{R}; \mathbb{R}^n)$ Lösungen von (6.6.5). 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 2.5.1 folgt

$$\frac{d\eta}{dt} = 2 \left\langle H, \frac{dH}{dt} \right\rangle = 2 \langle H, AH \rangle \leq 2 |H| \cdot |AH| \leq C_1 |H|^2 = C_1 \eta.$$

Somit gilt

$$\frac{d}{dt} (e^{-C_1 t} \eta(t)) = \left(\frac{d\eta}{dt} - C_1 \eta \right) e^{-C_1 t} \leq 0,$$

und Korollar 6.2.1.ii) ergibt

$$\forall t \geq 0 : e^{-C_1 t} \eta(t) \leq \eta(0) = |H(0)|^2 = 0.$$

Das heisst, $\eta(t) = 0$ und somit auch $H(t) = 0$ für alle $t \geq 0$; analog für $t \leq 0$.

ii) (Existenz.) Analog zu Beispiel 6.6.1.ii) machen wir für die Lösung von (6.6.5) den Ansatz

$$F(t) = \text{Exp}(At)F_0, \quad t \in \mathbb{R}, \quad (6.6.8)$$

wobei die Reihe (mit $A^0 = id$)

$$\text{Exp}(At) = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} =: e^{At}$$

in jeder Matrix-Norm analog zu Beispiel 3.7.2.i) für beliebige $t \in \mathbb{R}$ konvergiert. Weiter gilt analog zu Beispiel 6.4.3.i), dass $\text{Exp}(At) \in C^1(\mathbb{R}; \mathbb{R}^{n \times n})$ mit

$$\frac{d}{dt} \text{Exp}(At) = \sum_{k=0}^{\infty} \frac{d}{dt} \left(\frac{A^k t^k}{k!} \right) = A \text{Exp}(At);$$

das heisst,

$$\frac{dF}{dt} = AF, \quad F(0) = \underbrace{\text{Exp}(A \cdot 0)}_{=id} F_0 = F_0,$$

wie gewünscht. Schliesslich sind nach Satz 6.5.6 die Komponenten von e^{At} , also auch die von F analytisch. \square

Bemerkung 6.6.2. Für invertierbares A ist die Definition $A^0 = id$ offenbar sinnvoll. Eine beliebige quadratische Matrix A kann man durch invertierbare Matrizen approximieren, was die Vereinbarung $A^0 = id$ für beliebige quadratische Matrizen rechtfertigt.

Definition 6.6.2. Die Matrix-wertige Funktion

$$t \mapsto \Phi(t) = \text{Exp}(At) \quad \text{mit} \quad \frac{d\Phi}{dt} = A\Phi, \quad \Phi(0) = id$$

heisst **Fundamentallösung** von (6.6.4).

Die Darstellung (6.6.8) der allgemeinen Lösung von (6.6.4) erlaubt die folgende präzise Charakterisierung.

Satz 6.6.2. i) Der Lösungsraum

$$X = \{F \in C^1(\mathbb{R}; \mathbb{R}^n); \frac{dF}{dt} = AF\}$$

von (6.6.4) 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 (6.6.4) ein n -dimensionaler \mathbb{C} -Vektorraum.

Beweis. Gemäss Satz 6.6.1 ist jede Lösung von (6.6.4) von der Gestalt (6.6.8). Da $\Phi(0) = id$, ist die lineare Abbildung $L: \mathbb{R}^n \ni F_0 \mapsto \Phi(t)F_0 \in C^1(\mathbb{R}; \mathbb{R}^n)$ injektiv, ihr Bild $L(\mathbb{R}^n) = X$ also wiederum ein Vektorraum der Dimension n . Analog in \mathbb{C} . \square

Bemerkung 6.6.3. i) Ohne Vorbereitung kann man $\Phi(t) = \text{Exp}(At)$ nur mit Mühe berechnen. Falls man jedoch durch eine lineare Transformation $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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

$$T \cdot \text{Exp}(At)T^{-1} = \text{Exp}(TAT^{-1}t) = \text{Exp}(\Lambda t) = \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix},$$

und wir erhalten

$$\Phi(t) = T^{-1} \begin{pmatrix} e^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n t} \end{pmatrix} T.$$

ii) Allgemein kann man über dem Skalkörper \mathbb{C} jede Matrix A durch eine Ähnlichkeitstransformation $T: \mathbb{C}^n \rightarrow \mathbb{C}^n$ auf *Jordan-Normalform* bringen; das heisst,

$$TAT^{-1} = \begin{pmatrix} \Lambda_1 & 0 & \dots & 0 \\ 0 & \Lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Lambda_k \end{pmatrix} = \Lambda$$

mit Jordan-Blöcken

$$\Lambda_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda_i \end{pmatrix}$$

der Länge m_i , $1 \leq i \leq \ell$, und

$$e^{\Lambda t} = \begin{pmatrix} e^{\Lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & e^{\Lambda_k t} \end{pmatrix},$$

wobei für $1 \leq i \leq \ell$ gilt:

$$e^{\Lambda_i t} = e^{\lambda_i t} \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \frac{t^{m_i-1}}{(m_i-1)!} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & t \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix};$$

siehe die Vorlesung *Lineare Algebra*.

iii) Falls insbesondere (6.6.4) die äquivalente Form ist für eine skalare Differentialgleichung n -ter Ordnung

$$f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0f = 0 \quad (6.6.9)$$

mit

$$F = \begin{pmatrix} f \\ \dot{f} \\ \vdots \\ f^{(n-1)} \end{pmatrix} \in C^1(\mathbb{R}, \mathbb{R}^n), \quad A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & & 0 & 1 & 0 \\ -\alpha_0 & -\alpha_1 & \dots & & -\alpha_{n-1} \end{pmatrix},$$

so erwarten wir, dass die allgemeine komplexe Lösung f von (6.6.9) dargestellt werden kann als Linearkombination von Funktionen der Form $e^{\lambda_i t}$ oder $t^k e^{\lambda_i t}$, wobei $k < m_i$, die Vielfachheit des Eigenwerts λ_i von A . In der Tat führt der

Exponentialansatz $f(t) = e^{\lambda t}$ für eine Lösung von (6.6.8) auch direkt auf die Gleichung

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0 = 0,$$

wobei $p(\lambda) = \det(\lambda \text{id} - A)$ das charakteristische Polynom ist von A . Wir nennen p daher auch das *charakteristische Polynom* von (6.6.9).

Übertragen auf die Gleichung (6.6.9) ergibt Satz 6.6.2 das folgende Ergebnis.

Korollar 6.6.1. Seien $a_0, \dots, a_{n-1} \in \mathbb{R}$ (oder \mathbb{C}). Der Lösungsraum

$$Z = \{f \in C^n(\mathbb{R}; \mathbb{R}^n); f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0f = 0\}$$

der Gleichung (6.6.9) ist ein n -dimensionaler \mathbb{R} -Vektorraum (analog in \mathbb{C}).

Beweis. Sei (6.6.4) das zu (6.6.9) äquivalente System erster Ordnung, X der zugehörige Lösungsraum nach Satz 6.6.2. Die Abbildung

$$Z \ni f \mapsto F = (f, f', \dots, f^{(n-1)})^t \in X$$

ist offenbar ein linearer Isomorphismus. □

Beispiel 6.6.3. i) Betrachte die Gleichung

$$f^{(4)} - 3f^{(2)} + 2f = 0. \tag{6.6.10}$$

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 (6.6.10) sich in der Form darstellen

$$f(t) = ae^t + be^{-t} + ce^{\sqrt{2}t} + de^{-\sqrt{2}t}.$$

Durch Vorgabe von $F(0) = (f(0), \dots, f^{(3)}(0))^t$ ist f gemäss Satz 6.6.2 eindeutig bestimmt, und man kann die Konstanten a, \dots, d aus den vorgegebenen Werten für $f(0), \dots, f^{(3)}(0)$ bestimmen.

ii) Die Differentialgleichung (6.6.2) 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 (6.6.2) reell sind, sind mit $f \in C^1(\mathbb{R}; \mathbb{C})$ auch die Funktionen $\text{Re}(f), \text{Im}(f) \in C^1(\mathbb{R}; \mathbb{R})$ Lösungen; das heisst, die allgemeine reelle Lösung $f \in C^1(\mathbb{R}; \mathbb{R})$ hat die Gestalt

$$f(t) = a \cos(\omega_0 t) + b \sin(\omega_0 t), \quad t \in \mathbb{R}, \tag{6.6.11}$$

wobei $a, b \in \mathbb{R}$. Der \mathbb{R} -Lösungsraum von (6.6.2) 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 reelle Lösung von (6.6.2) von der Form (6.6.11).

Wir betrachten nun den Fall, dass das charakteristische Polynom $p(\lambda)$ der Gleichung (6.6.9) mit Koeffizienten $a_0, \dots, a_{n-1} \in \mathbb{R}$ oder \mathbb{C} mehrfache Nullstellen besitzt. Gemäss Satz 6.5.7 können wir p durch sukzessives Abspalten von Linearfaktoren darstellen als Produkt

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0 = \prod_{i=1}^l (\lambda - \lambda_i)^{m_i},$$

wobei $\lambda_1, \dots, \lambda_l \in \mathbb{C}$ die paarweise verschiedenen Nullstellen von p bezeichnen und $m_1, \dots, m_l \in \mathbb{N}$ deren Vielfachheit.

Satz 6.6.3. 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 (6.6.9) darstellbar als Linearkombination der n linear unabhängigen Funktionen

$$f_{ik}(t) = t^k e^{\lambda_i t}, \quad 1 \leq i \leq l, \quad 0 \leq k < m_i.$$

Beweis. Mit $D = \frac{d}{dt}$ können wir (6.6.9) in der Form schreiben

$$f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_0 f = p(D)f = 0,$$

wobei

$$p(D) = D^n + a_{n-1}D^{n-1} + \dots + a_0 \text{id} = \prod_{i=1}^l (D - \lambda_i \text{id})^{m_i}.$$

Hier fassen wir $D = \frac{d}{dt}$ auf als lineare Abbildung

$$D: C^\infty(\mathbb{R}, \mathbb{C}) \ni f \mapsto f' \in C^\infty(\mathbb{R}, \mathbb{C});$$

analog $\text{id} = D^0: f \mapsto f$. Für $q \in C^1(\mathbb{R})$ und beliebiges $\lambda \in \mathbb{C}$ gilt weiter:

$$(D - \lambda \text{id})(q(t)e^{\lambda t}) = (q'(t)e^{\lambda t} + \lambda q(t)e^{\lambda t}) - \lambda q(t)e^{\lambda t} = q'(t)e^{\lambda t}.$$

Nach Iteration folgt für $q \in C^k(\mathbb{R})$:

$$(D - \lambda \text{id})^k(q(t)e^{\lambda t}) = q^{(k)}(t)e^{\lambda t}, \quad k \in \mathbb{N}.$$

Für $i \in \{1, \dots, l\}$, $k < m_i$ erhalten wir damit

$$(D - \lambda_i \text{id})^{m_i}(t^k e^{\lambda_i t}) = \frac{d^{m_i} t^k}{dt^{m_i}} e^{\lambda_i t} = 0.$$

Beachte, dass für $\lambda, \mu \in \mathbb{C}$, $f \in C^2(\mathbb{R})$ stets gilt

$$(D - \lambda \text{id})(D - \mu \text{id})f = f'' - (\lambda + \mu)f' + \lambda\mu f = (D - \mu \text{id})(D - \lambda \text{id})f.$$

Es folgt

$$p(D)f_{ik} = \prod_{j \neq i} (D - \lambda_j \text{id})^{m_j} \left((D - \lambda_i \text{id})^{m_i} f_{ik} \right) = 0, \quad 1 \leq i \leq \ell, \quad 0 \leq k < m_i.$$

Also löst jede der Funktionen f_{ik} die Gleichung (6.6.9).

Behauptung. Die Funktionen $(f_{ik})_{1 \leq i \leq \ell, 0 \leq k < m_i}$ sind linear unabhängig.

Beweis (indirekt). Seien $b_{ik} \in \mathbb{C}$, $1 \leq i \leq \ell$, $0 \leq k < m_i$, mit

$$\forall t \in \mathbb{R} : \sum_{i=1}^{\ell} \sum_{k=0}^{m_i-1} b_{ik} f_{ik}(t) = 0,$$

wobei $b_{i_0 k_0} \neq 0$ für ein $i_0 \in \{1, \dots, \ell\}$, $k_0 < m_{i_0}$. OBdA gelte

$$k_0 = \max\{k < m_{i_0}; b_{i_0 k} \neq 0\}.$$

Gemäss Teil i) gelten die Gleichungen

$$\prod_{j \neq i_0} (D - \lambda_j \text{id})^{m_j} \sum_{i \neq i_0} \sum_{k=0}^{m_i-1} b_{ik} f_{ik} = 0, \quad (D - \lambda_{i_0} \text{id})^{k_0} \sum_{k < k_0} f_{i_0 k} = 0,$$

sowie

$$(D - \lambda_{i_0} \text{id})^{k_0} f_{i_0 k_0} = k_0! e^{\lambda_{i_0} t}.$$

Mit $\sum_{i,k} b_{ik} f_{ik} = 0$ erhalten wir dann

$$\begin{aligned} 0 &= \prod_{j \neq i_0} (D - \lambda_j \text{id})^{m_j} (D - \lambda_{i_0} \text{id})^{k_0} \left(\sum_{i,k} b_{ik} f_{ik}(t) \right) \\ &= \prod_{j \neq i_0} (D - \lambda_j \text{id})^{m_j} k_0! b_{i_0 k_0} e^{\lambda_{i_0} t} = k_0! b_{i_0 k_0} \prod_{j \neq i_0} (\lambda_{i_0} - \lambda_j)^{m_j} e^{\lambda_{i_0} t}, \end{aligned}$$

und $b_{i_0 k_0} = 0$ im Widerspruch zu unserer Annahme. \square

Da $\sum_{1 \leq i \leq \ell} m_i = n$, folgt mit Korollar 6.6.1 die Behauptung. \square

Beispiel 6.6.4. i) Die Gleichung

$$\ddot{f} - 2\dot{f} + f = 0 \tag{6.6.12}$$

hat das charakteristische Polynom

$$p(\lambda) = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2$$

mit der doppelten Nullstelle $\lambda_1 = 1$. Jede Lösung f der Differentialgleichung (6.6.12) ist also von der Form

$$f(t) = (a + bt)e^t, \quad t \in \mathbb{R},$$

wobei $a, b \in \mathbb{R}$.

ii) Ein Federpendel mit Dämpfung wird beschrieben durch die Differentialgleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 0 \quad (6.6.13)$$

mit Dämpfungskonstante $\delta > 0$ und $\omega_0 > 0$. Das charakteristische Polynom

$$p(\lambda) = \lambda^2 + 2\delta\lambda + \omega_0^2$$

hat die (im allgemeinen komplexen) Nullstellen

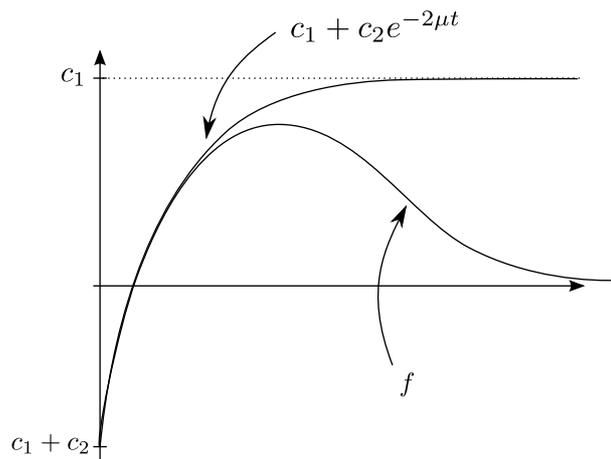
$$\lambda_{1,2} = -\delta \pm \sqrt{\delta^2 - \omega_0^2} = -\delta \pm i\sqrt{\omega_0^2 - \delta^2}.$$

Wir unterscheiden folgende Fälle:

a) Sei $\delta^2 > \omega_0^2$ ("superkritische Dämpfung"). Die allgemeine Lösung von (6.6.13) 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) Sei $\delta^2 = \omega_0^2$ ("kritische Dämpfung"). Analog zu Beispiel i) ist die allgemeine Lösung in diesem Fall von der Form

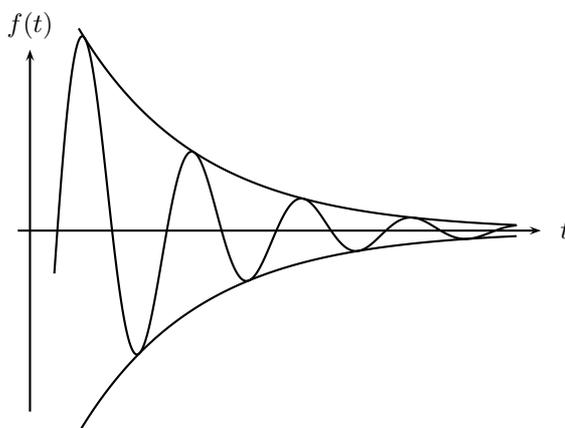
$$f(t) = c_1 e^{-\delta t} + c_2 t e^{-\delta t} = (c_1 + c_2 t) e^{-\delta t}.$$

Die Lösungen in den Fällen a) und b) sind also stets exponentiell abfallend mit höchstens einer Nullstelle (Stossdämpfertest).

c) Sei $\delta^2 < \omega_0^2$ ("subkritische 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 (6.6.13) in \mathbb{C} lautet

$$f(t) = e^{-\delta t} (c_1 e^{i\mu t} + c_2 e^{-i\mu t}),$$

wobei $c_{1,2} \in \mathbb{C}$. Die allgemeine reelle Lösung von (6.6.13) erhalten wir, indem wir hiervon den Realteil nehmen. Mit $c_1 + \overline{c_2} =: a + ib$, wobei $a, b \in \mathbb{R}$, ergibt



dies die Darstellung

$$\begin{aligned} f(t) &= \mathcal{R}e(e^{-\delta t}(c_1 e^{i\mu t} + c_2 e^{-i\mu t})) = \frac{1}{2}e^{-\delta t}((c_1 + \overline{c_2})e^{i\mu t} + (\overline{c_1} + c_2)e^{-i\mu t}) \\ &= \frac{1}{2}e^{-\delta t}((a + ib)e^{i\mu t} + (a - ib)e^{-i\mu t}) = e^{-\delta t}(a \cos(\mu t) - b \sin(\mu t)); \end{aligned}$$

die allgemeine reelle Lösung beschreibt also eine gedämpfte Schwingung.

6.7 Inhomogene Differentialgleichungen

Bisher haben wir nur homogene lineare Differentialgleichungen betrachtet. Sehr oft treten jedoch auch Zusatzterme in den Gleichungen auf.

Beispiel 6.7.1. 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

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = \beta_0 \cos(\omega t). \quad (6.7.1)$$

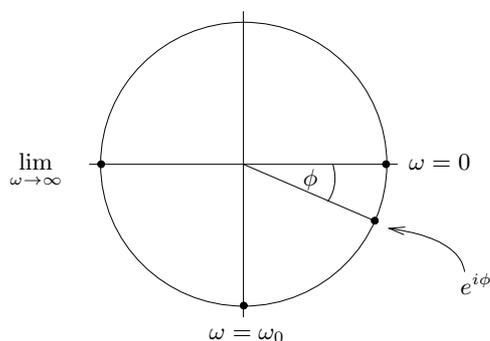
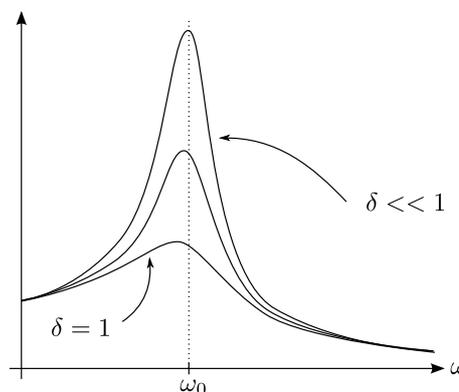
Als *partikuläre Lösung* dieser Gleichung erwarten wir eine Schwingung mit derselben Frequenz ω . Am leichtesten gelingt die Rechnung im Komplexen.

Für eine Lösung $f \in C^2(\mathbb{R}; \mathbb{C})$ der Gleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = \beta_0 e^{i\omega t} \quad (6.7.2)$$

machen wir den *Ansatz*

$$f(t) = ce^{i\omega t}, \quad t \in \mathbb{R},$$

Abbildung 6.1: Phasenwinkel ϕ Abbildung 6.2: Resonanzamplitude R

mit einer beliebigen Konstanten $c \in \mathbb{C}$. Einsetzen in (6.7.2) ergibt

$$\left((\omega_0^2 - \omega^2) + 2i\delta\omega \right) c - \beta_0 e^{i\omega t} = 0$$

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

$$\frac{1}{\omega_0^2 - \omega^2 + 2i\delta\omega} = \frac{\omega_0^2 - \omega^2 - 2i\delta\omega}{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2} = Re^{i\phi},$$

so können wir

$$R = \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\delta^2\omega^2}}$$

als "Resonanzamplitude" und

$$\phi = \arctan\left(\frac{-2\delta\omega}{\omega_0^2 - \omega^2}\right) \in]-\pi, 0].$$

als "Phasenverschiebung" gegenüber der von aussen wirkenden Kraft deuten. Schliesslich liefert

$$f_{part}(t) = \beta_0 R e^{i(\omega t + \phi)}, \quad t \in \mathbb{R},$$

die gesuchte partikuläre Lösung von (6.7.2), bzw.

$$\tilde{f}_{part}(t) = Re(f_{part}(t)) = \beta_0 R \cos(\omega t + \phi), \quad t \in \mathbb{R}$$

die gesuchte partikuläre Lösung von (6.7.1).

Beachte: Für $\delta \sim 0$ tritt bei $\omega = \omega_0$ ein unkontrolliertes Aufschaukeln der anregenden Schwingung ein (*Resonanzkatastrophe*).

ii) Eine partikuläre Lösung der Gleichung

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 1 \quad (6.7.3)$$

kann man ebenfalls leicht erraten. Die “Kraft” der Grösse 1 auf der rechten Seite von (6.7.3) führt zu einer Verschiebung der Ruhelage des Pendels um $\frac{1}{\omega_0^2}$; neu entspricht also 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

$$\ddot{f} + 2\delta\dot{f} + \omega_0^2 f = 1 + \beta_0 e^{i\omega t}. \quad (6.7.4)$$

Nach dem *Superpositionsprinzip* ergänzen sich die oben bestimmten partikulären Lösungen von (6.7.2) und (6.7.3) zu einer partikulären Lösung

$$f(t) = ce^{i\omega t} - \frac{1}{\omega_0^2}, \quad t \in \mathbb{R},$$

von (6.7.4), 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

$$\frac{dF}{dt} = AF + B, \quad (6.7.5)$$

wobei $B = B(t) \in C^0(\mathbb{R}; \mathbb{R}^n)$ (oder $\in C^0(\mathbb{R}; \mathbb{C}^n)$).

Satz 6.7.1. Sei $F_{part} \in C^1(\mathbb{R}; \mathbb{R}^n)$ eine beliebige (“partikuläre”) Lösung von (6.7.5). Dann ist jede Lösung F von (6.7.5) von der Form

$$F = F_{part} + F_{hom}, \quad (6.7.6)$$

wobei F_{hom} eine beliebige Lösung der homogenen Gleichung (6.6.4) ist. Insbesondere gibt es zu jedem $F_0 \in \mathbb{R}^n$ stets genau eine Lösung F von (6.7.5) mit $F(0) = F_0$. (Analog in \mathbb{C} .)

Beweis. i) Jedes F der Form (6.7.6) löst (6.7.5). Sind umgekehrt $F_1, F_2 \in C^1(\mathbb{R}; \mathbb{R}^n)$ Lösungen von (6.7.5), so gilt

$$\frac{d(F_1 - F_2)}{dt} = A(F_1 - F_2) + B - B = A(F_1 - F_2);$$

das heisst, jede Lösung von (6.7.5) ist von der Gestalt (6.7.6).

ii) Zu vorgegebenen Anfangswerten $F_0 \in \mathbb{R}^n$ sei F_{hom} die Lösung des Anfangswertproblems (6.6.5) mit $F_{hom}(0) = F_0 - F_{part}(0)$. Dann löst $F = F_{part} + F_{hom}$ das Anfangswertproblem (6.7.5) mit $F(0) = F_0$. Eindeutigkeit der Lösung F folgt mit i) und Satz 6.6.1. \square

Beispiel 6.7.2. i) Radioaktiver Zerfall mit konstanter Zufuhr wird beschrieben durch die Modellgleichung

$$\dot{f} = -\alpha f + s, \quad (6.7.7)$$

wobei $\alpha, s > 0$ gegeben sind. Eine partikuläre Lösung ist offenbar die *Gleichgewichtslösung*

$$f_{part}(t) = \frac{s}{\alpha}, \quad t \in \mathbb{R}.$$

Mit der Darstellung $f_{hom}(t) = ce^{-\alpha t}$ der allgemeinen Lösung der homogenen Gleichung aus Beispiel 6.6.1.ii) erhalten wir die allgemeine Lösung

$$f(t) = ce^{-\alpha t} + \frac{s}{\alpha}$$

von (6.7.7), 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

$$y(t) = \frac{1}{\frac{b}{a} + ce^{-at}} \rightarrow \frac{a}{b} \quad (t \rightarrow \infty).$$

Ein analoges Resultat gilt für allgemeine Differentialgleichungen vom "Bernoulli-Typ"

$$y' = ay - by^\alpha, \quad \alpha > 1.$$

Wie findet man im allgemeinen Fall eine partikuläre Lösung zu (6.7.5)? 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 7

Integration

7.1 Stammfunktionen

Seien $-\infty < a < b < \infty$, $f \in C^0(]a, b[)$.

Definition 7.1.1. Ein $F \in C^1(]a, b[)$ heisst **Stammfunktion** zu f , falls gilt

$$F' = \frac{dF}{dx} = f \quad \text{in }]a, b[.$$

Beispiel 7.1.1. i) Wegen $\log'(x) = \frac{1}{x}$, $x > 0$, ist $F(x) = \log(x)$ Stammfunktion zu $f(x) = \frac{1}{x}$, $x > 0$; ebenso $\tilde{F}(x) = \log(x) + 1$, 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 7.1.1. 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 6.2.1.i). \square

Sei $f \in C^0(]a, b[)$ und $F \in C^1(]a, b[)$ eine Stammfunktion zu f .

Definition 7.1.2. 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 7.1.1. i) Wegen Satz 7.1.1 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 in Kapitel 6 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 aus Kapitel 6 unmittelbar die folgenden Regeln.

Satz 7.1.2. 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 $\alpha f + \beta g$; das heisst,

$$\int (\alpha f + \beta g) dx = \alpha \int f dx + \beta \int g dx.$$

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. i) Summen- und Produktregel gemäss Satz 6.1.2 ergeben

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g.$$

ii) Analog folgt mit $F' = uv'$ sofort

$$(uv - F)' = u'v + uv' - F' = u'v.$$

□

Beispiel 7.1.2. i) Sei $p(x) = a_n x^n + \dots + a_1 x + a_0$. Mit $\int x^k dx = \frac{x^{k+1}}{k+1}$, $k \in \mathbb{N}_0$, und Satz 7.1.2.i) erhalten wir

$$\int p dx = \frac{a_n}{n+1} x^{n+1} + \dots + a_0 x.$$

ii) Seien $u(x) = x$, $v(x) = \log(x)$ mit $uv'(x) = x \cdot \frac{1}{x} = 1$, $x > 0$. Mit Satz 7.1.2.ii) folgt

$$\begin{aligned} \int \log(x) dx &= \int u'v dx = uv - \int uv' dx \\ &= x \log(x) - \int 1 dx = x \log(x) - x. \end{aligned}$$

iii) Mit $e^{-x} = -\frac{d}{dx}(e^{-x})$ erhalten wir nach k -facher partieller Integration

$$\begin{aligned} \int x^k e^{-x} dx &= -x^k e^{-x} + k \int x^{k-1} e^{-x} dx \\ &= -x^k e^{-x} - kx^{k-1} e^{-x} + k(k-1) \int x^{k-2} e^{-x} dx \\ &= \dots = -\sum_{l=1}^k \frac{k!}{l!} x^l e^{-x} + k! \underbrace{\int e^{-x} dx}_{-e^{-x}} = -\sum_{l=0}^k \frac{k!}{l!} x^l e^{-x}. \end{aligned}$$

Mit Beispiel 6.2.3.i) folgt die Identität

$$\Gamma(k+1) := \lim_{x \rightarrow \infty} \int_0^x x^k e^{-x} dx = k!, \quad k \in \mathbb{N}.$$

Satz 7.1.3. (Monotonie) Seien $f, g \in C^0([a, b])$ mit Stammfunktionen $F, G \in C^1([a, b])$, und sei $f \leq g$. Dann gilt für $a < x_0 < x_1 < b$:

$$\int_{x_0}^{x_1} f dx \leq \int_{x_0}^{x_1} g dx.$$

Beweis. OBdA $f = 0$. (Betrachte $\tilde{g} = g - f \geq 0$ und benutze Satz 7.1.2.i). Mit Korollar 6.2.1.ii) folgt aus $G' = \frac{dG}{dx} = g \geq 0$ und Definition 7.1.2

$$\int_{x_0}^{x_1} g dx = G(x_1) - G(x_0) \geq 0 = \int_{x_0}^{x_1} f dx.$$

□

Beispiel 7.1.3. i) Da für $0 \leq x \leq \pi/2$ und beliebiges $k \in \mathbb{N}_0$ gilt

$$\sin^k(x) - \sin^{k+1}(x) = \underbrace{\sin^k(x)}_{\geq 0} \underbrace{(1 - \sin(x))}_{\geq 0} \geq 0,$$

erhalten wir mit Satz 7.1.3 die Ungleichung

$$\forall k \in \mathbb{N}_0 : \int_0^{\pi/2} \sin^{k+1}(x) dx \leq \int_0^{\pi/2} \sin^k(x) dx.$$

ii) Mit partieller Integration gilt für jedes $k \in \mathbb{N}$ andererseits

$$\begin{aligned} \int_0^{\pi/2} \sin^{k+1}(x) dx &= \int_0^{\pi/2} \underbrace{\sin(x)}_{=\frac{d}{dx}(-\cos(x))} \sin^k(x) dx \\ &= -\cos(x) \sin^k(x) \Big|_{x=0}^{\pi/2} + k \int_0^{\pi/2} \underbrace{\cos^2(x)}_{=1-\sin^2(x)} \sin^{k-1}(x) dx \\ &= k \int_0^{\pi/2} \sin^{k-1}(x) dx - k \int_0^{\pi/2} \sin^{k+1}(x) dx; \end{aligned}$$

das heisst,

$$\forall k \in \mathbb{N} : \int_0^{\pi/2} \sin^{k+1}(x) dx = \frac{k}{k+1} \int_0^{\pi/2} \sin^{k-1}(x) dx.$$

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

$$\frac{(2^n n!)^2}{(2n+1)!} \leq \frac{(2n)!}{(2^n n!)^2} \cdot \frac{\pi}{2} \leq \frac{(2^{n-1}(n-1)!)^2}{(2n-1)!} = \frac{1}{2n} \frac{(2^n n!)^2}{(2n)!}$$

oder

$$\frac{(2^n n!)^4}{((2n)!)^2} \frac{2}{2n+1} \leq \pi \leq \frac{2}{2n} \frac{(2^n n!)^4}{((2n)!)^2}$$

und somit

$$\pi = \lim_{n \rightarrow \infty} \frac{1}{n} \left(\frac{2^{2n} (n!)^2}{(2n)!} \right)^2 \quad (\text{Wallissches Produkt}).$$

Einige Werte des Näherungsausdrucks:

$$\frac{1}{n} \left(\frac{2^{2n} (n!)^2}{(2n)!} \right)^2 = \begin{cases} 4, & n = 1, \\ 3.221 \dots, & n = 10, \\ 3.157 \dots, & n = 50, \\ 3.149 \dots, & n = 100. \end{cases}$$

Satz 7.1.4. (Gebietsadditivität des Integrals) Sei $f \in C^0([a, b])$ mit Stammfunktion $F \in C^1([a, b])$, und seien $a < x_0 \leq x_1 \leq x_2 < b$. 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 7.1.2 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. \square

Beispiel 7.1.4. (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 7.1.2.ii) Stammfunktion ist von $f(x) = \log(x)$, erhalten wir zudem für jedes $k \in \mathbb{N}$ die Identität

$$\int_{k-1/2}^{k+1/2} \log(x) dx = (x \log(x) - x) \Big|_{x=k-1/2}^{k+1/2}.$$

Entwickeln wir gemäss Bemerkung 6.5.1, so folgt

$$\begin{aligned} (k \pm \frac{1}{2}) \log(k \pm \frac{1}{2}) - (k \pm \frac{1}{2}) &= \\ &= (k \pm \frac{1}{2}) (\log(k) \pm \frac{1}{2k} - \frac{1}{8k^2} + t_k^\pm) - (k \pm \frac{1}{2}) \\ &= k \log(k) - k \pm \frac{1}{2} \log(k) + \frac{1}{8k} + s_k^\pm \end{aligned} \quad (7.1.1)$$

mit

$$|t_k^\pm| \leq \frac{1}{2k^3}, \quad |s_k^\pm| = \left| kt_k^\pm \pm \frac{1}{2} (t_k^\pm - \frac{1}{8k^2}) \right| \leq \frac{1}{k^2},$$

also

$$\int_{k-1/2}^{k+1/2} \log(x) dx = \log(k) - r_k$$

mit

$$|r_k| \leq |s_k^+| + |s_k^-| \leq \frac{2}{k^2}, \quad k \in \mathbb{N}. \quad (7.1.2)$$

Mit Satz 7.1.4 erhalten wir

$$\begin{aligned} \log(n!) &= \sum_{k=2}^n \left(\int_{k-1/2}^{k+1/2} \log(x) dx + r_k \right) \\ &= \int_1^{n+1/2} \log(x) dx + \sum_{k=2}^n r_k - \int_1^{3/2} \log(x) dx. \end{aligned}$$

Einsetzen von (7.1.1) liefert

$$\begin{aligned} \log(n!) &= (n + \frac{1}{2}) \log(n) - n + \frac{1}{8n} + s_n^+ + 1 \\ &\quad + \sum_{k=2}^n r_k - \int_1^{3/2} \log(x) dx \\ &=: (n + \frac{1}{2}) \log(n) - n + a_n, \end{aligned}$$

wobei wegen (7.1.2) die Folge (a_n) konvergiert. Sei $a = \lim_{n \rightarrow \infty} a_n$. Exponenzieren ergibt

$$n! = \sqrt{n} \cdot n^n e^{-n} e^{a_n}$$

mit

$$b_n = e^{a_n} \rightarrow b = e^a \quad (n \rightarrow \infty).$$

Überraschenderweise kann man b bestimmen, da gemäss Satz 3.2.2 und Beispiel 7.1.3.ii) gilt

$$\begin{aligned} b &= \lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} \frac{b_n^2}{b_{2n}} = \lim_{n \rightarrow \infty} \left(\frac{n! e^n}{\sqrt{nn^n}} \right)^2 \cdot \left(\frac{\sqrt{2n} 2^{2n} n^{2n}}{(2n)! e^{2n}} \right) \\ &= \lim_{n \rightarrow \infty} \sqrt{\frac{2}{n}} \frac{2^{2n} (n!)^2}{(2n)!} = \sqrt{2\pi}. \end{aligned}$$

Wir erhalten somit die Näherungsformel

$$n! \approx \sqrt{2\pi n} \cdot n^n e^{-n}.$$

Satz 7.1.5. (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 6.1.3)

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

folgt die Behauptung unmittelbar aus Definition 7.1.2. □

Bemerkung 7.1.2. Wir können formal in Satz 7.1.5 die Variable $y = g(x)$ "substituieren" mit " $dy = g'(x) dx$ ".

Beispiel 7.1.5. i)

$$\int_0^2 \frac{x}{\sqrt{1+x^2}} dx \stackrel{(y=1+x^2)}{=} \frac{1}{2} \int_1^5 \frac{dy}{\sqrt{y}} = \sqrt{y} \Big|_{y=1}^5 = \sqrt{5} - 1.$$

ii)

$$\int_0^1 \frac{x}{1+x^2} dx \stackrel{(y=1+x^2)}{=} \frac{1}{2} \int_1^2 \frac{dy}{y} = \frac{1}{2} \log(y) \Big|_{y=1}^2 = \frac{\log 2}{2}.$$

iii)

$$\int_{-\pi/2}^{\pi/2} \cos^2(\phi) d\phi \stackrel{(x=\sin(\phi))}{=} \int_{-1}^1 \underbrace{\sqrt{1-x^2}}_{=\sqrt{1-\sin^2(\phi)}=\cos(\phi)} \overbrace{dx}^{=\cos(\phi) d\phi}.$$

Andererseits folgt nach partieller Integration (mit $u = \sin \phi$, $v = \cos \phi$)

$$\begin{aligned} \int_{-\pi/2}^{\pi/2} \cos^2(\phi) d\phi &= \sin(\phi) \cos(\phi) \Big|_{\phi=-\pi/2}^{\pi/2} + \int_{-\pi/2}^{\pi/2} \sin^2(\phi) d\phi \\ &= \int_{-\pi/2}^{\pi/2} (1 - \cos^2(\phi)) d\phi = \pi - \int_{-\pi/2}^{\pi/2} \cos^2(\phi) d\phi; \end{aligned}$$

das heisst,

$$\int_{-1}^1 \sqrt{1-x^2} dx = \int_{-\pi/2}^{\pi/2} \cos^2(\phi) d\phi = \pi/2.$$

iv)

$$\int_{x_0}^{x_1} \frac{g'(x)}{g(x)} dx \stackrel{(y=g(x))}{=} \int_{g(x_0)}^{g(x_1)} \frac{dy}{y} = \log(y) \Big|_{y=g(x_0)}^{g(x_1)} = \log\left(\frac{g(x_1)}{g(x_0)}\right)$$

für $0 < g \in C^1(]a, b[)$, $a < x_0 < x_1 < b$.

v) Insbesondere können wir nun das Anfangswertproblem

$$f' = af, \quad f(0) = f_0 > 0 \tag{7.1.3}$$

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

$$\log\left(\frac{f(x)}{f(0)}\right) = \int_0^x \frac{f'}{f} d\xi = \int_0^x a(\xi) d\xi;$$

das heisst,

$$f(x) = f(0) \cdot \exp\left(\int_0^x a(\xi) d\xi\right)$$

analog zu Beispiel 6.6.1.ii).

vi) Die zu (7.1.3) gehörige inhomogene Gleichung

$$\dot{f} = af + b \tag{7.1.4}$$

mit $a, b \in C^0(\mathbb{R})$ kann man nun ebenfalls lösen. Für eine partikuläre Lösung von (7.1.4) machen wir den Ansatz

$$\begin{aligned} f_{part}(t) &= \int_0^t b(s) e^{\int_s^t a(\tau) d\tau} ds \\ &= e^{\int_0^t a(\tau) d\tau} \int_0^t b(s) e^{-\int_0^s a(\tau) d\tau} ds, \end{aligned} \tag{7.1.5}$$

sofern die Stammfunktionen existieren. Die rechte Seite in (7.1.5) kann man deuten als Überlagerung der Impulsantworten der Gleichung (7.1.3) auf die infinitesimalen Auslenkungen um den Wert $b(s)$ zu den Zeiten $0 \leq s \leq t$. Andererseits erhält man denselben Ausdruck auch leicht aus dem Ansatz

$$f_{part}(t) = c(t)e^{\int_0^t a(\tau) d\tau}$$

mit variablem $c = c(t)$. Daher trägt die Darstellung (7.1.5) auch den Namen *Variation-der-Konstanten Formel*. Wir verifizieren

$$\begin{aligned} \dot{f}_{part}(t) &= a(t)f_{part}(t) + e^{\int_0^t a(\tau) d\tau} \underbrace{\frac{d}{dt} \left(\int_0^t b(s)e^{-\int_0^s a(\tau) d\tau} ds \right)}_{=b(t)e^{-\int_0^t a(\tau) d\tau}} \\ &= a(t)f_{part}(t) + b(t). \end{aligned}$$

Im Fall $a \equiv -\alpha$, $b \equiv \beta \in \mathbb{R}$ erhalten wir so die partikuläre Lösung

$$f_{part}(t) = e^{-\alpha t} \int_0^t \beta e^{\alpha s} ds = e^{-\alpha t} \frac{\beta}{\alpha} (e^{\alpha t} - 1) = \frac{\beta}{\alpha} (1 - e^{-\alpha t})$$

von (6.7.7); vergleiche Beispiel 6.7.2.i).

vii) Analog können wir vorgehen im vektorwertigen Fall. Sei A eine reelle $n \times n$ -Matrix, $B \in C^0(\mathbb{R}; \mathbb{R}^n)$. Für eine partikuläre Lösung der inhomogenen linearen Differentialgleichung

$$F' = AF + B$$

machen wir den Ansatz

$$F_{part}(t) = e^{At} \int_0^t e^{-As} B(s) ds = \int_0^t e^{A(t-s)} B(s) ds,$$

wobei wir komponentenweise integrieren, und verifizieren

$$\begin{aligned} F'_{part}(t) &= \frac{d}{dt} \left(e^{At} \int_0^t e^{-As} B(s) ds \right) = AF_{part}(t) + e^{At} (e^{-As} B(s)) \Big|_{s=t} \\ &= AF_{part}(t) + B(t). \end{aligned}$$

Hier nehmen wir wieder an, dass die Funktion $e^{-As} B(s)$ eine Stammfunktion besitzt. Wie wir später sehen werden, ist dies für $B \in C^0(\mathbb{R}, \mathbb{R}^n)$ jedoch immer der Fall. Ein wenig Vorsicht ist dennoch geboten: Der Ansatz $\Phi(t) = \exp(\int_0^t A(s) ds)$ für die Fundamentallösung der homogenen Gleichung $F' = AF$ funktioniert im Allgemeinen nur im Falle konstanter Koeffizienten. Falls $A = A(t)$ mit der Zeit t variiert, so liefert die Produktregel beispielsweise für den quadratischen Term $Q(t) = (\int_0^t A(s) ds)^2$ in der Exponentialreihe die Identität

$$Q'(t) = \frac{d}{dt} \left(\int_0^t A(s) ds \cdot \int_0^t A(s) ds \right) = A(t) \int_0^t A(s) ds + \int_0^t A(s) ds A(t),$$

und die Gleichung $Q'(t) = 2A(t) \int_0^t A(s) ds$ gilt im Allgemeinen nur dann, wenn sämtliche Kommutatoren $[A(s), A(t)] = A(s)A(t) - A(t)A(s)$ verschwinden. Analog kann man die Identität

$$\frac{d}{dt} \left(\exp\left(\int_0^t A(s) ds\right) \right) = A(t) \exp\left(\int_0^t A(s) ds\right)$$

auch nur in diesem Fall erwarten. Später werden wir jedoch auch für beliebige $A = A(t)$ der Klasse C^0 die Existenz einer Fundamentallösung mit anderen Methoden beweisen.

viii) Mittels Separation wie in Beispiel v) kann man auch gewisse nichtlineare Differentialgleichungen lösen. Beispielsweise geht die Gleichung

$$y' = 2xy^2 \quad (7.1.6)$$

für eine Funktion $y = y(x) > 0$ mit der Anfangsbedingung $y(0) = 1$ nach Separation über in die Form

$$\frac{y'}{y^2} = 2x;$$

das heisst,

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

Integration unter Beachtung der Anfangsbedingung ergibt nun sofort

$$1 - \frac{1}{y(x)} = - \int_0^x \left(\frac{1}{y}\right)' dx = x^2,$$

also

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

Beachte, dass im Unterschied zu linearen Differentialgleichungen die Lösung y von (7.1.6) nur für $|x| < 1$ existiert.

ix) Allgemeiner seien $g, h : \mathbb{R} \rightarrow \mathbb{R}$ stetig mit $h > 0$. Weiter nehmen wir an, es existieren die Stammfunktionen

$$G(x) = \int_0^x g(\xi) d\xi, \quad H(y) = \int_{y_0}^y \frac{d\eta}{h(\eta)}.$$

Nach Separation erhalten wir aus der Gleichung

$$y'(x) = g(x)h(y)$$

für eine Funktion $y \in C^1(\mathbb{R})$ mit Anfangswert $y(0) = y_0$ die Gleichung

$$\frac{y'(x)}{h(y(x))} = g(x)$$

mit separierten Variablen x und y . Es folgt

$$\frac{d(H(y(x)))}{dx} = H'(y(x))y'(x) = \frac{y'(x)}{h(y(x))} = g(x) = G'(x);$$

das heisst,

$$H(y(x)) - H(y_0) = G(x) - G(0).$$

Da $H'(y) = 1/h(y) > 0$, ist H nach Satz 6.2.2 invertierbar mit $H^{-1} \in C^1$, und wir erhalten als Lösung unseres Anfangswertproblems

$$y(x) = H^{-1}(G(x) - G(0) + H(y_0)) \in C^1.$$

Partialbruchzerlegung Mit Hilfe von Satz 7.1.5 kann man die rationalen Funktionen $r(x) = \frac{p(x)}{q(x)}$ elementar integrieren.

Beispiel 7.1.6. Das unbestimmte Integral

$$\int \frac{dx}{1-x^2} = \operatorname{artanh}(x)$$

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

$$\int \frac{dx}{1-x^2} = \frac{1}{2} \int \frac{dx}{1-x} + \frac{1}{2} \int \frac{dx}{1+x} = \frac{1}{2} \log\left(\frac{1+x}{1-x}\right) + c.$$

Beispiel 7.1.6 lässt sich verallgemeinern mit Hilfe des folgenden Satzes.

Satz 7.1.6. Seien p, q Polynome in \mathbb{R} mit

$$\deg(p) < \deg(q) = n$$

und ohne gemeinsame Nullstelle. Sei weiter

$$q(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0 = \prod_{i=1}^l (x-x_i)^{m_i},$$

wobei x_1, \dots, x_l die paarweise verschiedenen Nullstellen $x_i \in \mathbb{C}$ mit Vielfachheit $m_i \in \mathbb{N}$ bezeichnen. Dann gibt es genau eine **Partialbruchzerlegung**

$$\frac{p(x)}{q(x)} = \sum_{i=1}^l \sum_{k=1}^{m_i} \frac{\gamma_{ik}}{(x-x_i)^k} \quad (7.1.7)$$

mit $\gamma_{ik} \in \mathbb{C}$, und für $1 \leq i \leq l$ erhalten wir für $k = m_i, m_i-1, \dots, 1$ iterativ

$$\gamma_{ik} = \lim_{x \rightarrow x_i, x \neq x_i} \left(\frac{p(x)}{q(x)} - \sum_{j=k+1}^{m_i} \frac{\gamma_{ij}}{(x-x_i)^j} \right) (x-x_i)^k. \quad (7.1.8)$$

Beweis. *i) Eindeutigkeit.* Offenbar folgt (7.1.8) aus (7.1.7).

ii) Existenz. Mache den Ansatz (7.1.7) und bringe die rechte Seite auf den Hauptnenner $q(x)$. Der Zähler ist ein Polynom \tilde{p} vom Grad $< n$. Die Gleichung

$$\frac{\tilde{p}(x)}{q(x)} = \frac{p(x)}{q(x)}, \quad x \in \mathbb{R},$$

also die Gleichung $\tilde{p}(x) = p(x)$ für $x \in \mathbb{R}$, liefert ein System von n Gleichungen für die n Koeffizienten $\gamma_{ik} \in \mathbb{C}$, $1 \leq i \leq l$, $1 \leq k \leq m_i$.

Da dieses Gleichungssystem linear ist und nach i) *höchstens* eine Lösung besitzt, folgt die Existenz einer Lösung mit der Rangformel. \square

Bemerkung 7.1.3. Da q nach Annahme reell ist, treten wegen der Identität

$$\prod_{i=1}^l (x - x_i)^{m_i} = q(x) = \overline{q(x)} = \prod_{i=1}^l (x - \overline{x_i})^{m_i}, \quad x \in \mathbb{R},$$

nicht reelle Nullstellen in komplex konjugierten 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

$$\sum_{i=1}^l \sum_{k=1}^{m_i} \frac{\gamma_{ik}}{(x - x_i)^k} = \frac{p(x)}{q(x)} = \overline{\left(\frac{p(x)}{q(x)}\right)} = \sum_{i=1}^l \sum_{k=1}^{m_i} \frac{\overline{\gamma_{ik}}}{(x - \overline{x_i})^k},$$

folgt wegen der Eindeutigkeit der Zerlegung (7.1.1) dann auch

$$\gamma_{(i+1)k} = \overline{\gamma_{ik}}, \quad 1 \leq k \leq m_i.$$

Wir können die entsprechenden Terme somit zusammenfassen

$$\begin{aligned} \frac{\gamma_{ik}}{(x - x_i)^k} + \frac{\gamma_{(i+1)k}}{(x - x_{i+1})^k} &= \frac{\gamma_{ik}(x - x_{i+1})^k + \gamma_{(i+1)k}(x - x_i)^k}{(x - x_i)^k(x - x_{i+1})^k} \\ &= \frac{2\operatorname{Re}(\gamma_{ik}(x - \overline{x_i})^k)}{|x - x_i|^{2k}}, \quad 1 \leq k \leq m_i. \end{aligned}$$

Stammfunktionen lassen sich nun wieder elementar angeben.

Beispiel 7.1.7. 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 7.1.6 und Bemerkung 7.1.3 führen auf die Zerlegung

$$\frac{1}{1 - x^4} = \frac{a}{1 - x} + \frac{b}{1 + x} + \frac{c + dx}{1 + x^2}$$

mit

$$\begin{aligned} a &= \lim_{x \rightarrow 1} \frac{1 - x}{1 - x^4} = \frac{1}{(1 + x)(1 + x^2)} \Big|_{x=1} = \frac{1}{4}, \\ b &= \lim_{x \rightarrow -1} \frac{1 + x}{1 - x^4} = \frac{1}{(1 - x)(1 + x^2)} \Big|_{x=-1} = \frac{1}{4}, \\ c + dx &= \left(\frac{1}{1 - x^4} - \frac{1}{4} \left(\frac{1}{1 - x} + \frac{1}{1 + x} \right) \right) (1 + x^2) \\ &= \frac{1}{1 - x^2} - \frac{1}{2} \cdot \frac{1 + x^2}{1 - x^2} = \frac{1}{2}, \end{aligned}$$

das heisst,

$$c = \frac{1}{2}, \quad d = 0.$$

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 Bemerkung 7.1.3 kann $\frac{3x-1}{x^2+1}$ im Reellen nicht weiter zerlegt werden. Wir erhalten

$$\begin{aligned} \int \frac{x^3 - 2x + 1}{x^2 + 1} dx &= \int x dx - \int \frac{3x}{x^2 + 1} dx + \int \frac{dx}{x^2 + 1} \\ &= \frac{x^2}{2} - \frac{3}{2} \log(1 + x^2) + \arctan(x) + c. \end{aligned}$$

7.2 Das Riemannsche Integral

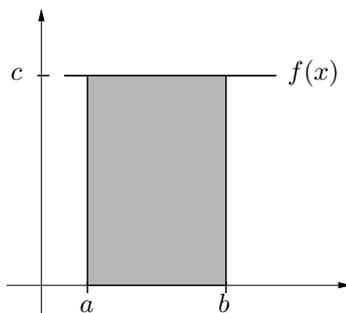
Wir wollen nun den Integralbegriff ausdehnen auf eine möglichst grosse Klasse von Funktionen $f: [a, b] \rightarrow \mathbb{R}$, wobei $-\infty < a < b < \infty$. Insbesondere wollen wir zeigen, dass jedes $f \in C^0([a, b])$ eine Stammfunktion besitzt.

Ausgangspunkt ist die folgende geometrische Interpretation der Stammfunktion.

Beispiel 7.2.1. 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 .

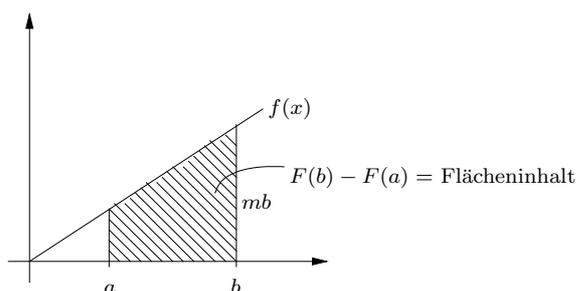


ii) Für $f \equiv c$ mit $c < 0$, $c \in \mathbb{R}$ gilt die Aussage i) analog, sofern wir den Flächeninhalt "mit der richtigen Orientierung" messen; vgl. später.

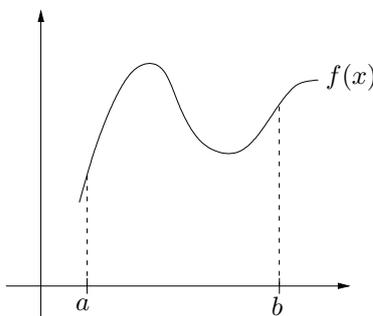
iii) Falls $f(x) = mx$ für ein $m \in \mathbb{R}$ mit Stammfunktion $F(x) = \frac{1}{2}mx^2$, $x \in \mathbb{R}$, so können wir für $a < b$ das Integral

$$\int_a^b mx \, dx = \frac{1}{2}mx^2 \Big|_{x=a}^b = \frac{1}{2}mb^2 - \frac{1}{2}ma^2$$

ebenfalls interpretieren als den “orientierten Flächeninhalt zwischen $[a, b]$ und $\mathcal{G}(f)$ ”.



Fragen: Für welche Funktionen $f: [a, b] \rightarrow \mathbb{R}$ kann man den Flächeninhalt zwischen $[a, b]$ und $\mathcal{G}(f)$ messen? - Liefert dies den gewünschten Integralbegriff?



Definition 7.2.1. i) Eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ heisst **Treppenfunktion**, falls für eine Zerlegung von $I = [a, b]$ in disjunkte (abgeschlossene, offene, halb-offene) Teilintervalle I_1, \dots, I_K mit Konstanten $c_k \in \mathbb{R}$ gilt

$$f(x) = c_k \text{ für } x \in I_k, 1 \leq k \leq K;$$

das heisst,

$$f = \sum_{k=1}^K c_k \chi_{I_k},$$

wobei

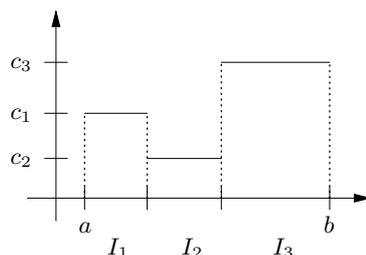
$$\chi_{I_k}(x) = \begin{cases} 1, & x \in I_k \\ 0, & x \notin I_k \end{cases}$$

die charakteristische Funktion von I_k ist.

ii) Das Integral einer Treppenfunktion $f = \sum_{k=1}^K c_k \chi_{I_k} : [a, b] \rightarrow \mathbb{R}$ ist

$$\int_a^b \left(\sum_{k=1}^K c_k \chi_{I_k} \right) dx = \sum_{k=1}^K c_k |I_k|,$$

wobei $|I_k|$ die Länge von I_k bezeichnet, $1 \leq k \leq K$.



Bemerkung 7.2.1. i) Die konstante Funktion $f \equiv c$, $c \in \mathbb{R}$, kann man auch schreiben in der Form einer Treppenfunktion

$$f = \sum_{k=1}^K c_k \chi_{I_k} \quad \text{mit } c_k = c, 1 \leq k \leq K,$$

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 $f = \sum_{k=1}^K c_k \chi_{I_k}$ 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 7.2.1. Sind $e, g : [a, b] \rightarrow \mathbb{R}$ Treppenfunktionen mit $e \leq g$, dann gilt

$$\int_a^b e \, dx \leq \int_a^b g \, dx.$$

Beweis. Seien $e = \sum_{k=1}^K c_k \chi_{I_k}$, $g = \sum_{l=1}^L d_l \chi_{J_l}$ 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

$$I_{kl} = I_k \cap J_l, \quad 1 \leq k \leq K, \quad 1 \leq l \leq L$$

sind dann ebenfalls disjunkt mit

$$I_k = \bigcup_{l=1}^L I_{kl}, \quad J_l = \bigcup_{k=1}^K I_{kl} \quad (7.2.1)$$

für jedes $1 \leq k \leq K$ und jedes $1 \leq l \leq L$, und es gilt

$$\bigcup_{k,l} I_{kl} = \bigcup_{k=1}^K \left(\bigcup_{l=1}^L I_{kl} \right) = \bigcup_{k=1}^K I_k = I.$$

Aus der Annahme $e \leq g$ folgt für beliebige $1 \leq k \leq K$, $1 \leq l \leq L$ die Ungleichung

$$\forall x \in I_{kl} : c_k = e(x) \leq g(x) = d_l;$$

also

$$c_k \leq d_l, \quad \text{falls } I_{kl} \neq \emptyset. \quad (7.2.2)$$

Da mit (7.2.1) auch gilt

$$|I_k| = \sum_{l=1}^L |I_{kl}|, \quad |J_l| = \sum_{k=1}^K |I_{kl}|,$$

erhalten wir mit Bemerkung 7.2.1

$$\begin{aligned} \int_a^b e \, dx &= \sum_{k=1}^K c_k |I_k| = \sum_{k,l} c_k |I_{kl}| \\ &\stackrel{(7.2.2)}{\leq} \sum_{k,l} d_l |I_{kl}| = \sum_{l=1}^L d_l |J_l| = \int_a^b g \, dx, \end{aligned}$$

wie gewünscht. □

Sei $f: [a, b] \rightarrow \mathbb{R}$ beschränkt; das heisst,

$$\exists c \in \mathbb{R} \forall x \in [a, b] : |f(x)| \leq c.$$

Dann gibt es stets (mindestens) ein Paar von Treppenfunktionen $e, g: [a, b] \rightarrow \mathbb{R}$ mit $e \leq f \leq g$, und die folgende Definition ist sinnvoll.

Definition 7.2.2. i) Für beschränktes $f: [a, b] \rightarrow \mathbb{R}$ bezeichnen

$$\int_a^b f \, dx = \sup \left\{ \int_a^b e \, dx; e \text{ Treppenfunktion, } e \leq f \right\},$$

bzw.

$$\int_a^b f \, dx = \inf \left\{ \int_a^b g \, dx; g \text{ Treppenfunktion, } f \leq g \right\}$$

das **untere**, bzw. **obere** Riemann-Integral (R-Integral) von f .

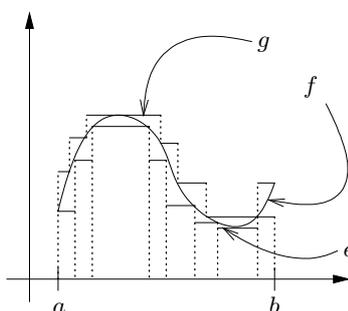
ii) Ein solches f heisst über $[a, b]$ **Riemann-integabel** (*R-integabel*), falls

$$\int_a^b f \, dx = \overline{\int_a^b f \, dx} =: A.$$

In diesem Fall heisst

$$A =: \int_a^b f \, dx$$

das **Riemann-Integral** (*R-Integral*) von f .



Bemerkung 7.2.2. i) Lemma 7.2.1 liefert für jedes beschränkte f die Ungleichung

$$\int_a^b f \, dx \leq \overline{\int_a^b f \, dx}.$$

ii) Eine Funktion f ist R-integabel genau dann, wenn zu jedem $\epsilon > 0$ Treppenfunktionen $e, g: [a, b] \rightarrow \mathbb{R}$ existieren mit $e \leq f \leq g$ und

$$\int_a^b g \, dx - \int_a^b e \, dx < \epsilon.$$

Beweis. i) Für Treppenfunktionen $e, g: [a, b] \rightarrow \mathbb{R}$ mit $e \leq f \leq g$ gilt nach Lemma 7.2.1

$$\int_a^b e \, dx \leq \int_a^b g \, dx;$$

also auch

$$\int_a^b f \, dx = \sup_{e \leq f, e \text{ Treppenfkt.}} \int_a^b e \, dx \leq \int_a^b g \, dx,$$

und die Behauptung folgt nach Übergang zum Infimum bzgl. $g \geq f$.

ii) Die Aussage ii) folgt nun unmittelbar aus der Ungleichung

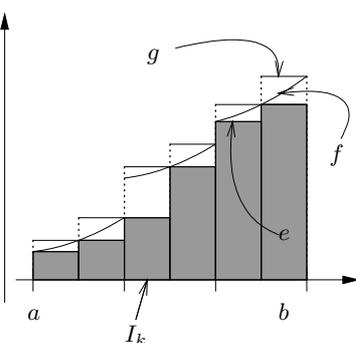
$$\int_a^b e \, dx \leq \int_a^b f \, dx \stackrel{i)}{\leq} \overline{\int_a^b f \, dx} \leq \int_a^b g \, dx$$

für alle Treppenfunktionen $e, g: [a, b] \rightarrow \mathbb{R}$ mit $e \leq f \leq g$.

□

Wir können nun für eine grosse Zahl von Funktionen $f: [a, b] \rightarrow \mathbb{R}$ zeigen, dass sie R-integrabel sind.

Satz 7.2.1. *Sei $f: [a, b] \rightarrow \mathbb{R}$ monoton. Dann ist f über $[a, b]$ R-integrabel.*



Beweis. OBdA sei f monoton wachsend, also

$$\forall a \leq x \leq y \leq b: f(a) \leq f(x) \leq f(y) \leq f(b).$$

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

$$a_k = a + (k-1) \frac{b-a}{K}, \quad b_k = a_k + \frac{b-a}{K} = a_{k+1}, \quad 1 \leq k \leq K,$$

und Länge

$$|I_k| = \frac{b-a}{K}, \quad 1 \leq k \leq K.$$

Dann sind $e = \sum_{k=1}^K c_k \chi_{I_k}$, $g = \sum_{k=1}^K d_k \chi_{I_k}$ mit

$$c_k = \inf_{I_k} f_k, \quad d_k = \sup_{I_k} f_k, \quad 1 \leq k \leq K,$$

Treppenfunktionen mit $e \leq f \leq g$.

Weiter gilt

$$d_k = \sup_{I_k} f \leq f(b_k) = f(a_{k+1}) \leq \inf_{I_{k+1}} f = c_{k+1}, \quad 1 \leq k < K;$$

also

$$\begin{aligned} \int_a^b g \, dx - \int_a^b e \, dx &= \sum_{k=1}^K (d_k - c_k) |I_k| = \frac{b-a}{K} \sum_{k=1}^K (d_k - c_k) \\ &= \frac{b-a}{K} \left(\underbrace{d_K - c_1}_{\leq 2c} + \sum_{k=1}^{K-1} \underbrace{(d_k - c_{k+1})}_{\leq 0} \right) \leq 2c \frac{b-a}{K} \rightarrow 0 \quad (K \rightarrow \infty). \end{aligned}$$

Die Behauptung folgt somit aus Bemerkung 7.2.2.ii). □

Satz 7.2.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist f über $[a, b]$ R -integrabel.

Beweis. Da $[a, b]$ kompakt, ist f nach Satz 5.3.3 und 5.5.2 beschränkt und gleichmässig stetig. Zu $\epsilon > 0$ wähle $\delta > 0$ mit

$$\forall x, y \in [a, b] : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon. \quad (7.2.3)$$

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

$$|I_k| = \frac{b-a}{K}, \quad 1 \leq k \leq K,$$

wie in Satz 7.2.1, und setze

$$c_k = \inf_{I_k} f \leq \sup_{I_k} f = d_k, \quad 1 \leq k \leq K.$$

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 (7.2.3) 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 7.2.2.ii). \square

Das Integral einer Funktion $f \in C^0([a, b])$ kann man numerisch bequem approximieren. Es gilt

Satz 7.2.3. 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**

$$\delta_n = \sup_{1 \leq k \leq K_n} |I_k^n| \rightarrow 0 \quad (n \rightarrow \infty)$$

und eine beliebige Auswahl von Punkten $x_k^n \in I_k^n$, $1 \leq k \leq K_n$, stets

$$\int_a^b \left(\sum_{k=1}^{K_n} f(x_k^n) \chi_{I_k^n} \right) dx = \sum_{k=1}^{K_n} f(x_k^n) |I_k^n| \rightarrow \int_a^b f dx \quad (n \rightarrow \infty)$$

Beweis. Zu $\epsilon > 0$ wähle $\delta \equiv \delta(\epsilon) > 0$ mit (7.2.3) wie in Satz 7.2.2, dazu $n_0 = n_0(\epsilon) \in \mathbb{N}$ mit

$$\forall n \geq n_0 : \delta_n < \delta.$$

Für $n \in \mathbb{N}$ setze weiter

$$c_k^n = \inf_{I_k^n} f \leq f(x_k^n) \leq \sup_{I_k^n} f = d_k^n, \quad 1 \leq k \leq K_n.$$

Wie in Satz 7.2.2 erhalten wir für $n \geq n_0$ die Abschätzung

$$d_k^n - c_k^n \leq \sup_{x,y \in I_k^n} |f(x) - f(y)| \leq \epsilon, \quad 1 \leq k \leq K_n.$$

Definiere die Treppenfunktionen

$$e_n = \sum_{k=1}^{K_n} c_k^n \chi_{I_k^n} \leq f_n = \sum_{k=1}^{K_n} f(x_k^n) \chi_{I_k^n} \leq g_n = \sum_{k=1}^{K_n} d_k^n \chi_{I_k^n}.$$

Da f gemäss Satz 7.2.2 R-integrierbar ist, können wir für $n \geq n_0(\epsilon)$ abschätzen

$$\begin{aligned} R_n &:= \int_a^b f dx - \int_a^b \left(\sum_{k=1}^{K_n} f(x_k^n) \chi_{I_k^n} \right) dx = \int_a^b f dx - \int_a^b f_n dx \\ &\leq \int_a^b g_n dx - \int_a^b e_n dx = \sum_{k=1}^{K_n} (d_k^n - c_k^n) |I_k^n| \leq \epsilon \sum_{k=1}^{K_n} |I_k^n| = (b-a)\epsilon. \end{aligned}$$

Analog erhalten wir $R_n \geq -(b-a)\epsilon$, und die Behauptung folgt. \square

Beispiel 7.2.2. i) Die stetige Funktion $f(x) = e^{x^2}$ ist über jedes Intervall $[a, b]$ R-integrierbar; eine Stammfunktion lässt sich jedoch nicht elementar berechnen.

ii) Die Funktion $f = \chi_{\mathbb{Q} \cap [0,1]}: [0,1] \rightarrow \mathbb{R}$ ist **nicht** R-integrierbar. Für jedes Intervall $I \subset [0,1]$ mit $|I| > 0$ gilt gemäss Beispiel 4.1.4.iv),v)

$$I \cap \mathbb{Q} \neq \emptyset \neq I \setminus \mathbb{Q}$$

und daher

$$0 = \inf_I f < \sup_I f = 1.$$

Für Treppenfunktionen $e = \sum_{k=1}^K c_k \chi_{I_k} \leq f \leq g = \sum_{l=1}^L d_l \chi_{J_l}: [0,1] \rightarrow \mathbb{R}$ folgt

$$\int_0^1 e dx = \sum_{k=1}^n c_k |I_k| \leq 0, \quad \int_0^1 g dx = \sum_{l=1}^L d_l |J_l| \geq 1;$$

also

$$\int_0^1 f dx \leq 0 < 1 \leq \int_0^1 f dx.$$

Mit Hilfe von Satz 7.2.3 können wir den Wert gewisser Summen approximativ berechnen, indem wir sie als “Riemann-Summen” deuten. Dabei benutzen wir, dass nach dem im folgenden Abschnitt bewiesenen Korollar 7.3.4 die Stammfunktion F einer Funktion $f = F'$ mit dem R-Integral übereinstimmt.

Beispiel 7.2.3. i) Für $\alpha > 0$ schreibe

$$\frac{\sum_{k=1}^n k^\alpha}{n^{\alpha+1}} = \frac{1}{n} \sum_{k=1}^n \left(\frac{k}{n}\right)^\alpha$$

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 $\frac{k}{n} = x_k^n$, $1 \leq k \leq n$, und definieren wir $f(x) = x^\alpha$, $x \in I$, so können wir die Summen deuten als Riemann-Summen für f . Gemäss Satz 7.2.3 erhalten wir somit

$$\frac{\sum_{k=1}^n k^\alpha}{n^{\alpha+1}} = \frac{1}{n} \sum_{k=1}^n \left(\frac{k}{n}\right)^\alpha \xrightarrow{(n \rightarrow \infty)} \int_0^1 x^\alpha dx = \frac{1}{\alpha + 1}.$$

ii) Analog erhalten wir

$$\frac{1}{n+1} + \dots + \frac{1}{2n} = \frac{1}{n} \sum_{k=n+1}^{2n} \frac{1}{\left(\frac{k}{n}\right)} \xrightarrow{(n \rightarrow \infty)} \int_1^2 \frac{dx}{x} = \log(2).$$

iii) Es gilt

$$\begin{aligned} \int_1^2 \log(x) dx &= x(\log(x) - 1) \Big|_{x=1}^2 = 2 \log(2) - 1 \\ &= \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n \log\left(1 + \frac{k}{n}\right) \right) = \lim_{n \rightarrow \infty} \log \left(\prod_{k=1}^n \left(1 + \frac{k}{n}\right) \right)^{\frac{1}{n}}; \end{aligned}$$

das heisst,

$$\lim_{n \rightarrow \infty} \prod_{k=1}^n \left(1 + \frac{k}{n}\right)^{\frac{1}{n}} = e^{2 \log(2) - 1} = \frac{4}{e}.$$

7.3 Integrationsregeln, Hauptsatz

Analog zu Satz 7.1.3 für das Rechnen mit Stammfunktionen gilt

Satz 7.3.1. (Monotonie des R-Integrals) Seien $f_{1,2}: [a, b] \rightarrow \mathbb{R}$ beschränkt und R-integrierbar mit $f_1 \leq f_2$. Dann gilt

$$\int_a^b f_1 dx \leq \int_a^b f_2 dx.$$

Beweis. Jede Treppenfunktion $g: [a, b] \rightarrow \mathbb{R}$ mit $f_2 \leq g$ erfüllt auch $f_1 \leq g$; also

$$\int_a^b f_1 \, dx = \overline{\int_a^b f_1 \, dx} \leq \int_a^b g \, dx.$$

Nach Übergang zum Infimum bzgl. derartiger Treppenfunktionen $g \geq f_2$ erhalten wir

$$\int_a^b f_1 \, dx \leq \overline{\int_a^b f_2 \, dx} = \int_a^b f_2 \, dx,$$

wie gewünscht. \square

Bemerkung 7.3.1. Mit der Interpretation des Integrals als (orientierter) Flächeninhalt ist Satz 7.3.1 auch geometrisch evident.

Satz 7.3.2. (Linearität des R-Integrals) Seien $f, f_1, f_2: [a, b] \rightarrow \mathbb{R}$ beschränkt und R-integrierbar, und sei $\alpha \in \mathbb{R}$. Dann sind die Funktionen αf , $f_1 + f_2$ über $[a, b]$ R-integrierbar, und

$$\begin{aligned} \int_a^b (\alpha f) \, dx &= \alpha \int_a^b f \, dx, \\ \int_a^b (f_1 + f_2) \, dx &= \int_a^b f_1 \, dx + \int_a^b f_2 \, dx. \end{aligned}$$

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

$$\begin{aligned} \overline{\int_a^b (\alpha f) \, dx} &\leq \inf \left\{ \underbrace{\int_a^b (\alpha g) \, dx}_{= \alpha \int_a^b g \, dx}; g \text{ Treppenfunktion, } g \geq f \right\} \\ &= \alpha \inf \left\{ \int_a^b g \, dx; g \text{ Treppenfunktion, } g \geq f \right\} = \alpha \overline{\int_a^b f \, dx} \\ &\stackrel{(f \text{ R-int.})}{=} \alpha \underline{\int_a^b f \, dx} = \alpha \sup \left\{ \int_a^b e \, dx; e \text{ Treppenfunktion, } e \leq f \right\} \\ &= \sup \left\{ \underbrace{\alpha \int_a^b e \, dx}_{= \int_a^b (\alpha e) \, dx}; e \text{ Treppenfunktion, } e \leq f \right\} \leq \underline{\int_a^b (\alpha f) \, dx}. \end{aligned}$$

Nach Bemerkung 7.2.2.i) ist die Funktion αf daher R-integrierbar mit

$$\int_a^b (\alpha f) \, dx = \alpha \int_a^b f \, dx.$$

Analog für $\alpha = -1$ (und damit auch für beliebiges $\alpha < 0$).

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:

$$\begin{aligned} \overline{\int_a^b (f_1 + f_2) dx} &\leq \inf \left\{ \underbrace{\int_a^b (g_1 + g_2) dx}_{= \int_a^b g_1 dx + \int_a^b g_2 dx}; g_i \text{ Treppenfkt.}, g_i \geq f_i, i = 1, 2 \right\} \\ &= \inf \left\{ \int_a^b g_1 dx; g_1 \text{ Treppenfunktion}, g_1 \geq f_1 \right\} \\ &\quad + \inf \left\{ \int_a^b g_2 dx; g_2 \text{ Treppenfunktion}, g_2 \geq f_2 \right\} \\ &= \underline{\int_a^b f_1 dx} + \underline{\int_a^b f_2 dx} = \underline{\int_a^b f_1 dx} + \underline{\int_a^b f_2 dx} \\ &= \sup \left\{ \int_a^b e_1 dx; e_1 \text{ Treppenfunktion}, e_1 \leq f_1 \right\} \\ &\quad + \sup \left\{ \int_a^b e_2 dx; e_2 \text{ Treppenfunktion}, e_2 \leq f_2 \right\} \\ &= \sup \left\{ \int_a^b (e_1 + e_2) dx; e_i \text{ Treppenfkt.}, e_i \leq f_i, i = 1, 2 \right\} \\ &\leq \underline{\int_a^b (f_1 + f_2) dx} \leq \overline{\int_a^b (f_1 + f_2) dx} . \end{aligned}$$

Also ist $f_1 + f_2$ R-integabel mit

$$\int_a^b (f_1 + f_2) dx = \int_a^b f_1 dx + \int_a^b f_2 dx .$$

□

Korollar 7.3.1. 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 Aussage aus Satz 7.3.1 und 7.3.2. □

Korollar 7.3.2. Seien $f, f_k \in C^0([a, b])$ mit $f_k \xrightarrow{glm} f$ ($k \rightarrow \infty$). Dann gilt

$$\left| \int_a^b f_k dx - \int_a^b f dx \right| \leq \int_a^b |f_k - f| dx \leq (b - a) \|f_k - f\|_{C^0} \rightarrow 0 \quad (k \rightarrow \infty).$$

Beweis. Unmittelbar aus Satz 7.3.2 und Korollar 7.3.1. \square

Gemäss Beispiel 5.7.2 sind Potenzreihen

$$p(x) = \sum_{k=0}^{\infty} c_k x^k$$

für jedes

$$r < \rho = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|c_k|}}$$

in $B_r(0)$ gleichmässig konvergent. Korollar 7.3.2 zusammen mit Korollar 7.3.4 ergibt somit für $-\rho < a < b < \rho$ die Darstellung

$$\begin{aligned} \int_a^b p(x) dx &= \int_a^b \left(\lim_{n \rightarrow \infty} \sum_{k=0}^n c_k x^k \right) dx = \lim_{n \rightarrow \infty} \int_a^b \sum_{k=0}^n c_k x^k dx \\ &= \lim_{n \rightarrow \infty} \left(\sum_{k=0}^n c_k \int_a^b x^k dx \right) = \sum_{k=0}^{\infty} \frac{c_k}{k+1} (b^{k+1} - a^{k+1}). \end{aligned}$$

Korollar 7.3.3. *Potenzreihen dürfen im Innern ihres Konvergenzkreises gliedweise integriert werden.*

Beispiel 7.3.1. Für $0 \leq b < 1$ erhalten wir

$$\begin{aligned} \log(1+b) &= \int_1^{1+b} \frac{dx}{x} = \int_1^{1+b} \frac{dx}{1-(1-x)} \\ &= \int_1^{1+b} \left(\sum_{k=0}^{\infty} (1-x)^k \right) dx = \sum_{k=0}^{\infty} \frac{-(1-x)^{k+1}}{k+1} \Big|_{x=1}^{1+b} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k b^{k+1}}{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1} b^k}{k}; \end{aligned}$$

das heisst, wir erhalten die Taylor-Reihe der Funktion \log um $x_0 = 1$. Beachte, dass nach Beispiel 3.5.1.iii) die alternierende Reihe $\sum_{k=0}^{\infty} (-1)^{k-1} \frac{b^k}{k}$ sogar für alle $0 \leq b \leq 1$ konvergiert mit Fehlerabschätzung

$$\left| \log(1+b) - \sum_{k=1}^n \frac{(-1)^{k-1} b^k}{k} \right| \leq \frac{b^{n+1}}{n+1}, \quad n \in \mathbb{N}.$$

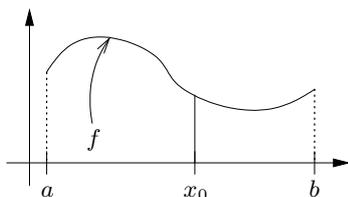
Grenzübergang $b \rightarrow 1$ und anschliessend $n \rightarrow \infty$ liefert die Summenformel

$$\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} = \log 2$$

für die alternierende harmonische Reihe.

Satz 7.3.3. (Gebietsadditivität) Sei $f: [a, b] \rightarrow \mathbb{R}$ beschränkt und R-integrierbar über $[a, b]$, und sei $x_0 \in [a, b]$. Dann sind die Funktionen $f|_{[a, x_0]}$, bzw. $f|_{[x_0, b]}$ R-integrierbar, 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 $f: [a, b] \rightarrow \mathbb{R}$ R-integrierbar. 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 $e \leq f \leq g$ auf $[a, x_0]$, bzw. auf $[x_0, b]$, und nach i) gilt weiter

$$\begin{aligned} & \underbrace{\left(\int_a^{x_0} g \, dx - \int_a^{x_0} e \, dx \right)}_{\geq 0} + \underbrace{\left(\int_{x_0}^b g \, dx - \int_{x_0}^b e \, dx \right)}_{\geq 0} \\ & = \int_a^b g \, dx - \int_a^b e \, dx < \epsilon. \end{aligned}$$

Also ist f über $[a, x_0]$ sowie über $[x_0, b]$ nach Bemerkung 7.2.2.ii) R-integrierbar.

Weiter gilt

$$\begin{aligned} A & := \int_a^b f \, dx - \left(\int_a^{x_0} f \, dx + \int_{x_0}^b f \, dx \right) \\ & \leq \int_a^b g \, dx - \left(\int_a^{x_0} e \, dx + \int_{x_0}^b e \, dx \right) = \int_a^b g \, dx - \int_a^b e \, dx < \epsilon; \end{aligned}$$

analog $A > -\epsilon$. Da $\epsilon > 0$ beliebig, folgt die Behauptung. □

Aus Satz 7.3.3 folgt nun der ‘‘Hauptsatz der Differential- und Integralrechnung’’.

Satz 7.3.4. Sei $f \in C^0([a, b])$. Setze

$$F: x \mapsto \int_a^x f \, dx, \quad x \in [a, b].$$

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 7.3.3 folgt:

$$F(x) - F(x_0) = \int_{x_0}^x f \, d\xi.$$

Schätze ab mittels Korollar 7.3.1:

$$\begin{aligned} \left| \int_{x_0}^x f \, d\xi - (x - x_0) \cdot f(x_0) \right| &= \left| \int_{x_0}^x (f(\xi) - f(x_0)) \, d\xi \right| \\ &\leq |x - x_0| \sup_{|y-x_0|<\delta} |f(y) - f(x_0)| \leq \epsilon |x - x_0|. \end{aligned}$$

Es folgt:

$$\sup_{x_0 < x < x_0 + \delta} \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \epsilon.$$

Analog für $x_0 - \delta < x < x_0$. Die Behauptung folgt. \square

Korollar 7.3.4. 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 7.3.4

$$\frac{d}{dx} \left(F(x) - \int_a^x f \, dx \right) = f - f = 0,$$

also wegen Korollar 6.2.1.i)

$$F(x) - \int_a^x f \, dx = \text{const} \stackrel{(x=a)}{=} F(a).$$

\square

7.3.1 Das R-Integral vektorwertiger Funktionen

Definition 7.3.1. $f = (f_i)_{1 \leq i \leq n} : [a, b] \rightarrow \mathbb{R}^n$ ist R-integrierbar über $[a, b]$ genau dann, wenn $f_i : [a, b] \rightarrow \mathbb{R}$ R-integrierbar ist für $1 \leq i \leq n$ und

$$\int_a^b f \, dx = \left(\int_a^b f_1 \, dx, \dots, \int_a^b f_n \, dx \right)^t.$$

Analog zu Korollar 7.3.1 gilt:

Satz 7.3.5. Sei $f \in C^0([a, b]; \mathbb{R}^n)$. Dann gilt:

$$\left| \int_a^b f \, dx \right| \leq \int_a^b |f| \, dx \leq |b - a| \|f\|_{C^0},$$

wobei $|\cdot|$ die euklidische Norm auf \mathbb{R}^n bezeichnet.

Beweis. Setze

$$P := \int_a^b f \, dx \in \mathbb{R}^n.$$

Dann gilt:

$$|P|^2 = \left(\int_a^b f \, dx \right) \cdot P \stackrel{(\text{Satz 7.3.2})}{=} \int_a^b (f \cdot P) \, dx \stackrel{(\text{Cauchy-Schwarz})}{\leq} \int_a^b |f| \, dx \cdot |P|;$$

also mit Korollar 7.3.1:

$$\left| \int_a^b f \, dx \right| = |P| \leq \int_a^b |f| \, dx \leq |b - a| \|f\|_{C^0}.$$

□

7.4 Uneigentliches Riemann-Integral

Sei $f:]a, b[\rightarrow \mathbb{R}$ über jedes kompakte Intervall $[c, d] \subset]a, b[$ R-integrierbar.

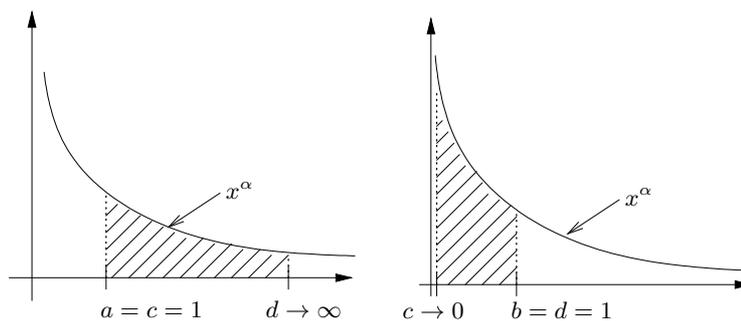
Definition 7.4.1. f heißt über $]a, b[$ **uneigentlich R-integrierbar**, falls

$$\int_a^b f \, dx := \lim_{c \downarrow a, d \uparrow b} \int_c^d f \, dx$$

existiert.

Beispiel 7.4.1. i) Für $\alpha < -1$ existiert

$$\int_1^\infty x^\alpha \, dx = \lim_{d \rightarrow \infty} \int_1^d x^\alpha \, dx = \lim_{d \rightarrow \infty} \frac{d^{\alpha+1} - 1}{\alpha + 1} = \frac{1}{|\alpha| - 1}.$$



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) Die uneigentlichen Integrale

$$\int_0^1 \frac{dx}{x} = \lim_{c \downarrow 0} \log\left(\frac{1}{c}\right), \quad \int_1^\infty \frac{dx}{x} = \lim_{d \rightarrow \infty} \log d$$

existieren nicht.

iv)

$$\int_0^\infty e^{-t} dt = \lim_{d \rightarrow \infty} (1 - e^{-d}) = 1.$$

v) Für alle $\alpha > 0$ existiert

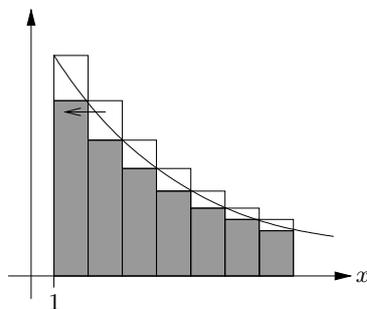
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt;$$

vgl. Beispiel 7.1.2.iii).

Die Konvergenz gewisser Reihen lässt sich auf die Konvergenz von uneigentlichen Integralen zurückführen.

Satz 7.4.1. Sei $f: [1, \infty[\rightarrow \mathbb{R}_+$ monoton fallend. Dann konvergiert die Reihe $\sum_{k=1}^\infty f(k)$ genau dann, wenn das uneigentliche Integral $\int_1^\infty f dx$ konvergiert, und in diesem Fall gilt

$$0 \leq \sum_{k=1}^\infty f(k) - \int_1^\infty f dx \leq f(1).$$



Beweis. Die Treppenfunktionen

$$e = \sum_{k=1}^\infty f(k+1) \chi_{[k, k+1[}, \quad g = \sum_{k=1}^\infty f(k) \chi_{[k, k+1[}$$

erfüllen wegen der Monotonie von f die Ungleichung $e \leq f \leq g$; somit folgt

$$\begin{aligned} \int_1^n e dx &= \sum_{k=1}^{n-1} f(k+1) = \sum_{k=1}^n f(k) - f(1) \leq \int_1^n f dx \\ &\leq \int_1^n g dx = \sum_{k=1}^{n-1} f(k) = \sum_{k=1}^n f(k) - f(n). \end{aligned}$$

Wir erhalten also

$$0 < f(n) \leq \sum_{k=1}^n f(k) - \int_1^n f \, dx \leq f(1).$$

Die Behauptung folgt. \square

Beispiel 7.4.2. i) $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$ existiert für alle $s > 1$ gemäss Satz 7.4.1 und Beispiel 7.4.1.i); vgl. Beispiel 3.7.4. Satz 7.4.1 liefert zudem die Abschätzung

$$0 \leq \zeta(s) - \int_1^{\infty} x^{-s} \, dx = \zeta(s) - \frac{1}{s-1} \leq 1$$

für alle $s > 1$.

ii) Die Reihe $\sum_{k=2}^{\infty} \frac{1}{k(\log k)^s}$ konvergiert gemäss Satz 7.4.1 für alle $s > 1$, da nach Beispiel 7.4.1.i)

$$\int_2^{\infty} \frac{dx}{x \log^s x} \stackrel{(y=\log x)}{=} \int_{\log 2}^{\infty} \frac{dy}{y^s} < \infty.$$

7.5 Differentialgleichungen

Sei $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig, $u_0 \in \mathbb{R}^n$. Gesucht ist eine Lösung $u \in C^1([0, T], \mathbb{R}^n)$ der Differentialgleichung

$$\dot{u} = \frac{du}{dt} = f(t, u(t)), \quad 0 \leq t < T, \quad (7.5.1)$$

mit Anfangsbedingung

$$u(0) = u_0. \quad (7.5.2)$$

Beispiel 7.5.1. i) Die allgemeine lineare Differentialgleichung (6.6.4) lässt sich in der Form (7.5.1) schreiben mit $f(t, y) = Ay$, $y \in \mathbb{R}^n$.

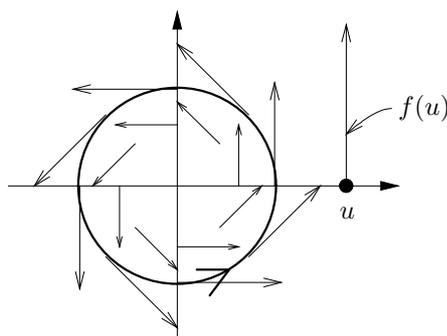
ii) Die Form (7.5.1) umfasst aber auch nichtlineare Gleichungen wie in Beispiel 6.7.2.ii) oder inhomogene Gleichungen.

Geometrisch können wir die Lösungen u von (7.5.1) als “Integralkurven” des durch f gegebenen “Richtungsfeldes” deuten.

Beispiel 7.5.2. i) Sei $n = 2$, $\mathbb{R}^2 \cong \mathbb{C}$, $f: \mathbb{C} \rightarrow \mathbb{C}$ 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 (7.5.1) das Gleichungssystem

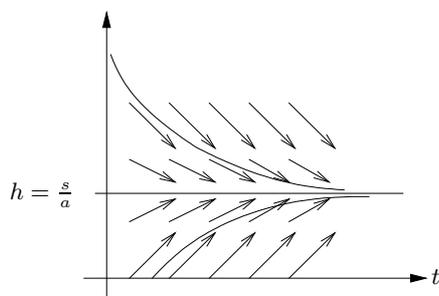
$$\dot{U} = \begin{pmatrix} \frac{dt}{dt} \\ \frac{du}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ f(t, u(t)) \end{pmatrix} = F(U(t))$$

mit $F(t, u) = (1, f(t, u)) \in C^0(\mathbb{R}^2, \mathbb{R}^2)$.

Im Falle $f(t, u) = s - au$ mit Konstanten $s, a > 0$ erhält man ein Richtungsfeld, das sehr schön zeigt, wie jede Lösung $u \in C^1(\mathbb{R})$ der Gleichung

$$\dot{u} = s - au$$

für $t \rightarrow \infty$ gegen die Gleichgewichtslage s/a konvergiert; vgl. Beispiel 6.7.2.i).



Fragen Was für lineare Differentialgleichungen selbverständlich war, muss im allgemeinen Fall des Anfangswertproblems (7.5.1), (7.5.2) 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 (7.5.1), (7.5.2) 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 7.5.3. i) Sei $n = 1$, $f(u) = u^2$, $u_0 > 0$. Nach Separation erhält man zu dem Anfangswertproblem

$$\dot{u} = u^2, \quad u(0) = u_0 \quad (7.5.3)$$

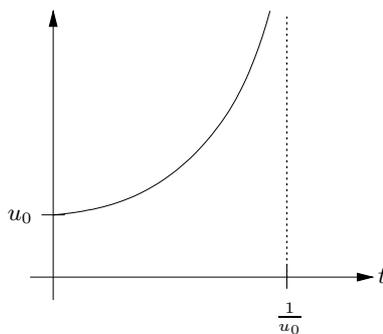
die äquivalente Form

$$\frac{d}{dt} \left(-\frac{1}{u} \right) = \frac{\dot{u}}{u^2} = 1, \quad u(0) = u_0$$

mit der eindeutigen Lösung

$$u(t) = \frac{1}{1/u_0 - t}, \quad 0 \leq t < 1/u_0.$$

Offenbar gilt $u(t) \rightarrow \infty$ für $t \uparrow 1/u_0$; die Lösung u lässt sich demnach nicht für alle $t > 0$ fortsetzen.



Beachte, dass gilt $f(u)/u \rightarrow \infty$ ($n \rightarrow \infty$).

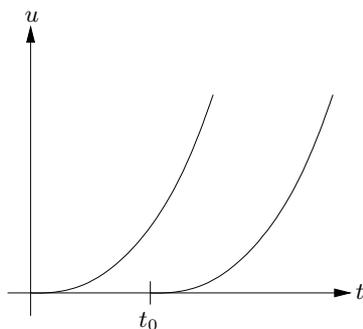
ii) Sei $n = 1$. Das Anfangswertproblem

$$\dot{u} = 2\sqrt{|u|}, \quad u(0) = 0, \quad (7.5.4)$$

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 (7.5.4), wobei der Parameter $t_0 \geq 0$ beliebig gewählt werden kann.



Beachte, dass die Funktion $f(u) = 2\sqrt{|u|}$ in keiner Umgebung von $u_0 = 0$ Lipschitz stetig ist, da

$$\frac{|f(u) - f(0)|}{|u|} = \frac{1}{\sqrt{|u|}} \rightarrow \infty \quad (u \rightarrow 0).$$

Wenn wir uns nun der Aufgabe zuwenden, das Anfangswertproblem (7.5.1), (7.5.2) zu lösen, müssen wir also einerseits damit rechnen, dass Lösungen von (7.5.1), (7.5.2) im Allgemeinen in endlicher Zeit “explodieren”; andererseits gilt es, geeignete zusätzliche Voraussetzungen an f zu finden, welche die Eindeutigkeit der Lösungen garantieren.

Satz 7.5.1. (Picard-Lindelöf) Sei $f = f(t, u): \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $t_0 > 0$ existieren $r_0 > 0$, $T_0 > 0$ und eine Konstante L , so dass

$$|f(t, u) - f(t, v)| \leq L|u - v| \text{ für } |t - t_0| < T_0 \text{ und alle } u, v \in B_{r_0}(u_0). \quad (7.5.5)$$

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 (7.5.1), (7.5.2).

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 (7.5.1) für $0 \leq t \leq T$ erklärt, und mit einer Konstanten $C \in \mathbb{R}$ gilt

$$\|u(t; u_0) - u(t; u_1)\|_{C^1([0, T])} \leq C|u_1 - u_0|.$$

wobei $T > 0$ wie in i) und $r_0 > 0$ wie in (7.5.5) zu u_0 und $t_0 = 0$ gewählt sind.

Beispiel 7.5.4. i) Die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(u) = u^2$ aus Beispiel 7.5.3.i) zeigt, dass wir im Allgemeinen nur die Existenz einer *lokalen Lösung* von (7.5.1), (7.5.2) erwarten dürfen, das heisst, einer Lösung in einer Umgebung von $t = 0$.

ii) Das Beispiel der Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ mit $f(u) = \sqrt{|u|}$ aus Beispiel 7.5.3.ii) zeigt, dass die Annahme (7.5.5) im Allgemeinen notwendig ist für die Eindeutigkeit der Lösungen von (7.5.1), (7.5.2).

iii) Die Annahme (7.5.5) ist für “autonome” Funktionen $f = f(u) \in C^1(\mathbb{R})$ stets erfüllt. Für jedes $R > 0$ und beliebige $u, v \in B_R(0)$ gibt es nämlich gemäss dem Mittelwertsatz, Satz 6.2.1, ein $w \in B_R(0)$ mit

$$f(u) - f(v) = f'(w)(u - v).$$

Zum Nachweis von (7.5.5) genügt es daher,

$$L = \sup_{w \in B_R(0)} |f'(w)| < \infty$$

zu setzen.

iv) Sei $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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

$$|f(t, u) - f(t, v)| \leq |A(t)| |u - v| \leq \|A\|_{C^0([-T, T])} |u - v|;$$

das heisst, wir erhalten (7.5.5) mit $L = \|A\|_{C^0([-T, T])}$.

Den Beweis der Aussage i) von Satz 7.5.1 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.

7.5.1 Der Banachsche Fixpunktsatz

Sei (M, d) ein metrischer Raum.

Definition 7.5.1. Eine Abbildung $\Phi : M \rightarrow M$ heisst **kontrahierend**, falls für eine Konstante $q < 1$ gilt

$$\forall x, y \in M : d(\Phi(x), \Phi(y)) \leq q \cdot d(x, y).$$

Satz 7.5.2. (Banachscher Fixpunktsatz, Kontraktionsprinzip) Sei (M, d) vollständig, und sei $\Phi : M \rightarrow M$ kontrahierend mit Konstante $q < 1$. Dann gibt es genau ein $\bar{x} \in M$ mit $\Phi(\bar{x}) = \bar{x}$. Zudem gilt für jedes $x_0 \in M$ und die Folge

$$x_1 = \Phi(x_0), \dots, x_k = \Phi(x_{k-1}) = \dots = \Phi^k(x_0), \quad k \in \mathbb{N}, \quad (7.5.6)$$

die Abschätzung

$$d(x_k, \bar{x}) \leq q^k d(x_0, \bar{x}), \quad k \in \mathbb{N}.$$

Beispiel 7.5.5. Sei $(X, \|\cdot\|_X)$ ein Banachraum, $M \subset X$ mit der induzierten Metrik

$$d(x, y) = \|x - y\|_X, \quad x, y \in M.$$

(Im Beweis von Satz 7.5.1 werden wir $X = C^0([0, T]; \mathbb{R}^n)$ wählen.) Dann ist (M, d) vollständig im Sinne von Definition 3.5.3.iii), falls $M \subset X$ abgeschlossen ist im Sinne von Satz 4.2.1, das heisst, falls gilt:

$$\forall (x_k) \subset M : x_k \rightarrow x \quad (k \rightarrow \infty) \Rightarrow x \in M.$$

Weiter ist eine Abbildung $\Phi : M \rightarrow M$ kontrahierend, falls gilt

$$\exists q < 1 \quad \forall x, y \in M : \|\Phi(x) - \Phi(y)\|_X \leq q \|x - y\|_X,$$

das heisst, falls Φ Lipschitz stetig ist mit Lipschitz Konstante $q < 1$.

Bemerkung 7.5.1. Falls $(X, \|\cdot\|_X)$ ein Banachraum, $M \subset X$ abgeschlossen, $\Phi : M \rightarrow M$ kontrahierend, so existiert gemäss Satz 7.5.2 genau ein $\bar{x} \in M$ mit

$\Phi(\bar{x}) = \bar{x}$, und für jedes $x_0 \in M$ und die gemäss (7.5.6) definierte Folge (x_k) gilt

$$\|x_k - \bar{x}\|_X \leq q^k \|x_0 - \bar{x}\|_X, \quad k \in \mathbb{N}.$$

Beweis von Satz 7.5.2. i) Wähle ein $x_0 \in M$ und definiere $(x_k)_{k \in \mathbb{N}} \subset M$ wie in (7.5.6). Schätze ab

$$\begin{aligned} d(x_k, x_{k+1}) &= d(\Phi(x_{k-1}), \Phi(x_k)) \leq q d(x_{k-1}, x_k) \\ &\leq \dots \leq q^k d(x_0, x_1), \quad k \in \mathbb{N}. \end{aligned}$$

Für $l \geq k \in \mathbb{N}$ folgt so

$$\begin{aligned} d(x_k, x_l) &\leq \sum_{j=k}^{l-1} d(x_j, x_{j+1}) \leq \sum_{j=k}^{l-1} q^j d(x_0, x_1) \\ &\leq \frac{q^k}{1-q} d(x_0, x_1) \rightarrow 0 \quad (l \geq k \rightarrow \infty); \end{aligned}$$

das heisst, $(x_k)_{k \in \mathbb{N}}$ ist Cauchy-Folge in M .

Da M nach Annahme vollständig ist, existiert $\bar{x} = \lim_{k \rightarrow \infty} x_k \in M$. Da Φ insbesondere stetig ist, ergibt (7.5.6) die gewünschte Beziehung

$$\bar{x} = \lim_{k \rightarrow \infty} x_k = \lim_{k \rightarrow \infty} \Phi(x_{k-1}) = \Phi(\bar{x}).$$

Die behauptete Fehlerabschätzung erhält man nun mittels

$$d(x_k, \bar{x}) = d(\Phi(x_{k-1}), \Phi(\bar{x})) \leq q d(x_{k-1}, \bar{x}) \leq \dots \leq q^k d(x_0, \bar{x}).$$

ii) Zum Beweis der Eindeutigkeit von \bar{x} seien $x, y \in M$ Fixpunkte von Φ . Mit

$$d(x, y) = d(\Phi(x), \Phi(y)) \leq q d(x, y)$$

folgt $d(x, y) = 0$, also $x = y$.

□

Die Annahme der Vollständigkeit von M und die Kontraktionsbedingung kann man im Allgemeinen nicht weiter abschwächen, wie die folgenden Beispiele zeigen.

Beispiel 7.5.6. i) Die Abbildung

$$f:]0, 1] \ni x \mapsto \frac{x}{2} \in]0, 1]$$

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 $\bar{x} = 0$ als Fixpunkt.)

ii) Die Funktion $f: \{0, 1\} \rightarrow \{0, 1\}$ mit $f(0) = 1$, $f(1) = 0$ ist Lipschitz-stetig mit Lipschitz-Konstante $L = 1$, hat aber keinen Fixpunkt.

iii) Sei

$$f: [1, \infty[\rightarrow [1, \infty[, \quad x \mapsto x + \frac{1}{x}$$

mit

$$|f(x) - f(y)| = \left| (x - y) + \frac{y - x}{xy} \right| = \left(1 - \frac{1}{xy}\right) |x - y|.$$

Beachte, dass

$$\forall x, y \geq 1: 0 \leq 1 - \frac{1}{xy} < 1;$$

jedoch besitzt f offenbar keinen Fixpunkt.

iv) Sei $1 \leq a \leq 2$, und sei $f: [1, \infty[\rightarrow \mathbb{R}$ mit

$$f(x) = \frac{1}{2} \left(x + \frac{a}{x}\right);$$

vgl. Beispiel 3.3.1.iii). Beachte

$$f'(x) = \frac{1}{2} \left(1 - \frac{a}{x^2}\right) \in \left[-\frac{1}{2}, \frac{1}{2}\right], \quad f''(x) = \frac{a}{x^3} > 0 \text{ für } x \geq 1;$$

insbesondere hat f genau eine Minimalstelle bei $x = \sqrt{a}$, und wir erhalten

$$f(x) \geq f(\sqrt{a}) = \sqrt{a} \geq 1 \text{ für } x \geq 1;$$

also $f: [1, \infty[\rightarrow [1, \infty[$. Weiter gibt es zu $1 \leq x < y$ gemäss Mittelwertsatz, Satz 6.2.1, ein $\xi \in]x, y[$ mit

$$|f(y) - f(x)| = |f'(\xi)| |y - x| \leq \frac{1}{2} |y - x|,$$

und f ist kontrahierend.

Gemäss Satz 7.5.2 ist der Fixpunkt $\bar{x} = \sqrt{a}$ von f eindeutig bestimmt. Satz 7.5.2 liefert zudem ein Verfahren zur näherungsweise Berechnung von \sqrt{a} für jedes $a \in [1, 2]$ mit der Fehlerabschätzung

$$|x_k - \sqrt{a}| \leq \left(\frac{1}{2}\right)^k |x_0 - \sqrt{a}|$$

für die gemäss (7.5.6) bestimmte Folge von Näherungen x_k , $k \in \mathbb{N}_0$.

7.5.2 Beweis von Satz 7.5.1

Zu gegebenem $u_0 \in \mathbb{R}^n$ existieren nach Voraussetzung $r_0 > 0$, $T_0 > 0$, $L \in \mathbb{R}$ so, dass gilt

$$|f(t, v) - f(t, w)| \leq L |v - w| \quad (7.5.7)$$

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

$$T = \min\left\{T_0, \frac{r_0}{2C_0}, \frac{1}{2L}\right\} > 0$$

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 $\Phi_{u_1}: M \rightarrow X$ 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 7.3.4 zum Nachweis, dass dieser Fixpunkt die gesuchte Lösung $u = u(t; u_1) \in C^1([0, T], \mathbb{R}^n)$ von (7.5.1) mit $u(0) = u_1$ ergibt. Dazu verifizieren wir zunächst die Voraussetzungen des Satzes 7.5.2.

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 7.3.1

$$\begin{aligned} |(\Phi_{u_1}(v))(t) - u_0| &\leq |u_1 - u_0| + \left| \int_0^t f(s, v(s)) \, ds \right| \\ &\leq r_0/2 + T \sup_{0 \leq s \leq T} |f(s, v(s))|, \quad 0 \leq t \leq T. \end{aligned}$$

Da wegen (7.5.7) 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

$$|(\Phi_{u_1}(v))(t) - u_0| \leq r_0, \quad 0 \leq t \leq T,$$

und $\Phi_{u_1}(v) \in M$ wie gewünscht. \square

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 (7.5.7) und Satz 7.3.5 ab

$$\begin{aligned} &|(\Phi_{u_1}(v))(t) - (\Phi_{u_2}(w))(t)| \\ &= \left| u_1 - u_2 + \int_0^t (f(s, v(s)) - f(s, w(s))) \, ds \right| \\ &\leq |u_1 - u_2| + L \int_0^t |v(s) - w(s)| \, ds \\ &\leq |u_1 - u_2| + LT \|v - w\|_{C^0([0, T])}, \end{aligned} \tag{7.5.8}$$

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 $\Phi_{u_1}: M \rightarrow M$ kontrahierend.

Beweis. Bei Wahl von $u_1 = u_2 \in B_{r_0/2}(u_0)$ ergibt sich wegen $LT \leq 1/2$ aus (7.5.8) sofort

$$\forall v, w \in M : \|\Phi_{u_1}(v) - \Phi_{u_1}(w)\|_{C^0} \leq \frac{1}{2} \|v - w\|_{C^0}.$$

□

Wir können nun den Beweis vollenden.

Beweis von Satz 7.5.1. i) Da $\Phi_{u_1} : M \rightarrow M$ kontrahierend, hat Φ_{u_1} gemäss Satz 7.5.2 für alle $u_1 \in B_{r_0/2}(u_0)$ genau einen Fixpunkt $u = u(t) \in M$ mit

$$u(t) = u_1 + \int_0^t f(s, u(s)) ds, \quad 0 \leq t \leq T. \quad (7.5.9)$$

Nach Satz 7.3.4 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 (7.5.1) mit $u(0) = u_1$.

Da umgekehrt jede Lösung dieses Anfangswertproblems auch (7.5.9) erfüllt, folgt mit Satz 7.5.2 auch die Eindeutigkeit dieser Lösung, und Aussage i) von Satz 7.5.1 ist bewiesen.

ii) Seien $v = \Phi_{u_1}(v)$, $w = \Phi_{u_2}(w)$ die Lösungen des Anfangswertproblems (7.5.1) mit $v(0) = u_1$, bzw. $w(0) = u_2$. Mit (7.5.8) folgt zunächst

$$\|v - w\|_{C^0} = \|\Phi_{u_1}(v) - \Phi_{u_2}(w)\|_{C^0} \leq |u_1 - u_2| + \frac{1}{2} \|v - w\|_{C^0};$$

das heisst,

$$\|v - w\|_{C^0} \leq 2|u_1 - u_2|.$$

Gleichung (7.5.1) zusammen mit (7.5.7) 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

$$\begin{aligned} \|v - w\|_{C^1} &\leq \|v - w\|_{C^0} + L\|v - w\|_{C^0} \\ &= (1 + L)\|v - w\|_{C^0} \leq 2(1 + L)|u_1 - u_2|, \end{aligned}$$

und somit Aussage ii).

□

7.5.3 Stetige Abhängigkeit von den Daten

Die Lösung von (7.5.1), (7.5.2) hängt auch stetig ab von der rechten Seite f im folgenden Sinne.

Satz 7.5.3. Seien $f, f_k : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig und bezüglich $u \in \mathbb{R}^n$ lokal Lipschitz stetig, lokal gleichmässig in $t \in \mathbb{R}$, gleichmässig in $k \in \mathbb{N}$, und nimm an,

$$f_k \rightarrow f \quad \text{lokal gleichmässig auf } \mathbb{R} \times \mathbb{R}^n \quad (k \rightarrow \infty).$$

Zu $u_0, u_{0k} \in \mathbb{R}^n$ mit $u_{0k} \rightarrow u_0$ ($k \rightarrow \infty$) gibt es dann $T = T(u_0) > 0$ und Lösungen $u, u_k \in C^1([0, T]; \mathbb{R}^n)$ von (7.5.1), (7.5.2), bzw.

$$\dot{u}_k(t) = f_k(t, u_k(t)), \quad 0 \leq t \leq T, \quad u_k(0) = u_{0k}, \quad (7.5.10)$$

und $u_k \rightarrow u$ in $C^1([0, T]; \mathbb{R}^n)$ für $k \rightarrow \infty$.

Beweis. Nach Annahme bezüglich f, f_k können die Zahlen r_0, T_0, L in (7.5.7) gleichmässig für f und alle f_k gewählt werden. Wie in Satz 7.5.1 erhalten wir dann eindeutige lokale Lösungen $u, u_k \in C^1([0, T]; \mathbb{R}^n)$ von (7.5.1), (7.5.2), bzw. (7.5.10) auf einem von k unabhängigen Intervall $[0, T]$, wobei

$$u(t), u_k(t) \in \overline{B_{r_0}(u_0)}, \quad 0 \leq t \leq T, \quad k \in \mathbb{N}.$$

Für $0 \leq t \leq T$ schätze ab:

$$\begin{aligned} |f_k(t, u_k(t)) - f(t, u(t))| &\leq |f_k(t, u_k(t)) - f_k(t, u(t))| + |f_k(t, u(t)) - f(t, u(t))| \\ &\leq L|u_k(t) - u(t)| + I_k, \end{aligned} \quad (7.5.11)$$

mit Fehlerschranke

$$I_k := \sup_{\substack{0 \leq t \leq T \\ v \in \overline{B_{r_0}(u_0)}}} |f_k(t, v) - f(t, v)| \rightarrow 0 \quad (k \rightarrow \infty).$$

Es folgt:

$$\begin{aligned} |u_k(t) - u(t)| &= \left| u_{0k} + \int_0^t f_k(s, u_k(s)) ds - \left(u_0 + \int_0^t f(s, u(s)) ds \right) \right| \\ &\leq |u_{0k} - u_0| + \int_0^t \underbrace{|f_k(s, u_k(s)) - f(s, u(s))|}_{\leq L\|u_k - u\|_{C^0} + I_k} ds \\ &\leq |u_{0k} - u_0| + LT\|u_k - u\|_{C^0} + TI_k, \end{aligned}$$

gleichmässig in $0 \leq t \leq T$. Für $T \leq \frac{1}{2L}$ erhalten wir somit

$$\frac{1}{2}\|u_k - u\|_{C^0([0, T])} \leq |u_{0k} - u_0| + TI_k,$$

und es folgt

$$\|u_k - u\|_{C^0([0, T])} \rightarrow 0 \quad (k \rightarrow \infty).$$

Mit (7.5.1), (7.5.2), bzw. (7.5.10) und (7.5.11) folgt zudem

$$\begin{aligned}\|\dot{u}_k - \dot{u}\|_{C^0([0,T])} &= \|f_k(t, u_k(t)) - f(t, u(t))\|_{C^0([0,T])} \\ &\leq L\|u_k - u\|_{C^0([0,T])} + I_k \rightarrow 0 \quad (k \rightarrow \infty).\end{aligned}$$

□

Beispiel 7.5.7. Mit Satz 7.5.3 erhalten wir insbesondere auch die stetige Abhängigkeit der Lösung des Anfangswertproblems für (7.5.1) von der Anfangszeit. Beachte, dass wir analog zu Satz 7.5.1 das Anfangswertproblem auch für negative Zeiten lösen können. Für $t_k \rightarrow 0$, $u_{0k} \rightarrow u_0$ betrachte Lösungen v_k des Anfangswertproblems

$$\dot{v}_k(t) = f(t, v_k(t)), \quad v_k(t_k) = u_{0k}.$$

Dann sind die Funktionen $u_k(t) = v_k(t + t_k)$, $k \in \mathbb{N}$, Lösungen des Anfangswertproblems

$$\dot{u}_k(t) = f_k(t, u_k(t)), \quad u_k(0) = u_{0k},$$

wobei

$$f_k(t, y) = f(t + t_k, y) \rightarrow f(t, y) \text{ lokal gleichmässig in } t \text{ und } y.$$

Satz 7.5.3 ergibt Konvergenz $u_k \rightarrow u$ in $C^1([-T_0, T_0], \mathbb{R}^n)$ für geeignetes $T_0 = T(u_0) > 0$, wobei $u \in C^1([-T_0, T_0], \mathbb{R}^n)$ die Lösung ist von (7.5.1), (7.5.2). Zu vorgegebenem $0 < T < T_0$ wähle $k_0 \in \mathbb{N}$ so, dass $T + |t_k| \leq T_0$ ($k \geq k_0$). Für $k \geq k_0$, $-T \leq t \leq T$ erhalten wir dann die gleichmässige Abschätzung

$$\begin{aligned}|v_k(t) - u(t)| &\leq |u_k(t - t_k) - u_k(t)| + |u_k(t) - u(t)| \\ &\leq |t_k| \|u_k\|_{C^1([-T_0, T_0])} + \|u_k - u\|_{C^1([-T_0, T_0])} \rightarrow 0 \quad (k \rightarrow \infty);\end{aligned}$$

das heisst, $v_k \rightarrow u \in C^0([-T, T], \mathbb{R}^n)$. Mit (7.5.1), (7.5.7) sehen wir nun, dass auch $\dot{v}_k \rightarrow \dot{u} \in C^0([-T, T], \mathbb{R}^n)$, also $v_k \rightarrow u \in C^1([-T, T], \mathbb{R}^n)$ mit $k \rightarrow \infty$ für jedes $T < T_0$.

7.5.4 Globale Fortsetzbarkeit

Was kann man über den Verlauf der Lösungen “im Grossen” sagen?

Satz 7.5.4. Sei $f: \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ wie in Satz 7.5.1, 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 (7.5.1), (7.5.2) gemäss Satz 7.5.1. 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 (7.5.1), (7.5.2), und entweder gilt

$$T_{max} = \infty,$$

oder

$$|u_{max}(t)| \rightarrow \infty \quad (t \rightarrow T_{max}).$$

Beweis. Setze

$$T_{max} = \sup\{T; \exists u \in C^1([0, T]; \mathbb{R}^n) \text{ mit (7.5.1), (7.5.2)}\}.$$

Wegen der Eindeutigkeitsaussage in Satz 7.5.1 stimmen je zwei Lösungen $u^{(i)} \in C^1([0, T_i]; \mathbb{R}^n)$, $i = 1, 2$, von (7.5.1), (7.5.2) 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 (7.5.1), (7.5.2) ist auf einem Intervall $[0, T]$ mit $T \geq t$, und u_{max} löst (7.5.1), (7.5.2) auf $[0, T_{max}[$.

Nimm an, $T_{max} < \infty$. Falls wir widerspruchswise annehmen, dass

$$\liminf_{t \uparrow T_{max}} |u_{max}(t)| < \infty,$$

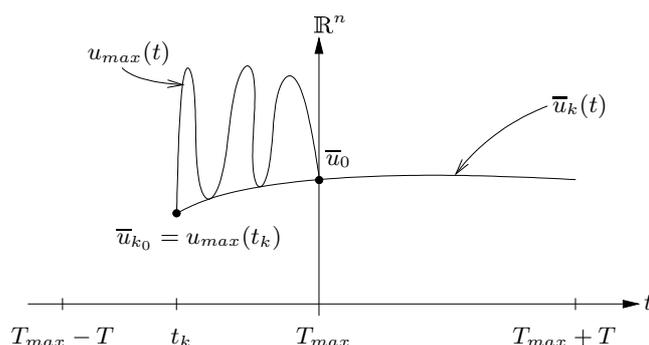
so gibt es $(t_k)_{k \in \mathbb{N}}$ mit $t_k \rightarrow T_{max}$, so dass

$$\bar{u}_{0k} = u_{max}(t_k) \rightarrow \bar{u}_0 \quad (k \rightarrow \infty)$$

für ein $\bar{u}_0 \in \mathbb{R}^n$. Wählen wir im Beweis von Satz 7.5.1 die Konstanten $r_0 > 0$, $T_0 > 0$, $L \in \mathbb{R}$ so, dass (7.5.7) gilt für alle $v, w \in B_{r_0}(\bar{u}_0)$, $t \in \mathbb{R}$ mit $|t - T_{max}| \leq T_0$, so liefert der Beweis ein von k unabhängiges $\bar{T} > 0$ und Lösungen $\bar{u}_k \in C^1([t_k, T_{max} + \bar{T}])$ des Anfangswertproblems (7.5.1) mit $\bar{u}_k(t_k) = \bar{u}_{0k}$, $k \geq k_0$, sofern $k_0 \in \mathbb{N}$ so gewählt ist, dass

$$t_k \geq T_{max} - \bar{T}, \quad \bar{u}_{0k} \in B_{r_0/2}(\bar{u}_0) \text{ für alle } k \geq k_0.$$

Wegen der Eindeutigkeit der Lösung u des Anfangswertproblems (7.5.1) mit Anfangswert $u(t_k) = \bar{u}_{0k}$ stimmt \bar{u}_k für $k \geq k_0$ auf $[t_k, T_{max}[$ überein mit u_{max} . Wir können daher u_{max} durch \bar{u}_{k_0} auf das Intervall $[0, T_{max} + \bar{T}]$ fortsetzen, im Widerspruch zur angenommenen Maximalität von T_{max} . \square



Bemerkung 7.5.2. In jedem Fall ist gemäss Satz 7.5.4 das maximale Existenzintervall $[0, T_{max}[$ der Lösung u von (7.5.1), (7.5.2) rechtsseitig offen.

Beispiel 7.5.8. 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 (7.5.1), (7.5.2) 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

$$\frac{1}{2} \frac{d}{dt} (1 + |u(t)|^2) = (u \cdot \frac{d}{dt} u)(t) = u(t) \cdot f(t, u(t)) \leq C(1 + |u(t)|^2);$$

das heisst,

$$\frac{d}{dt} (\log(1 + |u(t)|^2) - 2Ct) \leq 0, \quad 0 \leq t < T_{max}.$$

Korollar 6.2.1 ergibt nun für alle $0 \leq t < T_{max}$ die Abschätzung

$$\log(1 + |u(t)|^2) \leq 2Ct + \log(1 + |u_0|^2);$$

insbesondere erhalten wir die gleichmässige Abschätzung

$$1 + |u(t)|^2 \leq (1 + |u_0|^2) e^{2CT_{max}}$$

im Widerspruch zur erwarteten Divergenz

$$|u(t)| \rightarrow \infty \quad (t \uparrow T_{max})$$

gemäss Satz 7.5.4. □

ii) Wählen wir insbesondere $b = 0$, $u_0 = e_i$, $1 \leq i \leq n$, so erhalten wir mit Beispiel i) nun auch für variable Koeffizienten $A \in C^0(\mathbb{R}; \mathbb{R}^{n \times n})$ eine global definierte *Fundamentallösung* $\Phi = \Phi(t, t_0)$ der Gleichung (6.6.4) mit der Eigenschaft, dass

$$\frac{d}{dt} \Phi(t, t_0) = A(t) \Phi(t, t_0), \quad \Phi(t, t_0)|_{t=t_0} = id$$

für alle $t \in \mathbb{R}$ und jedes vorgegebene $t_0 \in \mathbb{R}$. Wegen der Eindeutigkeit der Lösung des Anfangswertproblems folgt zusätzlich die Identität

$$\forall s, t, t_0 \in \mathbb{R} : \Phi(t, t_0) = \Phi(t, s) \Phi(s, t_0), \quad (7.5.12)$$

denn für festes s und beliebig vorgegebenes $t_0 \in \mathbb{R}$ ist sowohl die Funktion $U_1(t) = \Phi(t, t_0)$ als auch die Funktion $U_2(t) = \Phi(t, s) \Phi(s, t_0)$ eine Lösung von

$$\frac{d}{dt} U(t) = A(t) U(t), \quad U(t)|_{t=s} = \Phi(s, t_0).$$

Wie in Beispiel 7.1.5.vi) erhalten wir dann mit (7.5.12) zu vorgegebenem $b \in C^0(\mathbb{R}; \mathbb{R}^n)$ eine partikuläre Lösung der inhomogenen Gleichung

$$\frac{d}{dt} u = Au + b,$$

indem wir setzen

$$u_{part}(t) = \int_0^t \Phi(t, s) b(s) ds = \Phi(t, 0) \int_0^t \Phi(0, s) b(s) ds.$$

iii) *Lorenz-Attraktor*. Sei $f = f(u): \mathbb{R}^3 \rightarrow \mathbb{R}^3$ gegeben durch

$$f(u) = \begin{pmatrix} a(u_2 - u_1) \\ bu_1 - u_2 - u_1u_3 \\ u_1u_2 - cu_3 \end{pmatrix}$$

für $u = (u_1, u_2, u_3) \in \mathbb{R}^3$, mit Konstanten $a, b, c \in \mathbb{R}$. Dann gilt

$$u \cdot f(u) = (a + b)u_1u_2 - (au_1^2 + u_2^2 + cu_3^2) \leq C|u|^2, \quad \forall u \in \mathbb{R}^3.$$

Also sind wie in Beispiel i) die Lösungen auf ganz \mathbb{R} fortsetzbar.

Für $a \neq 0 \neq c$ gibt es zudem genau drei Lösungen $u = (u_1, u_2, u_3)$ der Gleichung $f(u) = 0$, nämlich $u_1 = u_2 = u_3 = 0$ oder $u_1 = u_2 = \pm\sqrt{c(b-1)}$.

Speziell für die Wahl

$$a = 10, \quad b = 28, \quad c = 8/3$$

streben alle Lösungen für $t \rightarrow \infty$ hin zu einem kompakten “Attraktor” K , der diese Gleichgewichtslösungen enthält, wobei das Langzeitverhalten der Bahnen $u(t; u_0)$ sehr empfindlich auf kleinste Variationen des Startwerts u_0 reagiert. Da nach Satz 7.5.1 die Lösungen auf jedem kompakten Zeitintervall stetig vom Anfangswert abhängen, spricht man von “deterministischem Chaos”.

Dieses System entstand als einfaches Modell des globalen Klimageschehens; die Sensitivität bzgl. der Daten führte zur Bezeichnung “Schmetterlingseffekt”.

Unter dem Stichwort “Lorenz-Attraktor” findet man bei Wikipedia eine Java-Animation im web.

Kapitel 8

Differentialrechnung im \mathbb{R}^n

8.1 Partielle Ableitungen und Differential

Wie kann man die Konzepte der Differentialrechnung in einer reellen Variablen auf Funktionen $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ erweitern?

Beispiel 8.1.1. Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y) = xe^y, \quad x, y \in \mathbb{R},$$

und sei $(x_0, y_0) \in \mathbb{R}^2$. Fassen wir $y \in \mathbb{R}$ als Parameter einer Schar von Funktionen $f(\cdot, y): \mathbb{R} \rightarrow \mathbb{R}$ auf, so können wir f für festes $y = y_0 \in \mathbb{R}$ "partiell" nach x differenzieren und erhalten so die "partielle Ableitung"

$$\begin{aligned} f_x(x_0, y_0) &= \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0, x \neq x_0} \frac{xe^{y_0} - x_0e^{y_0}}{x - x_0} = e^{y_0}; \end{aligned}$$

ebenso

$$f_y(x_0, y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = x_0e^{y_0}.$$

Sei $\Omega \subset \mathbb{R}^n$ offen, $x_0 = (x_0^i)_{1 \leq i \leq n} \in \Omega$, $e_i = (0, \dots, \underbrace{0, 1, 0}_{i\text{-te Stelle}}, \dots, 0)$, $1 \leq i \leq n$.

Allgemein definieren wir:

Definition 8.1.1. Die Funktion $f: \Omega \rightarrow \mathbb{R}$ heißt an der Stelle x_0 in Richtung e_i (bzw. nach x^i) **partiell differenzierbar**, falls der Limes

$$\begin{aligned} \frac{\partial f}{\partial x^i}(x_0) &= f_{x^i}(x_0) = \lim_{h \rightarrow 0, h \neq 0} \frac{f(x_0 + he_i) - f(x_0)}{h} \\ &= \lim_{h \rightarrow 0, h \neq 0} \frac{f(x_0^1, \dots, x_0^i + h, \dots, x_0^n) - f(x_0^1, \dots, x_0^i, \dots, x_0^n)}{h} \end{aligned}$$

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

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

erhalten wir sofort

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{|f(x) - (f(x_0) + f'(x_0)(x - x_0))|}{x - x_0} = 0.$$

Gilt eine vergleichbare Approximationseigenschaft auch für $n > 1$?

Beispiel 8.1.2. i) Sei $f(x, y) = xe^y$ wie in Beispiel 8.1.1 und sei $(x_0, y_0) \in \mathbb{R}^2$ gegeben.

Für $(x, y) \in \mathbb{R}^2$ erhalten wir mit Satz 6.2.1

$$\begin{aligned} f(x, y) - f(x_0, y_0) &= f(x, y) - f(x_0, y) + f(x_0, y) - f(x_0, y_0) \\ &= \frac{\partial f}{\partial x}(\xi(y), y)(x - x_0) + \frac{\partial f}{\partial y}(x_0, \eta)(y - y_0) \\ &= \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0) + R(x, y) \end{aligned}$$

mit geeigneten Zwischenstellen $\xi = \xi(y)$, η und Restterm

$$R(x, y) = \left(\frac{\partial f}{\partial x}(\xi(y), y) - \frac{\partial f}{\partial x}(x_0, y_0) \right) (x - x_0) + \left(\frac{\partial f}{\partial y}(x_0, \eta) - \frac{\partial f}{\partial y}(x_0, y_0) \right) (y - y_0).$$

Wegen der Stetigkeit der Funktionen

$$\frac{\partial f}{\partial x}(x, y) = e^y, \quad \frac{\partial f}{\partial y}(x, y) = xe^y$$

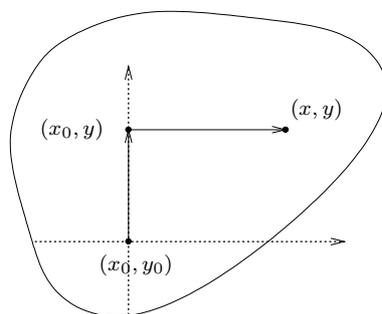
können wir den "Fehler" $R(x, y)$ leicht abschätzen

$$\frac{|R(x, y)|}{|x - x_0| + |y - y_0|} \leq \sup_{\substack{|\xi - x_0| < |x - x_0| \\ |\eta - y_0| < |y - y_0|}} \left(|e^y - e^{y_0}| + |x_0| |e^\eta - e^{y_0}| \right) \rightarrow 0$$

für $(x, y) \rightarrow (x_0, y_0)$, $(x, y) \neq (x_0, y_0)$; das heisst, es gilt

$$\frac{f(x, y) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) - \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)}{|x - x_0| + |y - y_0|} \rightarrow 0 \quad (8.1.1)$$

für $(x, y) \rightarrow (x_0, y_0)$, $(x, y) \neq (x_0, y_0)$.



ii) Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 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

$$f(x, x) = 1 \not\rightarrow f(0, 0) = 0 \quad (x \rightarrow 0, x \neq 0);$$

die Funktion f ist also bei $(x_0, y_0) = (0, 0)$ noch nicht einmal stetig; schon gar nicht kann man die Approximationseigenschaft (8.1.1) erwarten.

Definition 8.1.2. Die Funktion $f: \Omega \rightarrow \mathbb{R}$ heißt an der Stelle $x_0 \in \Omega$ **differenzierbar**, falls eine lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}$ existiert mit

$$\lim_{x \rightarrow x_0, x \neq x_0} \frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0.$$

In diesem Fall heißt $df(x_0) := A$ das **Differential** von f an der Stelle x_0 .

Bemerkung 8.1.1. i) Falls f differenzierbar ist an der Stelle x_0 , so gilt für jede Folge $x_k \rightarrow x_0$ offenbar $f(x_k) \rightarrow f(x_0)$; also ist f insbesondere stetig an der Stelle x_0 . Weiter existieren sämtliche partiellen Ableitungen $\frac{\partial f}{\partial x^i}(x_0) = df(x_0)e_i$, $1 \leq i \leq n$; die Umkehrung gilt aber nicht (vgl. Beispiel 8.1.2).

ii) Ist f an der Stelle x_0 differenzierbar, so hat $df(x_0)$ die Darstellung

$$df(x_0) = \left(\frac{\partial f}{\partial x^1}(x_0), \dots, \frac{\partial f}{\partial x^n}(x_0) \right).$$

Notation: Es zeigt sich nun, wie vorteilhaft es ist, Zeilen- und Spaltenvektoren zu unterscheiden. Schreiben wir nämlich $x = (x^i)_{1 \leq i \leq n} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix}$ für einen Vektor $x \in \mathbb{R}^n$ und $A = (A_1, \dots, A_n)$ für die Darstellung einer linearen Abbildung

$A: \mathbb{R}^n \rightarrow \mathbb{R}$ bezüglich der Standardbasis, so ist

$$A(x - x_0) = \sum_{i=1}^n A_i(x^i - x_0^i) = (A_1, \dots, A_n) \begin{pmatrix} x^1 - x_0^1 \\ \vdots \\ x^n - x_0^n \end{pmatrix}$$

interpretierbar als Matrixmultiplikation des **co-Vektors** $A = (A_1, \dots, A_n)$ mit dem **Vektor** $x - x_0$.

Diese Schreibweise lädt ein zur **Einsteinschen Summenkonvention**: Über doppelt auftretende obere und untere Indizes wird stillschweigend summiert.

Beispiel 8.1.3. 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 $df(x_0) = A$, denn es gilt

$$\forall x, x_0 \in \mathbb{R}^n : f(x) - f(x_0) - A(x - x_0) = 0.$$

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

$$dx^i|_{x=x_0} = (0, \dots, 0, \overbrace{1}^{i\text{-te Stelle}}, 0, \dots, 0), \quad 1 \leq i \leq n.$$

Die Differentiale dx^1, \dots, dx^n bilden also an jeder Stelle $x_0 \in \mathbb{R}^n$ eine Basis des Raumes

$$L(\mathbb{R}^n; \mathbb{R}) = \{A: \mathbb{R}^n \rightarrow \mathbb{R}; A \text{ linear}\},$$

wobei wir $A \in L(\mathbb{R}^n; \mathbb{R})$ mit der Darstellung $A = (A_1, \dots, A_n)$ bzgl. der Standardbasis e_1, \dots, e_n des \mathbb{R}^n identifizieren, und mit $A_i = Ae_i$, $1 \leq i \leq n$.

Da offenbar gilt

$$dx^i e_j = \frac{\partial x^i}{\partial x^j} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

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

$$df(x_0) = \frac{df}{dx}(x_0)dx = f'(x_0)dx;$$

das heisst, $f'(x_0)$ ist die Darstellung von $df(x_0)$ bzgl. der Basis dx von $L(\mathbb{R}; \mathbb{R})$.

iv) Die Funktion $f(x, y) = xe^y: \mathbb{R}^2 \rightarrow \mathbb{R}$ ist an jeder Stelle $(x_0, y_0) \in \mathbb{R}^2$ gemäss Beispiel 8.1.2.i) differenzierbar, und es gilt

$$df(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right) = (e^{y_0}, x_0 e^{y_0}).$$

(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 8.1.2.ii) ist an der Stelle $(x_0, y_0) = (0, 0)$ nicht stetig, nach i) also auch nicht differenzierbar.

8.1.1 Funktionen der Klasse C^1

Die Funktionen in den Beispielen 8.1.2.i) und ii) sind beide partiell differenzierbar, aber nur die Funktion in Beispiel 8.1.2.i) ist auch differenzierbar. Was macht den Unterschied aus?

Definition 8.1.3. Die Funktion $f: \Omega \rightarrow \mathbb{R}$ 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 $x \mapsto \frac{\partial f}{\partial x^i}(x)$, $1 \leq i \leq n$, auf Ω stetig sind.

Beispiel 8.1.4. Die Funktion $f(x, y) = x e^y: \mathbb{R}^2 \rightarrow \mathbb{R}$ aus Beispiel 8.1.2.i) ist von der Klasse C^1 , die Funktion aus Beispiel 8.1.2.ii) nicht.

Satz 8.1.1. 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 8.1.2.i). Für $x = (x^i)_{1 \leq i \leq n}$ schätze ab mit Satz 6.2.1

$$\begin{aligned} & f(x) - f(x_0) - df(x_0)(x - x_0) \\ &= \sum_{i=1}^n \left(f(x^1, \dots, x^i, x_0^{i+1}, \dots, x_0^n) - f(x^1, \dots, x^i, \dots, x_0^n) - \frac{\partial f}{\partial x^i}(x_0)(x^i - x_0^i) \right) \\ &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x^i}(x^1, \dots, \xi^i, x_0^{i+1}, \dots, x_0^n) - \frac{\partial f}{\partial x^i}(x_0) \right) (x^i - x_0^i) \end{aligned}$$

für geeignete Punkte ξ^i zwischen x_0^i und x^i , $1 \leq i \leq n$. Wir erhalten somit für $x \neq x_0$ die Abschätzung

$$\frac{f(x) - f(x_0) - df(x_0)(x - x_0)}{|x - x_0|} \leq n \sup_{|z^i - x_0^i| < |x^i - x_0^i|} \left| \frac{\partial f}{\partial x^i}(z) - \frac{\partial f}{\partial x^i}(x_0) \right|,$$

und die rechte Seite strebt wegen der Stetigkeit der partiellen Ableitungen mit $x \rightarrow x_0$ gegen 0. Das heisst, f ist an der Stelle $x_0 \in \Omega$ differenzierbar. \square

Beispiel 8.1.5. 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 $x = (x^i)_{1 \leq i \leq n} \in \mathbb{R}^n$ setzt

$$x^\alpha := \prod_{i=1}^n (x^i)^{\alpha_i}.$$

Somit kann man ein Polynom $p: \mathbb{R}^n \rightarrow \mathbb{R}$ vom Grad N in der Form schreiben

$$p(x) = \sum_{|\alpha| \leq N} a_\alpha x^\alpha, \quad x \in \mathbb{R}^n,$$

wobei $|\alpha| = \alpha_1 + \dots + \alpha_n$.

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 analog zu Abschnitt 6.4 für beschränktes, offenes $\Omega \subset \mathbb{R}^n$ den Raum

$$C^1(\overline{\Omega}) = \{f \in C^1(\Omega); f \text{ und } \frac{\partial f}{\partial x^i} \text{ sind stetig auf } \overline{\Omega} \text{ ergänzbar, } 1 \leq i \leq n\}$$

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 6.4.3 zeigt man nun die Vollständigkeit von $C^1(\overline{\Omega})$.

Satz 8.1.2. *Sei $\Omega \subset \mathbb{R}^n$ beschränkt. Dann ist der Raum $(C^1(\overline{\Omega}), \|\cdot\|_{C^1(\overline{\Omega})})$ ein Banachraum.*

Beweis. Offenbar definiert die $\|\cdot\|_{C^1}$ -Norm eine Norm. Sei $(f_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in $C^1(\overline{\Omega})$. Dann sind $(f_k)_{k \in \mathbb{N}}$ sowie $(\frac{\partial f_k}{\partial x^i})_{k \in \mathbb{N}}$ Cauchy-Folgen in $C^0(\overline{\Omega})$, und nach Satz 5.7.4 existieren die gleichmässigen Limites

$$f = \lim_{k \rightarrow \infty} f_k, \quad g_i = \lim_{k \rightarrow \infty} \frac{\partial f_k}{\partial x^i} \in C^0(\overline{\Omega}), \quad 1 \leq i \leq n.$$

Nach Satz 6.4.1 ist f in jeder Richtung x^i partiell differenzierbar mit

$$\frac{\partial f}{\partial x^i} = g_i \in C^0(\overline{\Omega}), \quad 1 \leq i \leq n.$$

Nach Satz 8.1.1 ist $f \in C^1(\overline{\Omega})$ und

$$\|f_k - f\|_{C^1(\overline{\Omega})} = \|f_k - f\|_{C^0(\overline{\Omega})} + \sum_{i=1}^n \left\| \frac{\partial f_k}{\partial x^i} - g_i \right\|_{C^0(\overline{\Omega})} \rightarrow 0 \quad (k \rightarrow \infty).$$

□

8.1.2 Landau-Symbole

Viele Grenzwertbetrachtungen werden in höheren Dimensionen schnell unübersichtlich, wenn man versucht, alle Terme explizit mitzuführen. Suggestive Abkürzungen liefern die **Landau-Symbole**.

Definition 8.1.4. *Für $k \in \mathbb{N}_0$ bezeichnen wir mit $O(s^k)$ für $s \downarrow 0$, beziehungsweise $o(s^k)$ für $s \downarrow 0$ irgendwelche Terme, abhängig von $s > 0$, mit*

$$\limsup_{s \downarrow 0} \frac{|O(s^k)|}{s^k} < \infty,$$

bzw.

$$\lim_{s \downarrow 0} \frac{o(s^k)}{s^k} = 0.$$

Das heisst, für Terme $O(s^k)$ ($s \downarrow 0$) existiert $C > 0$ mit

$$|O(s^k)| \leq C s^k \quad \text{für } s \downarrow 0$$

und

$$\frac{o(s^k)}{s^k} \rightarrow 0 \quad (s \downarrow 0).$$

Beispiel 8.1.6. i) Sei $f : \Omega \rightarrow \mathbb{R}$ differenzierbar an der Stelle $x_0 \in \Omega$. Dann gelten die Abschätzungen

$$f(x) - f(x_0) - df(x_0)(x - x_0) = o(|x - x_0|) \quad (x \rightarrow x_0),$$

bzw.

$$f(x) - f(x_0) = df(x_0)(x - x_0) + o(|x - x_0|) = O(|x - x_0|) \quad (x \rightarrow x_0).$$

ii) Sei $f : \Omega \rightarrow \mathbb{R}$ stetig an der Stelle $x_0 \in \Omega$. Dann gilt

$$f(x) - f(x_0) = o(|x - x_0|^0) = o(1) \quad (x \rightarrow x_0).$$

iii) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$, und sei $x_0 = 1$. Es gilt

$$f(x) - f(x_0) = x^2 - 1 = (x + 1)(x - 1) = O(|x - 1|) \quad (x \rightarrow 1),$$

bzw.

$$\begin{aligned} f(x) - f(x_0) - f'(x_0)(x - x_0) &= x^2 - 1 - 2(x - 1) = (x - 1)^2 \\ &= O(|x - 1|^2) = o(|x - 1|) \quad (x \rightarrow 1). \end{aligned}$$

iv) Sei $g : \Omega \rightarrow \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar. Dann gilt $g(x) \rightarrow g(x_0)$ für $x \rightarrow x_0$, und die Notation $o(|g(x) - g(x_0)|)$ ($x \rightarrow x_0$) ist sinnvoll. Weiter gilt

$$o(|g(x) - g(x_0)|) = o(|x - x_0|) \quad (x \rightarrow x_0),$$

denn für $g(x) \neq g(x_0)$ folgt mit Beispiel i):

$$\frac{o(|g(x) - g(x_0)|)}{|x - x_0|} = \underbrace{\frac{o(|g(x) - g(x_0)|)}{|g(x) - g(x_0)|}}_{\rightarrow 0 \ (x \rightarrow x_0)} \underbrace{\frac{|g(x) - g(x_0)|}{|x - x_0|}}_{\leq C} = o(1) \xrightarrow{(x \rightarrow x_0)} 0.$$

Bemerkung 8.1.2. i) Das obige Beispiel 8.1.6.ii) zeigt, wie wichtig die korrekte Syntax beim Einsatz der Landau-Symbole ist. Ohne die Angabe des betrachteten Grenzprozesses $x \rightarrow x_0$ ist das Landau-Symbol $o(1)$ sinnlos.

ii) Wie Beispiel iii) zeigt, erlauben die Landau-Symbole nur Abschätzungen; die "Identität"

$$O(|x - 1|^2) = o(|x - 1|) \quad (x \rightarrow 1)$$

sagt nur aus, dass

$$\frac{O(|x - 1|^2)}{|x - 1|} \rightarrow 0 \quad (x \rightarrow 1).$$

8.2 Differentiationsregeln

Sei $\Omega \subset \mathbb{R}^n$ offen.

Satz 8.2.1. Seien $f, g: \Omega \rightarrow \mathbb{R}$ 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) $d(f \cdot g)(x_0) = g(x_0)df(x_0) + f(x_0)dg(x_0)$.

Falls zusätzlich $g(x_0) \neq 0$, so ist auch f/g an der Stelle x_0 differenzierbar mit

iii)

$$d(f/g)(x_0) = \frac{g(x_0)df(x_0) - f(x_0)dg(x_0)}{(g(x_0))^2}.$$

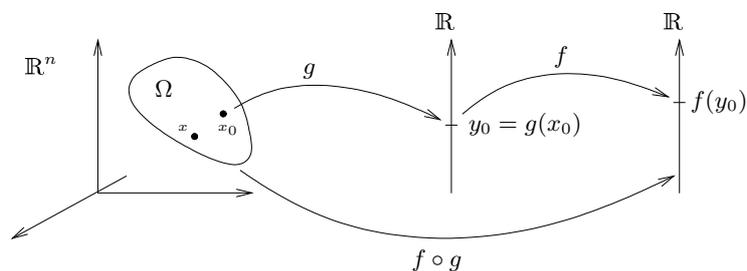
Beweis. Analog zu Satz 6.1.2. Zum Beispiel erhalten wir ii), indem wir entwickeln

$$\begin{aligned} & (fg)(x) - (fg)(x_0) - (g(x_0)df(x_0) + f(x_0)dg(x_0))(x - x_0) \\ &= f(x) \underbrace{(g(x) - g(x_0) - dg(x_0)(x - x_0))}_{=o(|x-x_0|)} \\ &+ g(x_0) \underbrace{(f(x) - f(x_0) - df(x_0)(x - x_0))}_{=o(|x-x_0|)} \\ &+ \underbrace{(f(x) - f(x_0))}_{=O(|x-x_0|)} \underbrace{dg(x_0)(x - x_0)}_{=O(|x-x_0|)} \\ &= o(|x - x_0|). \end{aligned}$$

□

Satz 8.2.2. (Kettenregel, 1. Version) Sei $g: \Omega \rightarrow \mathbb{R}$ an der Stelle $x_0 \in \Omega$ differenzierbar, und sei $f: \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar an der Stelle $g(x_0)$. Dann ist die Funktion $f \circ g: \Omega \rightarrow \mathbb{R}$ 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 $x \rightarrow x_0$, $x \in \Omega$, gilt $g(x) \rightarrow g(x_0)$. Da f bei $g(x_0)$ differenzierbar, folgt mit Beispiel 8.1.6.i) und iv) für $x \rightarrow x_0$ die Abschätzung

$$\begin{aligned} & f(g(x)) - f(g(x_0)) - f'(g(x_0))(g(x) - g(x_0)) \\ &= o(|g(x) - g(x_0)|) = o(|x - x_0|). \end{aligned}$$

Ebenso gilt

$$g(x) - g(x_0) - dg(x_0)(x - x_0) = o(|x - x_0|) \quad (x \rightarrow x_0).$$

Wir erhalten

$$\begin{aligned} (f \circ g)(x) - (f \circ g)(x_0) - f'(g(x_0))dg(x_0)(x - x_0) \\ = f(g(x)) - f(g(x_0)) - f'(g(x_0))(g(x) - g(x_0)) \\ + f'(g(x_0))(g(x) - g(x_0) - dg(x_0)(x - x_0)) \\ = o(|x - x_0|) \quad (x \rightarrow x_0), \end{aligned}$$

und die Behauptung folgt. \square

Beispiel 8.2.1. Sei $h: \mathbb{R}^2 \rightarrow \mathbb{R}$ 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 8.2.2 in der Form

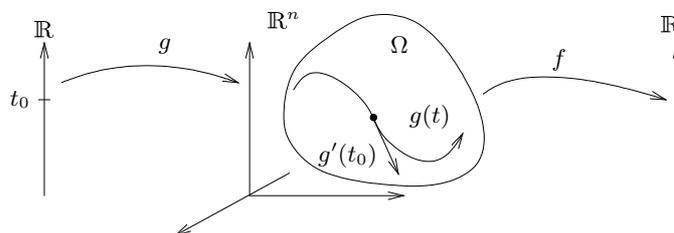
$$dh(x, y) = f'(g(x, y)) \cdot dg(x, y) = e^{xy} \cdot (y, x).$$

Satz 8.2.3. (Kettenregel, 2. Teil) Sei $I \subset \mathbb{R}$ und sei $\Omega \subset \mathbb{R}^n$ offen. Sei weiter $g: I \subset \mathbb{R} \rightarrow \Omega$ an der Stelle $t_0 \in I$ differenzierbar, und sei $f: \Omega \rightarrow \mathbb{R}$ an der Stelle $x_0 = g(t_0)$ differenzierbar. Dann ist die Funktion $f \circ g: I \rightarrow \mathbb{R}$ 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 8.2.1. 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 $df(g(t_0)): \mathbb{R}^n \rightarrow \mathbb{R}$ wirkt. Im zweiten Fall deuten wir $dg(t_0)$ als Differential der vektorwertigen Funktion g , das heisst, als lineare Abbildung $dg(t_0): \mathbb{R} \rightarrow \mathbb{R}^n$, die wir mit der linearen Abbildung $df(g(t_0)): \mathbb{R}^n \rightarrow \mathbb{R}$ verknüpfen.

Beweis. Der Beweis ist vollkommen analog zum Beweis von Satz 8.2.2. \square

Beispiel 8.2.2. i) Sei $f(x, y) = x^2 + y^2$, $(x, y) \in \mathbb{R}^2$, und sei $g: \mathbb{R} \rightarrow \mathbb{R}^2$ die Kurve

$$g(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad t \in \mathbb{R}.$$

Dann gilt

$$\forall t \in \mathbb{R}: (f \circ g)(t) = \cos^2(t) + \sin^2(t) = 1,$$

also

$$\forall t \in \mathbb{R}: \frac{d}{dt}(f \circ g)(t) = 0.$$

Dasselbe Ergebnis erhalten wir auch mit Satz 8.2.3, denn

$$\begin{aligned} \forall t \in \mathbb{R}: \frac{d}{dt}(f \circ g)(t) &= df(g(t)) \frac{dg}{dt}(t) \\ &= (2x, 2y) \Big|_{(x,y)=g(t)} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \\ &= 2(\cos t, \sin t) \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = 0. \end{aligned}$$

ii) Sei $f: \Omega \rightarrow \mathbb{R}$ differenzierbar an der Stelle $x_0 \in \Omega$, und sei $e \in \mathbb{R}^n \setminus \{0\}$. Betrachte die Gerade $g(t) = x_0 + te$, $t \in \mathbb{R}$, durch x_0 mit Richtungsvektor $\frac{dg}{dt}(t_0) = e$, $\forall t_0 \in \mathbb{R}$.

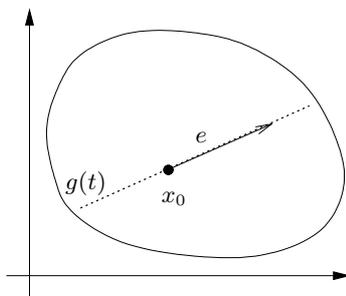
Dann ist die Funktion $f \circ g$ in einer Umgebung von $t_0 = 0$ definiert, und nach Satz 8.2.3 ist $f \circ g$ an der Stelle $t_0 = 0$ differenzierbar mit

$$\frac{d}{dt}(f \circ g)(0) = df(g(0)) \frac{dg}{dt}(0) = df(x_0)e.$$

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 8.1.1.

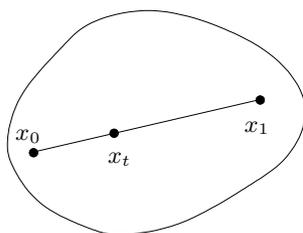


Satz 8.2.4. Sei $f: B_r(0) \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar, und seien $x_0, x_1 \in B_r(0)$. Dann existiert $0 < \vartheta < 1$ mit

$$f(x_1) - f(x_0) = df(x_\vartheta)(x_1 - x_0),$$

wobei

$$x_t = (1-t)x_0 + tx_1, \quad 0 \leq t \leq 1.$$



Beweis. Setze $g(t) = x_t$, $0 \leq t \leq 1$. Dann ist nach Satz 5.3.1 und Satz 8.2.3 die Funktion $f \circ g: [0, 1] \rightarrow \mathbb{R}$ stetig und in $]0, 1[$ differenzierbar. Gemäss Satz 6.2.1 gibt es $\vartheta \in]0, 1[$ mit

$$\begin{aligned} f(x_1) - f(x_0) &= f(g(1)) - f(g(0)) = \frac{d}{dt}(f \circ g)(\vartheta) \\ &= df(g(\vartheta)) \frac{dg}{dt}(\vartheta) = df(x_\vartheta)(x_1 - x_0). \end{aligned}$$

□

Korollar 8.2.1. 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 $\overline{B_r(x_0)} \subset \Omega$. Nach Satz 8.2.4 gilt

$$\forall x, y \in B_r(x_0) : |f(y) - f(x)| \leq L |x - y|,$$

mit $L = \sup\{|df(x)|; x \in B_r(x_0)\} < \infty$.

□

Integrale mit Parametern. Sei $h = h(s, t): \mathbb{R}^2 \rightarrow \mathbb{R}$ 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 lässt sich in diesem Fall die Ableitung \dot{u} bestimmen?

Deute $u = f \circ g$ mit

$$f(x, y) = \int_0^x h(s, y) ds: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(t) = \begin{pmatrix} t \\ t \end{pmatrix}: \mathbb{R} \rightarrow \mathbb{R}^2.$$

Offenbar ist $g \in C^1(\mathbb{R}, \mathbb{R}^2)$, und nach Satz 7.3.4 ist f partiell bzgl. 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 $x > 0$, $y_0 \in \mathbb{R}$ und $y \neq y_0$ gilt gemäss Satz 6.2.1

$$f(x, y) - f(x, y_0) = \int_0^x (h(s, y) - h(s, y_0)) ds = \int_0^x \frac{\partial h}{\partial y}(s, y(s))(y - y_0) ds$$

mit Zwischenstellen $y(s)$ zwischen y_0 und y , $0 \leq s \leq x$. Mit Korollar 7.3.1 folgt

$$\begin{aligned} \left| \frac{f(x, y) - f(x, y_0)}{y - y_0} - \int_0^x \frac{\partial h}{\partial y}(x, y_0) ds \right| &\leq \int_0^x \left| \frac{\partial h}{\partial y}(x, y(s)) - \frac{\partial h}{\partial y}(x, y_0) \right| ds \\ &\leq x \sup_{0 \leq s \leq x, |\eta - y_0| < |y - y_0|} \left| \frac{\partial h}{\partial y}(s, \eta) - \frac{\partial h}{\partial y}(s, y_0) \right| \rightarrow 0 \quad (y \rightarrow y_0), \end{aligned}$$

da $\frac{\partial h}{\partial y}$ auf dem kompakten Streifen $\{(s, y); 0 \leq s \leq x; |y - y_0| \leq 1\}$ gemäss Satz 5.5.2 gleichmässig stetig ist.

Ebenso erhält man die Stetigkeit von $\frac{\partial f}{\partial y}$ aus der Abschätzung

$$\begin{aligned} \left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y_0) \right| &\leq \left| \frac{\partial f}{\partial y}(x, y) - \frac{\partial f}{\partial y}(x_0, y) \right| + \left| \frac{\partial f}{\partial y}(x_0, y) - \frac{\partial f}{\partial y}(x_0, y_0) \right| \\ &= \left| \int_{x_0}^x \frac{\partial h}{\partial y}(s, y) ds \right| + \left| \int_0^{x_0} \left(\frac{\partial h}{\partial y}(s, y) - \frac{\partial h}{\partial y}(s, y_0) \right) ds \right| \\ &\leq \sup_{|s - x_0| \leq 1, |y - y_0| \leq 1} \left| \frac{\partial h}{\partial y}(s, y) \right| \cdot |x - x_0| + x_0 \sup_{0 \leq s \leq x_0} \left| \frac{\partial h}{\partial y}(s, y) - \frac{\partial h}{\partial y}(s, y_0) \right| \\ &\rightarrow 0 \quad (x \rightarrow x_0, y \rightarrow y_0). \end{aligned}$$

□

Satz 8.2.3 ergibt somit $u \in C^1$ mit

$$\begin{aligned} \dot{u}(t) &= \frac{d}{dt}(f \circ g)(t) = \left(\frac{\partial f}{\partial x}(g(t)), \frac{\partial f}{\partial y}(g(t)) \right) \frac{dg}{dt}(t) \\ &= \left(h(t, t), \int_0^t \frac{\partial h}{\partial t}(s, t) ds \right) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = h(t, t) + \int_0^t \frac{\partial h}{\partial t}(s, t) ds. \end{aligned} \tag{8.2.1}$$

Beispiel 8.2.3. 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 7.1.5.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) Die Gleichung (8.2.1) ermöglicht einen neuen Zugang zur Variation-der-Konstanten Formel zur Lösung der inhomogenen linearen Differentialgleichung

$$\dot{u}(t) = A(t)u(t) + b(t) \quad (8.2.2)$$

mit $A \in C^0(\mathbb{R}; \mathbb{R}^{n \times n})$, $b \in C^0(\mathbb{R}; \mathbb{R}^n)$. Nach Beispiel 7.5.8.ii) besitzt das Anfangswertproblem

$$\frac{d\Phi}{dt}(t; s) = A(t)\Phi(t; s), \quad \Phi(s; s) = id \quad (8.2.3)$$

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 $s, t \in \mathbb{R}$.)

Den Ansatz

$$u(t) = \int_0^t \Phi(t; s)b(s) ds, \quad t \in \mathbb{R} \quad (8.2.4)$$

aus Beispiel 7.5.8.ii) für eine partikuläre Lösung von (8.2.2) können wir nun auch ohne Benutzung von (7.5.12) leicht verifizieren. Mit (8.2.1) erhalten wir $u \in C^1(\mathbb{R}; \mathbb{R}^n)$ mit

$$\begin{aligned} \dot{u}(t) &= \Phi(t; t)b(t) + \int_0^t \frac{d\Phi}{dt}(t; s)b(s) ds \\ &\stackrel{(8.2.3)}{=} b(t) + A(t) \int_0^t \Phi(t; s)b(s) ds = A(t)u(t) + b(t), \end{aligned}$$

wie gewünscht.

Ebenso wie die Darstellung (7.1.5) einer partikulären Lösung im Falle $n = 1$ erhält man auch in Dimensionen $n > 1$ die Lösungsformel (8.2.4) aus dem Ansatz $u(t) = \Phi(t; 0)c(t)$ mit variablem $c = c(t)$.

8.3 Differentialformen und Vektorfelder

In Abschnitt 7.5 haben wir bereits Funktionen $v: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $\lambda: \Omega \subset \mathbb{R}^n \rightarrow L(\mathbb{R}^n; \mathbb{R})$ betrachten, welche jedem $x \in \Omega$ eine lineare Abbildung $\lambda(x): T_x\mathbb{R}^n \cong \mathbb{R}^n \rightarrow \mathbb{R}$ zuordnen. Bezüglich der Basis dx^1, \dots, 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 Abbildung

$$\lambda = (\lambda_1, \dots, \lambda_n): \Omega \rightarrow \mathbb{R}^n$$

identifizieren.

Definition 8.3.1. Eine Abbildung $\lambda: \Omega \rightarrow L(\mathbb{R}^n; \mathbb{R})$ heisst eine **Differentialform vom Grad 1** (kurz **1-Form** oder *Pfaffsche Form*).

Beispiel 8.3.1. 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 . Bezüglich der Standardbasis hat λ die Darstellung $\lambda(x, y, z) = (3, 2z, xy)$.

Bemerkung 8.3.1. i) Mit Hilfe des Skalarprodukts $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ kann man ein Vektorfeld $v = (v^i)_{1 \leq i \leq n}: \Omega \rightarrow \mathbb{R}^n$ in eine 1-Form λ verwandeln. Setze dazu für jedes $x \in \Omega$

$$\forall w \in T_x\mathbb{R}^n : \lambda(x)w := \langle v(x), w \rangle_{\mathbb{R}^n} \quad (8.3.1)$$

Bezüglich der Standardbasis gilt $\lambda = (\lambda_i)_{1 \leq i \leq n}$ mit

$$\forall x \in \Omega : \lambda_i(x) = \lambda(x)e_i = \langle v(x), e_i \rangle_{\mathbb{R}^n} = v^i(x). \quad (8.3.2)$$

ii) Umgekehrt kann man via (8.3.1) (z.B. mit dem Riesz'schen Darstellungssatz; siehe Beispiel 8.5.5.i)) auch 1-Formen λ auf Ω in Vektorfelder $v: \Omega \rightarrow \mathbb{R}^n$ umwandeln; vergleiche Beispiel 8.5.5.i).

Speziell für $\lambda = df$ ergibt Bemerkung 8.3.1.ii) die folgende Definition

Definition 8.3.2. Sei $f \in C^1(\Omega)$. Das durch die Gleichung

$$\forall w \in \mathbb{R}^n : \langle \nabla f(x), w \rangle_{\mathbb{R}^n} = df(x)w$$

definierte Vektorfeld $\nabla f: \Omega \rightarrow \mathbb{R}^n$ heisst **Gradientenfeld von f bzgl. $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$** .

Bezüglich der Standardbasis e_1, \dots, e_n des \mathbb{R}^n folgt mit (8.3.2) die Darstellung

$$\forall x \in \Omega : \nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x^1}(x) \\ \vdots \\ \frac{\partial f}{\partial x^n}(x) \end{pmatrix}.$$

Bemerkung 8.3.2. 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

$$\begin{aligned} df(x_0)e &\stackrel{(\text{Def})}{=} \langle \nabla f(x_0), e \rangle_{\mathbb{R}^n} \stackrel{(\text{Cauchy-Schwarz})}{\leq} |\nabla f(x_0)| \\ &= \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{|\nabla f(x_0)|} \right\rangle = df(x_0) \frac{\nabla f(x_0)}{|\nabla f(x_0)|}, \end{aligned}$$

mit Gleichheit für $e = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$. □

Beispiel 8.3.2. 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

$$|\nabla f(x_0, y_0)| = \sqrt{2}, \quad \left(\frac{\nabla f}{|\nabla f|} \right)(x_0, y_0) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

ii) Sei $f \in C^1(\mathbb{R}^2)$ mit

$$\begin{aligned} df(0) \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= df(0)e_2 = \frac{\partial f}{\partial x^2}(0) = 3, \\ df(0) \begin{pmatrix} 1 \\ -1 \end{pmatrix} &= \frac{\partial f}{\partial x^1}(0) - \frac{\partial f}{\partial x^2}(0) = 1. \end{aligned}$$

Es folgt $\frac{\partial f}{\partial x^1}(0) = 4$; das heisst,

$$\nabla f(0) = \begin{pmatrix} 4 \\ 3 \end{pmatrix}.$$

8.4 Wegintegrale

Sei $\Omega \subset \mathbb{R}^n$ offen, $\gamma: [0, 1] \rightarrow \Omega$ ein “Weg” in Ω von der Klasse C^1 (kurz $\gamma \in C^1([0, 1]; \Omega)$) mit “Geschwindigkeitsvektorfeld”

$$\dot{\gamma}(t) = \frac{d\gamma}{dt}(t) = \begin{pmatrix} \frac{d\gamma^1}{dt}(t) \\ \vdots \\ \frac{d\gamma^n}{dt}(t) \end{pmatrix}, \quad 0 \leq t \leq 1.$$

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

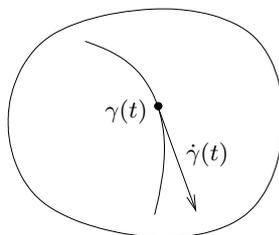
$$t \mapsto \lambda(\gamma(t))\dot{\gamma}(t) = \sum_{i=1}^n \lambda_i(\gamma(t)) \frac{d\gamma^i}{dt}(t)$$

eine stetige Funktion auf $[0, 1]$ definiert.

Definition 8.4.1. *Der Ausdruck*

$$\int_{\gamma} \lambda := \int_0^1 \lambda(\gamma(t))\dot{\gamma}(t) dt$$

heißt **Wegintegral** von λ längs γ .



Bemerkung 8.4.1. i) Anstelle $[0, 1]$ kann man ein beliebiges Intervall $[a, b]$ mit $a < b$ als Parameterbereich betrachten.

ii) Das Wegintegral $\int_{\gamma} \lambda$ ist unabhängig von orientierungserhaltenden Umparametrisierungen von γ . Bei orientierungsumkehrenden Umparametrisierungen $\tilde{\gamma}$ von γ gilt

$$\int_{\tilde{\gamma}} \lambda = - \int_{\gamma} \lambda.$$

Beweis. Sei $\gamma \in C^1([0, 1]; \Omega)$, $\varphi \in C^1([a, b]; [0, 1])$. Offenbar ist der Parameterwechsel $\varphi \in C^1([a, b]; [0, 1])$ orientierungserhaltend, falls $\varphi(a) = 0$, $\varphi(b) = 1$; er ist orientierungsumkehrend, falls $\varphi(a) = 1$, $\varphi(b) = 0$.

Sei zunächst φ orientierungserhaltend mit $\varphi(a) = 0$, $\varphi(b) = 1$. Betrachte

$$\tilde{\gamma} = \gamma \circ \varphi \in C^1([a, b]; \Omega).$$

Für $\lambda \in C^0(\Omega; \mathbb{R}^n)$ gilt:

$$\begin{aligned} \int_{\gamma} \lambda &= \int_0^1 \lambda(\gamma(t)) \dot{\gamma}(t) dt \\ &\stackrel{(t=\varphi(s))}{=} \int_a^b \underbrace{\lambda(\gamma(\varphi(s)))}_{=\tilde{\gamma}(s)} \underbrace{\frac{d\gamma}{dt}(\varphi(s)) \frac{d\varphi}{ds}(s)}_{=\frac{d\tilde{\gamma}}{ds}(s)} ds \\ &= \int_a^b \lambda(\tilde{\gamma}(s)) \frac{d\tilde{\gamma}}{ds}(s) ds = \int_{\tilde{\gamma}} \lambda. \end{aligned}$$

Für den Fall orientierungsumkehrender Umparametrisierungen genügt es, die Abbildung $\varphi(t) = 1 - t$ zu betrachten, mit $\tilde{\gamma}(t) = \gamma(1 - t)$, $0 \leq t \leq 1$. Es gilt:

$$\begin{aligned} \int_{\tilde{\gamma}} \lambda &= \int_0^1 \lambda(\gamma(1-t)) \underbrace{\frac{d\tilde{\gamma}}{dt}(t)}_{=-\dot{\gamma}(1-t)} dt \stackrel{(s=1-t)}{=} \int_1^0 \lambda(\gamma(s)) \dot{\gamma}(s) ds \\ &= - \int_0^1 \lambda(\gamma(s)) \dot{\gamma}(s) ds = - \int_{\gamma} \lambda. \end{aligned}$$

□

Wir bezeichnen den obigen Weg $\tilde{\gamma}(t) = \gamma(1 - t)$, $0 \leq t \leq 1$, als $-\gamma$.

iii) Wege $\gamma_1, \gamma_2 \in C^1([0, 1]; \Omega)$ mit $\gamma_1(1) = \gamma_2(0)$ kann man aneinanderhängen zu einem stetigen, stückweise C^1 -Weg $\gamma = \gamma_1 + \gamma_2: [0, 2] \rightarrow \Omega$ mit

$$t \mapsto \begin{cases} \gamma_1(t), & 0 \leq t \leq 1 \\ \gamma_2(t - 1), & 1 \leq t \leq 2. \end{cases}$$

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 Englisch "piece-wise"), und es gilt

$$\int_{\gamma_1 + \gamma_2} \lambda = \int_{\gamma_1} \lambda + \int_{\gamma_2} \lambda.$$

Beispiel 8.4.1. i) Sei $\gamma \in C^1([0, 2\pi]; \mathbb{R}^2)$ mit

$$\gamma(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad 0 \leq t \leq 2\pi,$$

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

$$\int_{\gamma} \lambda = \int_0^{2\pi} (-\sin t, \cos t) \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = \int_0^{2\pi} (\sin^2(t) + \cos^2(t)) dt = 2\pi.$$

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 8.2.3 gilt

$$df(\gamma(t))\dot{\gamma}(t) = \frac{d}{dt}(f \circ \gamma)(t);$$

das heisst,

$$\int_{\gamma} df = \int_0^1 \frac{d}{dt}(f \circ \gamma)(t) dt = f(\gamma(1)) - f(\gamma(0))$$

hängt nur von Anfangs- und Endpunkt des Weges γ ab.

Beispiel 8.4.1.ii) liefert ein Analogon zu Korollar 6.2.1.i).

Satz 8.4.1. Sei $\Omega \subset \mathbb{R}^n$ offen und (C^1 -)wegzusammenhängend im Sinne von Definition 5.6.2, 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 8.4.1.ii) erhalten wir

$$f(x_1) - f(x_0) = f(\gamma(1)) - f(\gamma(0)) = \int_{\gamma} df = 0.$$

□

Bemerkung 8.4.2. Analog zu Satz 5.6.8 ist eine offene und zusammenhängende Menge $\Omega \subset \mathbb{R}^n$ stets C^1 -wegzusammenhängend. Das Beispiel

$$\Omega = \{(x, \sin(1/x)); 0 < x \leq 1\} \cup \{(0, y); -1 \leq y \leq 1\} \subset \mathbb{R}^2$$

zeigt jedoch, dass dies im Allgemeinen nicht mehr gilt, wenn Ω nicht offen ist.

Wie kann man entscheiden, ob eine Differentialform λ von der Form $\lambda = df$ ist für ein $f \in C^1(\Omega)$?

Satz 8.4.2. Sei $\lambda \in C^0(\Omega; \mathbb{R}^n)$. Es sind äquivalent:

i) $\exists f \in C^1(\Omega): \lambda = df$. (“ λ ist **exakt** mit Potential f .”)

ii) Für je zwei Wege $\gamma_{1,2} \in C_{pw}^1([0, 1]; \Omega)$ 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 $\gamma \in C_{pw}^1([0, 1]; \Omega)$ mit $\gamma(0) = \gamma(1)$ gilt

$$\int_{\gamma} \lambda = 0.$$

Beweis. $i) \Rightarrow ii)$: Dies ist Beispiel 8.4.1.ii).

$ii) \Rightarrow i)$: Fixiere $p_0 \in \Omega$. Setze $f(p_0) = 0$. Für $x \in \Omega$ sei $\gamma \in C_{pw}^1([0, 1]; \Omega)$ 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 $x_0 \in \Omega$, $\gamma_0 \in C_{pw}^1([0, 1]; \Omega)$ 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

$$\gamma(t) = x_0 + t h e_i \in C^1([0, 1]; \Omega)$$

und

$$f(x_0 + h e_i) - f(x_0) = \int_{\gamma_0 + \gamma} \lambda - \int_{\gamma_0} \lambda = \int_{\gamma} \lambda = \int_0^1 \lambda(x_0 + t h e_i) h e_i dt.$$

Da λ stetig ist, existiert

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h e_i) - f(x_0)}{h} = \lim_{h \rightarrow 0} \int_0^1 \lambda(x_0 + t h e_i) e_i dt = \lambda(x_0) e_i;$$

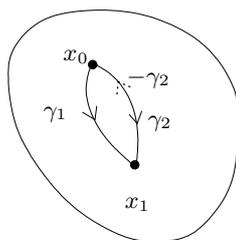
das heisst, f ist auf Ω partiell in Richtung e_i differenzierbar mit

$$\frac{\partial f}{\partial x^i}(x_0) = \lambda(x_0) e_i \in C^0(\Omega), \quad 1 \leq i \leq n,$$

und $f \in C^1(\Omega)$ mit $df = \lambda$. □

ii) \Rightarrow iii): Sei $\gamma \in C_{pw}^1([0, 1]; \Omega)$ geschlossen mit $\gamma(0) = \gamma(1) = x_0$. Wähle zum Vergleich den konstanten Weg $\gamma_1(t) = x_0$, $0 \leq t \leq 1$. Mit ii) folgt

$$\int_{\gamma} \lambda = \int_{\gamma_1} \lambda = 0.$$



iii) \Rightarrow ii): Seien $\gamma_{1,2} \in C_{pw}^1([0, 1]; \Omega)$ mit $\gamma_1(0) = \gamma_2(0)$, $\gamma_1(1) = \gamma_2(1)$. Definiere den Weg $-\gamma_2(t) := \gamma_2(1 - t) \in C_{pw}^1([0, 1]; \Omega)$ 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 $\gamma = \gamma_1 - \gamma_2 = \gamma_1 + (-\gamma_2) \in C_{pw}^1([0, 1]; \Omega)$ ist geschlossen, also

$$0 = \int_{\gamma} \lambda = \int_{\gamma_1} \lambda + \int_{-\gamma_2} \lambda = \int_{\gamma_1} \lambda - \int_{\gamma_2} \lambda.$$

□

Bemerkung 8.4.3. Der Beweis von Satz 8.4.2 liefert offenbar ein Verfahren zur Berechnung des Potentials f der 1-Form $\lambda = df$.

Beispiel 8.4.2. 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 $f(x, 0)$, indem wir setzen

$$f(x, 0) = \int_{\gamma} \lambda, \quad \text{wo } \gamma(t) = (tx, 0), \quad 0 \leq t \leq 1.$$

Da $y = 0$ längs γ , folgt

$$f(x, 0) = \int_0^1 \lambda(\gamma(t)) \dot{\gamma}(t) dt = \int_0^1 (2tx \cdot 0, 2t^2 x^2 \cdot 0) \cdot \begin{pmatrix} x \\ 0 \end{pmatrix} dt = 0.$$

Anschliessend machen wir für beliebiges $(x, y) \in \mathbb{R}^2$ den Ansatz

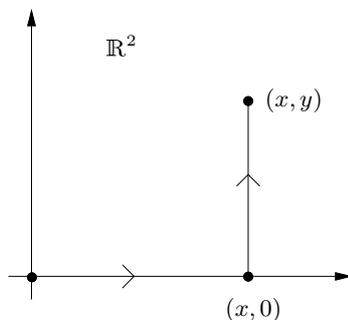
$$f(x, y) = f(x, 0) + \int_{\gamma} \lambda,$$

wobei wir nun den Weg γ wählen mit $\gamma(t) = (x, ty)$, $0 \leq t \leq 1$. Dies ergibt

$$f(x, y) = 0 + \int_0^1 (2x(ty)^2, 2x^2(ty)) \begin{pmatrix} 0 \\ y \end{pmatrix} dt = \int_0^1 2x^2 ty^2 dt = x^2 y^2.$$

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

$$f(x, y) = \int_0^1 2ty^2 dt = y^2, \quad (x, y) \in \mathbb{R}^2.$$

Die Probe versagt jedoch, da $df(x, y) = (0, 2y) \neq \lambda(x, y)$, falls $x \neq 0 \neq y$.

iii) Ebenso besitzt die 1-Form $\lambda(x, y) = -y dx + x dy$ aus Beispiel 8.4.1 kein Potential auf \mathbb{R}^2 .

8.4.1 Konservative Vektorfelder

In Bemerkung 8.3.1 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

$$\forall x \in \Omega, w \in T_x \mathbb{R}^n : \lambda(x)w = \langle v(x), w \rangle_{\mathbb{R}^n},$$

und sei $\gamma \in C^1([0, 1]; \Omega)$.

Definition 8.4.2. Das Wegintegral von v längs γ ist erklärt als

$$\int_{\gamma} v \cdot d\vec{s} = \int_{\gamma} \lambda = \int_0^1 \langle v(\gamma(t)), \dot{\gamma}(t) \rangle_{\mathbb{R}^n} dt.$$

Das Symbol $d\vec{s}$ heisst **gerichtetes Längenelement** (mit der Darstellung $d\vec{s} = \dot{\gamma}(t)dt$).

Definition 8.4.3. Das Vektorfeld $v \in C^0(\Omega; \mathbb{R}^n)$ heisst **konservativ**, falls für jeden "geschlossenen" Weg $\gamma \in C_{pw}^1([0, 1]; \Omega)$ 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 8.4.2 folgt unmittelbar:

Satz 8.4.3. 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 .

Beispiel 8.4.3. Sei $v: \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$ das **Zentralfeld** mit

$$v(x) = x g(|x|^2), \quad x \neq 0,$$

wobei $g \in C^0(]0, \infty[)$.

Behauptung. Das Vektorfeld v ist konservativ, ein Potentialfeld.

Beweis. Sei G die Stammfunktion

$$G(r) = \int_1^r g(s) ds, \quad 0 < r < \infty.$$

Setze

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

Mit der Kettenregel folgt:

$$\frac{\partial f}{\partial x^i}(x) = G'(|x|^2)x^i = g(|x|^2)x^i, \quad 1 \leq i \leq n;$$

das heisst,

$$v = \begin{pmatrix} \frac{\partial f}{\partial x^1} \\ \vdots \\ \frac{\partial f}{\partial x^n} \end{pmatrix} = \nabla f.$$

Zentralfelder sind also konservativ. \square

8.5 Höhere Ableitungen

Sei $\Omega \subset \mathbb{R}^n$ offen, $f \in C^1(\Omega)$.

Definition 8.5.1. Die Funktion f heisst von der Klasse C^2 , $f \in C^2(\Omega)$, falls $\frac{\partial f}{\partial x^i} \in C^1(\Omega)$, $1 \leq i \leq n$.

Beispiel 8.5.1. Sei $f(x, y) = x^2 + 4xy + y$, $(x, y) \in \mathbb{R}^2$. Dann ist $f \in C^2(\mathbb{R}^2)$ mit

$$\frac{\partial f}{\partial x}(x, y) = 2x + 4y, \quad \frac{\partial f}{\partial y}(x, y) = 4x + 1$$

und

$$\frac{\partial^2 f}{(\partial x)^2}(x, y) = 2, \quad \frac{\partial^2 f}{\partial x \partial y}(x, y) = 4 = \frac{\partial^2 f}{\partial y \partial x}(x, y), \quad \frac{\partial^2 f}{(\partial y)^2}(x, y) = 0.$$

Frage: Sind die ‘‘gemischten’’ Ableitungen stets vertauschbar?

Satz 8.5.1. Sei $f \in C^2(\Omega)$. Dann gilt

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial}{\partial x^i} \left(\frac{\partial f}{\partial x^j} \right) = \frac{\partial^2 f}{\partial x^j \partial x^i}, \quad 1 \leq i, j \leq n.$$

Bemerkung 8.5.1. 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. Offenbar gilt $f \in C^2(\mathbb{R}^2 \setminus \{(0, 0)\})$. Aus der Abschätzung

$$|f(x, y) - f(0, 0)| \leq |xy| = o(|(x, y)|) \quad ((x, y) \rightarrow (0, 0))$$

folgt zudem, dass f an der Stelle $(x_0, y_0) = (0, 0)$ differenzierbar ist mit $df(0, 0) = 0$. Weiter gilt

$$\frac{\partial f}{\partial x}(x, 0) = 0, \quad \frac{\partial f}{\partial x}(0, y) = -y, \quad \frac{\partial f}{\partial y}(x, 0) = x, \quad \frac{\partial f}{\partial y}(0, y) = 0;$$

also existieren an der Stelle $(x_0, y_0) = (0, 0)$ auch die zweiten Ableitungen. Jedoch gilt

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1 \neq \frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1.$$

Beweis von Satz 8.5.1. Sei $x_0 \in \Omega$, $i \neq j$. Für genügend kleine $h, k > 0$ 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

$$\gamma(t) = x_0 + tke_j, \quad 0 \leq t \leq 1,$$

und ebenso für den ersten Term, erhalten wir zunächst

$$I = k \int_0^1 \left(\frac{\partial f}{\partial x^j}(x_0 + he_i + tke_j) - \frac{\partial f}{\partial x^j}(x_0 + tke_j) \right) dt.$$

Nach einer analogen Umformung des Integranden ergibt sich schliesslich

$$I = hk \int_0^1 \left(\int_0^1 \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j) ds \right) dt.$$

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

$$I = hk \int_0^1 \left(\int_0^1 \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0 + she_i + tke_j) dt \right) ds.$$

Da $\frac{\partial^2 f}{\partial x^i \partial x^j}$ nach Voraussetzung stetig, ergibt Korollar 7.3.1 für den Term

$$R_{ij} := \int_0^1 \left(\int_0^1 \left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j) - \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) \right) ds \right) dt$$

die Abschätzung

$$|R_{ij}| \leq \sup_{|x-x_0| < h+k} \left| \frac{\partial^2 f}{\partial x^i \partial x^j}(x) - \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) \right| \rightarrow 0 \quad (h, k \rightarrow 0);$$

analog $R_{ji} \rightarrow 0$ ($h, k \rightarrow 0$). Subtraktion der obigen beiden Ausdrücke für I und Division durch hk ergibt somit

$$\begin{aligned} 0 &= \int_0^1 \left(\int_0^1 \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + she_i + tke_j) ds \right) dt \\ &\quad - \int_0^1 \left(\int_0^1 \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0 + she_i + tke_j) dt \right) ds \\ &= \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) + R_{ij} - \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0) - R_{ji}. \end{aligned}$$

Nach Grenzübergang $h, k \rightarrow 0$ folgt

$$\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) = \frac{\partial^2 f}{\partial x^j \partial x^i}(x_0),$$

wie gewünscht. \square

Mit Satz 8.5.1 erhalten wir eine einfach zu handhabende notwendige Bedingung für ein konservatives Vektorfeld.

Korollar 8.5.1. Sei $v = (v^i)_{1 \leq i \leq n} \in C^1(\Omega; \mathbb{R}^n)$ konservativ. Dann gilt

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial v^j}{\partial x^i}, \quad 1 \leq i, j \leq n.$$

Beweis. Nach Voraussetzung ist $v = \nabla f$ für ein $f \in C^2(\Omega)$. Mit Satz 8.5.1 folgt

$$\frac{\partial v^i}{\partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i} = \frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial v^j}{\partial x^i}, \quad 1 \leq i, j \leq n.$$

\square

Beispiel 8.5.2. 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 8.4.2.ii).

Für beliebiges $m \in \mathbb{N}$ definieren wir induktiv analog zu Definition 8.5.1:

Definition 8.5.2. Die Funktion $f \in C^1(\Omega)$ heisst von der Klasse C^m , $f \in C^m(\Omega)$, falls $\frac{\partial f}{\partial x^i} \in C^{m-1}(\Omega)$, $1 \leq i \leq n$.

Bemerkung 8.5.2. Gemäss Satz 8.5.1 sind für $f \in C^m(\Omega)$ partielle Ableitungen der Ordnung $\leq m$ beliebig vertauschbar.

Beispiel 8.5.3. Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ die Funktion

$$f(x, y, z) = e^x \sin(z) + \arctan(y) \cos^2(x).$$

Dann folgt $\frac{\partial^3 f}{\partial x \partial y \partial z} = 0$ auf Ω , da $f = f_1 + f_2$ mit

$$\frac{\partial f_1}{\partial y} = 0, \quad \frac{\partial f_2}{\partial z} = 0$$

auf Ω .

8.5.1 Taylor-Näherung m -ter Ordnung

Sei $f \in C^m(\Omega)$, und seien $x_0, x_1 \in \Omega$ mit

$$x_t = (1-t)x_0 + tx_1 \in \Omega, \quad 0 \leq t \leq 1. \quad (8.5.1)$$

Gemäss Satz 8.2.3 ist die Funktion

$$\varphi(t) = f(x_t) \in C^m([0, 1])$$

mit

$$\begin{aligned} \frac{d\varphi}{dt}(t) &= df(x_t)(x_1 - x_0) = \sum_{i=1}^n \frac{\partial f}{\partial x^i}(x_t)(x_1^i - x_0^i) \\ \frac{d^2\varphi}{dt}(t) &= \sum_{i_1, i_2=1}^n \frac{\partial^2 f}{\partial x^{i_1} \partial x^{i_2}}(x_t)(x_1^{i_1} - x_0^{i_1})(x_1^{i_2} - x_0^{i_2}) \\ &\vdots \\ \frac{d^m\varphi}{dt}(t) &= \sum_{i_1, \dots, i_m=1}^n \frac{\partial^m f}{\partial x^{i_1} \dots \partial x^{i_m}}(x_t) \prod_{j=1}^m (x_1^{i_j} - x_0^{i_j}). \end{aligned}$$

Satz 6.5.2, angewandt auf φ , ergibt

Satz 8.5.2. (Taylorformel) Sei $f \in C^m(\Omega)$, und seien $x_0, x_1 \in \Omega$ mit (8.5.1). Dann gibt es eine Zahl $0 < \vartheta < 1$ mit

$$\begin{aligned} f(x_1) &= f(x_0) + df(x_0)(x_1 - x_0) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x_1^i - x_0^i)(x_1^j - x_0^j) \\ &+ \dots + \frac{1}{m!} \sum_{i_1, \dots, i_m=1}^n \frac{\partial^m f}{\partial x^{i_1} \dots \partial x^{i_m}}(x_\vartheta) \prod_{j=1}^m (x_1^{i_j} - x_0^{i_j}). \end{aligned}$$

Beispiel 8.5.4. Sei $f(x, y) = x^2 + 2xy + y^3$, $(x, y) \in \mathbb{R}^2$, mit

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= 2x + 2y, & \frac{\partial f}{\partial y}(x, y) &= 2x + 3y^2, & \frac{\partial^2 f}{(\partial x)^2}(x, y) &= 2, \\ \frac{\partial^2 f}{\partial x \partial y}(x, y) &= 2, & \frac{\partial^2 f}{(\partial y)^2}(x, y) &= 6y, & \frac{\partial^3 f}{(\partial y)^3}(x, y) &= 6; \end{aligned}$$

alle weiteren Ableitungen verschwinden. Mit Satz 8.5.2 erhalten wir für $m = 4$ und $(x_0, y_0) = (0, 0)$ die Darstellung

$$\begin{aligned} f(x, y) &= \frac{1}{2} \left(\frac{\partial^2 f}{(\partial x)^2}(x, y)x^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(0, 0)xy \right) + \frac{1}{6} \frac{\partial^3 f}{(\partial y)^3}(0, 0)y^3 \\ &= x^2 + 2xy + y^3. \end{aligned}$$

Bemerkung 8.5.3. Mit der Multi-Index Schreibweise

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{(\partial x^1)^{\alpha_1} \dots (\partial x^n)^{\alpha_n}}$$

für $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ analog zu Beispiel 8.1.5.i) kann man den Ausdruck

$$\sum_{i_1, \dots, i_m=1}^n \frac{\partial^m f}{\partial x^{i_1} \dots \partial x^{i_m}}(x_0) \prod_{j=1}^m (x_1^{i_j} - x_0^{i_j}) = \sum_{|\alpha|=m} \partial^\alpha f(x_0)(x_1 - x_0)^\alpha$$

kompakt schreiben.

Wie in Abschnitt 5.5 definieren wir das **Taylor-Polynom m -ter Ordnung**

$$\begin{aligned} T_m f(x, x_0) &= f(x_0) + df(x_0)(x - x_0) + \dots \\ &+ \frac{1}{m!} \sum_{i_1, \dots, i_m=1}^n \frac{\partial^m f}{\partial x^{i_1} \dots \partial x^{i_m}}(x_0) \prod_{j=1}^m (x_1^{i_j} - x_0^{i_j}) \\ &= f(x_0) + df(x_0)(x - x_0) + \dots + \frac{1}{m!} \sum_{|\alpha|=m} \partial^\alpha f(x_0)(x_1 - x_0)^\alpha. \end{aligned}$$

Bemerkung 8.5.4. i) Gemäss Satz 8.5.2 gilt für $f \in C^m(\Omega)$

$$f(x) = T_m f(x; x_0) + r_m f(x, x_0)$$

mit

$$\begin{aligned} |r_m f(x_1; x_0)| &\leq \frac{n^m}{m!} \sup_{0 < \vartheta < 1, |\alpha|=m} |\partial^\alpha f(x_\vartheta) - \partial^\alpha f(x_0)| |x_1 - x_0|^m \\ &= o(|x_1 - x_0|^m) \quad (x_1 \rightarrow x_0). \end{aligned}$$

ii) Falls $f \in C^{m+1}(\Omega)$, so liefert Satz 8.5.2 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

$$\begin{aligned} f(x_1) &= f(x_0) + df(x_0)(x_1 - x_0) + \\ &\frac{1}{2} \sum_{i,j=1}^2 \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x_1^i - x_0^i)(x_1^j - x_0^j) + r_2 f(x_1, x_0) \end{aligned}$$

mit Fehler

$$r_2 f(x_1, x_0) = o(|x_1 - x_0|^2) \quad (x_1 \rightarrow x_0).$$

Analog zu Korollar 6.5.1 für $n = 1$ gilt nun:

Satz 8.5.3. 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”**

$$Hess_f(x_0) = \left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) \right)_{1 \leq i, j \leq n}$$

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 $\xi \in \mathbb{R}^n \setminus \{0\}$, so ist x_0 eine strikte lokale Minimalstelle von f .

Beweis. i) Sei x_0 eine lokale Minimalstelle von f , und nimm widerspruchswise an, $df(x_0) \neq 0$. Setze $e = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$. Dann hat die Funktion

$$\varphi(t) = f(x_0 + te), \quad |t| \ll 1,$$

bei $t = 0$ ein lokales Minimum; jedoch gilt

$$\frac{d\varphi}{dt}(0) = df(x_0)e = |\nabla f(x_0)| > 0$$

im Widerspruch zu Korollar 6.5.1.i).

ii) Da $S^{n-1} = \partial B_1(0; \mathbb{R}^n)$ kompakt, gibt es $\lambda > 0$ mit

$$\forall \xi \in S^{n-1} : Hess_f(x_0)(\xi, \xi) \geq \lambda |\xi|^2.$$

Mit Bemerkung 8.5.4.iii) folgt für $x \neq x_0$ genügend nahe bei x_0 die Ungleichung

$$f(x) - f(x_0) \geq \lambda |x - x_0|^2 + r_2 f(x; x_0) \geq \frac{\lambda}{2} |x - x_0|^2 > 0,$$

wie gewünscht. □

Definition 8.5.3. i) Ein Punkt $x_0 \in \Omega$ mit $df(x_0) = 0$ heisst **kritischer Punkt**, der zugehörige Wert $f(x_0)$ **kritischer Wert** von f .

ii) Die quadratische Form $\xi \mapsto Hess_f(x_0)(\xi, \xi)$ heisst **Hessesche Form**.

Bemerkung 8.5.5. Analog zu Satz 8.5.3.ii) ist ein kritischer Punkt x_0 von f , wo die Hesse-Matrix **negativ definit** ist mit

$$\forall \xi \in \mathbb{R}^n \setminus \{0\} : Hess_f(x_0)(\xi, \xi) < 0$$

eine strikte lokale Maximalstelle von f .

Beispiel 8.5.5. i) (Rieszscher Darstellungssatz) Sei $\lambda \in L(\mathbb{R}^n; \mathbb{R})$. Betrachte die Funktion $f \in C^\infty(\mathbb{R}^n)$ mit

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

Für $x \in \mathbb{R}^n$ mit $|x| \geq 2|\lambda| =: R$ gilt

$$f(x) \geq \frac{1}{2}|x|^2 - |\lambda||x| = \left(\frac{1}{2}|x| - |\lambda|\right)|x| \geq 0 = f(0).$$

Also existiert $x_0 \in \overline{B_R(0)}$ mit

$$f(x_0) = \min_{|x| \leq R} f(x) = \min_{x \in \mathbb{R}^n} f(x).$$

Mit Satz 8.5.3.i) folgt

$$\forall w \in T_x \mathbb{R}^n : df(x_0)w = \langle x_0, w \rangle_{\mathbb{R}^n} - \lambda w = 0.$$

Weiter gilt

$$\forall \xi \in T_x \mathbb{R}^n \setminus \{0\} : Hess_f(x_0)(\xi, \xi) = |\xi|^2 > 0.$$

ii) 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

$$\forall (x, y) \in \mathbb{R}^2 : f(x, y) \geq 0.$$

Andererseits gilt offenbar

$$\forall y \in \mathbb{R} : f(-\beta y, y) = 0;$$

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 $f(p) > 0 > f(q)$.

8.5.2 Der Raum $C^m(\overline{\Omega})$

Falls $\Omega \subset \mathbb{R}^n$ beschränkt, so ist analog zu Satz 6.5.1 und Satz 8.1.2 der Raum

$$C^m(\overline{\Omega}) = \{f \in C^m(\Omega); f \text{ und alle partiellen Ableitungen bis zur Ordnung } m \text{ sind auf } \overline{\Omega} \text{ stetig ergänzbar.}\}$$

ein Banachraum bezüglich der Norm

$$\|f\|_{C^m(\overline{\Omega})} = \|f\|_{C^0(\overline{\Omega})} + \sum_{|\alpha| \leq m} \|\partial^\alpha f\|_{C^0(\overline{\Omega})}.$$

8.5.3 Reell analytische Funktionen

Falls $f \in C^\infty(\Omega)$, so können wir an jedem $x_0 \in \Omega$ die formale Reihe bilden

$$p_f(x; x_0) = \sum_{m=0}^{\infty} \frac{1}{m!} \sum_{|\alpha|=m} \partial^\alpha f(x_0) (x - x_0)^\alpha.$$

Definition 8.5.4. $p_f(\cdot; x_0)$ heisst Taylor-Reihe von f um x_0 .

Satz 8.5.4. Falls $M \in \mathbb{R}$ existiert mit

$$\forall m \in \mathbb{N}_0 \forall |\alpha| = m : |\partial^\alpha f(x_0)| \leq M^m,$$

dann konvergiert $p_f(x; x_0)$ für jedes $x \in \mathbb{R}^n$.

Beweis. Schätze ab

$$|p_f(x; x_0)| \leq \sum_{m=0}^{\infty} a_m |x - x_0|^m$$

mit

$$a_m = \frac{1}{m!} \sum_{|\alpha|=m} |\partial^\alpha f(x_0)| \leq \frac{n^m M^m}{m!}, \quad m \in \mathbb{N}_0.$$

Somit ist $p_f(x; x_0)$ für jedes $x \in \mathbb{R}^n$ konvergent. \square

Beispiel 8.5.6. i) Sei

$$f(x, y, z) = \sum_{k=0}^{\infty} \frac{(x + y + z)^k}{k!}, \quad (x, y, z) \in \mathbb{R}^3.$$

Offenbar gilt an jeder Stelle $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ für jedes $\alpha \in \mathbb{N}_0^n$:

$$|\partial^\alpha f(p_0)| \leq \sum_{k=|\alpha|}^{\infty} \frac{(|x_0| + |y_0| + |z_0|)^{k-|\alpha|}}{(k-|\alpha|)!} \leq M_0.$$

Nach Satz 8.5.4 konvergiert $p_f(\cdot; p_0)$ also auf ganz \mathbb{R}^3 .

ii) Analog für $\exp, \sin, \cos: \mathbb{C} \cong \mathbb{R}^2 \rightarrow \mathbb{C}$.

Definition 8.5.5. Eine Funktion $f \in C^\infty(\Omega)$ heisst **reell analytisch**, falls für alle $x_0 \in \Omega$ in einer Umgebung U von x_0 gilt:

$$f(x) = p_f(x; x_0), \quad x \in U.$$

Beispiel 8.5.7. i) Polynome, rationale Funktionen, die Funktionen $\exp, \sin, \cos: \mathbb{C} \cong \mathbb{R}^2 \rightarrow \mathbb{C}$ sind reell analytisch.

ii) Die Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} e^{-1/(1-|x|^2)}, & |x| < 1 \\ 0, & \text{sonst} \end{cases}$$

ist von der Klasse $C^\infty(\mathbb{R}^n)$ aber nicht reell analytisch; vergleiche Beispiel 6.5.10.

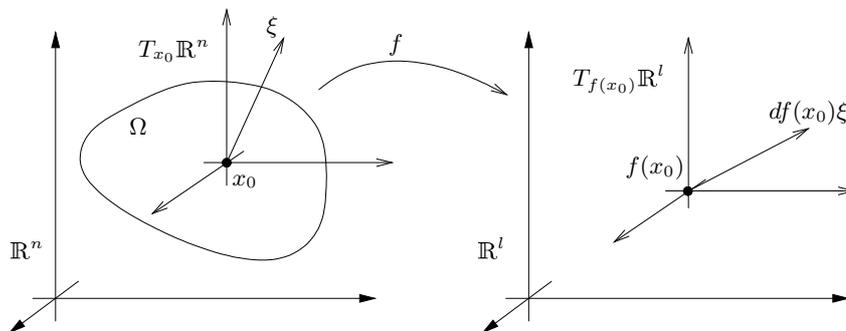
8.6 Vektorwertige Funktionen

Sei $\Omega \subset \mathbb{R}^n$ offen, $f = (f^i)_{1 \leq i \leq l} : \Omega \rightarrow \mathbb{R}^l$.

Definition 8.6.1. i) Die Funktion f heisst an der Stelle $x_0 \in \Omega$ **differenzierbar**, falls jede Komponente f^i , $1 \leq i \leq l$, an der Stelle x_0 differenzierbar ist. Das Differential $df(x_0)$ hat die Gestalt

$$df(x_0) = \begin{pmatrix} df^1(x_0) \\ \vdots \\ df^l(x_0) \end{pmatrix} : T_{x_0}\mathbb{R}^n \cong \mathbb{R}^n \rightarrow T_{f(x_0)}\mathbb{R}^l \cong \mathbb{R}^l.$$

ii) f heisst auf Ω **differenzierbar**, bzw. **von der Klasse C^m** , $m \geq 1$, falls jedes f^i differenzierbar ist, bzw. $f^i \in C^m(\Omega)$, $1 \leq i \leq l$.



Notation: $C^m(\Omega; \mathbb{R}^l) = \{f = (f^i)_{1 \leq i \leq l}; f^i \in C^m(\Omega), 1 \leq i \leq l\}$.

Bemerkung 8.6.1. i) Bzgl. der Standardbasis dx^j , $1 \leq j \leq n$, erhalten wir mit

$$df^i(x_0) = \sum_{j=1}^n \frac{\partial f^i}{\partial x^j}(x_0) dx^j = \left(\frac{\partial f^i}{\partial x^1}(x_0), \dots, \frac{\partial f^i}{\partial x^n}(x_0) \right)$$

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

$$df(x_0) = \left(\frac{\partial f^i}{\partial x_j}(x_0) \right)_{1 \leq i \leq l, 1 \leq j \leq n}$$

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 $A: \mathbb{R}^n \rightarrow \mathbb{R}^l$ existiert mit

$$\lim_{x \rightarrow x_0, 0 \neq x \in \Omega} \frac{f(x) - f(x_0) - A(x - x_0)}{|x - x_0|} = 0,$$

analog zu Definition 8.1.2.

Beispiel 8.6.1. i) Sei $f = f_1: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 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

$$df(x, y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}, \quad (x, y) \in \mathbb{R}^2.$$

ii) Sei $f = f_2: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ 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 8.6.1. Seien $f, g: \Omega \rightarrow \mathbb{R}^l$ an der Stelle $x_0 \in \Omega$ differenzierbar, $\lambda \in \mathbb{R}$. Dann sind die Funktionen λf , $f + g$ sowie das Skalarprodukt von f und g an der Stelle x_0 differenzierbar, und

i) $d(\lambda f)(x_0) = \lambda df(x_0)$,

ii) $d(f + g)(x_0) = df(x_0) + dg(x_0)$,

iii) $d(f \cdot g)(x_0) = f(x_0) \cdot dg(x_0) + g(x_0) \cdot df(x_0)$,

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 8.2.1 lässt sich unmittelbar übertragen. \square

Beispiel 8.6.2. Seien $f(x) = g(x) = x$ mit $df(x) = id$, $x \in \mathbb{R}^n$. Satz 8.6.1 ergibt

$$\forall x \in \mathbb{R}^n, \xi \in \mathbb{R}^n : d(|x|^2)\xi = 2x \cdot \xi.$$

Satz 8.6.2. (Kettenregel, 3. Teil) Seien $g: \Omega \rightarrow \mathbb{R}^l$ an der Stelle $x_0 \in \Omega$ und $f: \mathbb{R}^l \rightarrow \mathbb{R}^m$ an der Stelle $y_0 = g(x_0) \in \mathbb{R}^l$ differenzierbar. Dann ist die Funktion $f \circ g: \Omega \rightarrow \mathbb{R}^m$ 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 8.2.2. \square

Bemerkung 8.6.2. i) Falls $g: \mathbb{R}^n \rightarrow \mathbb{R}^l$, $f: \mathbb{R}^l \rightarrow \mathbb{R}^m$ linear mit

$$f(y) = Ay, \quad y \in \mathbb{R}^l, \quad g(x) = Bx, \quad x \in \mathbb{R}^n,$$

so ist $f \circ g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear mit

$$(f \circ g)(x) = ABx, \quad x \in \mathbb{R}^n.$$

ii) Im folgenden Abschnitt sind die Rollen von f und g meist vertauscht. Das heisst, wir betrachten differenzierbare Funktionen $f: \Omega \rightarrow \mathbb{R}^l$, $g: \mathbb{R}^l \rightarrow \mathbb{R}^m$ und wenden Satz 8.6.2 an auf die Funktion $g \circ f$ mit

$$d(g \circ f)(x_0) = dg(f(x_0)) \cdot df(x_0). \quad (8.6.1)$$

In Koordinaten und mit $y_0 = f(x_0)$ können wir die Formel (8.6.1) schreiben

$$\frac{\partial (g \circ f)^i}{\partial x^k}(x_0) = \sum_{j=1}^l \frac{\partial g^i(y_0)}{\partial y^j} \cdot \frac{\partial f^j(x_0)}{\partial x^k}, \quad 1 \leq i \leq m, \quad 1 \leq k \leq n.$$

Beispiel 8.6.3. Betrachte die Funktionen $g = f_1: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $f = f_2: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ aus Beispiel 8.6.1. Die Funktion

$$(g \circ f)(x, y, z) = \begin{pmatrix} (x^2 + y^2 + z^2)^2 - x^2 y^2 z^2 \\ 2(x^2 + y^2 + z^2)xyz \end{pmatrix}: \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

ist differenzierbar mit

$$\begin{aligned} d(g \circ f)(x, y, z) &= dg(f(x, y, z)) \cdot df(x, y, z) \\ &= \begin{pmatrix} 2(x^2 + y^2 + z^2) & -2xyz \\ 2xyz & 2(x^2 + y^2 + z^2) \end{pmatrix} \begin{pmatrix} 2x & 2y & 2z \\ yz & xz & xy \end{pmatrix} \\ &= \begin{pmatrix} 4x(x^2 + y^2 + z^2) - 2xy^2z^2 & \dots & \dots \\ \dots & \dots & \dots \end{pmatrix}. \end{aligned}$$

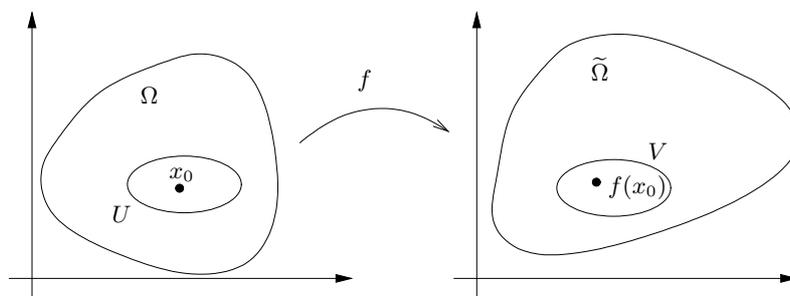
Probe: Differenziere direkt.

8.7 Der Umkehrsatz

Sei $\Omega \subset \mathbb{R}^n$, $f \in C^1(\Omega, \mathbb{R}^n)$ injektiv, $\tilde{\Omega} = f(\Omega)$ das Bild von Ω unter f .

Fragen i) Unter welchen Bedingungen ist die Umkehrabbildung $f^{-1}: \tilde{\Omega} \rightarrow \Omega$ 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 8.6.2 eine notwendige Bedingung: Falls $f \in C^1(\Omega; \mathbb{R}^n)$, $g \in C^1(\tilde{\Omega}; \mathbb{R}^n)$ 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 $df(x_0) \in L(\mathbb{R}^n; \mathbb{R}^n)$ muss also invertierbar sein.

Gemäss dem folgenden Satz ist diese Bedingung auch hinreichend für die lokale Invertierbarkeit von f .

Satz 8.7.1. (Umkehrsatz) Sei $f \in C^1(\Omega; \mathbb{R}^n)$ und sei $df(x_0): \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 $g = (f|_U)^{-1}$; das heisst,

$$\forall x \in U : g(f(x)) = x, \quad \forall y \in V : f(g(y)) = y.$$

Weiter gilt für alle $x \in U$ gemäss (8.6.1) die Beziehung

$$dg(f(x)) = (df(x))^{-1}.$$

Bevor wir diesen Satz beweisen, diskutieren wir die Aussage durch Vergleich mit dem Fall $n = 1$ und anhand von Beispielen.

Beispiel 8.7.1. 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 Existenz von $r > 0$ mit

$$\forall x \in]x_0 - r, x_0 + r[: f'(x) > 0.$$

Nach Satz 6.2.2 ist $f:]x_0 - r, x_0 + r[\rightarrow]c, d[$ invertierbar mit $g = f^{-1} \in C^1(]c, d[)$, und

$$\frac{d(f^{-1})}{dy} \Big|_{y=f(x)} = \left(\frac{df}{dx}(x) \right)^{-1}.$$

Das heisst, das Differential $d(f^{-1})(y)$ 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 8.6.1.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

$$\det(df(x, y)) = 4(x^2 + y^2) > 0,$$

ist f nach Satz 8.7.1 lokal um jeden Punkt $(x_0, y_0) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ 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 $\mathbb{R}^2 \setminus \{(0, 0)\}$. Satz 8.7.1 zeigt jedoch, dass dies lokal auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ möglich ist. Entsprechend kann man **lokal** auf $\mathbb{C} \setminus \{0\}$ eine Quadratwurzelfunktion definieren.

iii) **Polarkoordinaten in \mathbb{R}^2** . Die Abbildung $f:]0, \infty[\times \mathbb{R} \rightarrow \mathbb{R}^2$ mit

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

erfüllt

$$df(r, \phi) = \left(\frac{\partial f}{\partial r}(r, \phi), \frac{\partial f}{\partial \phi}(r, \phi) \right) = \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix},$$

also

$$\det(df(r, \phi)) = r(\cos^2 \phi + \sin^2 \phi) = r > 0.$$

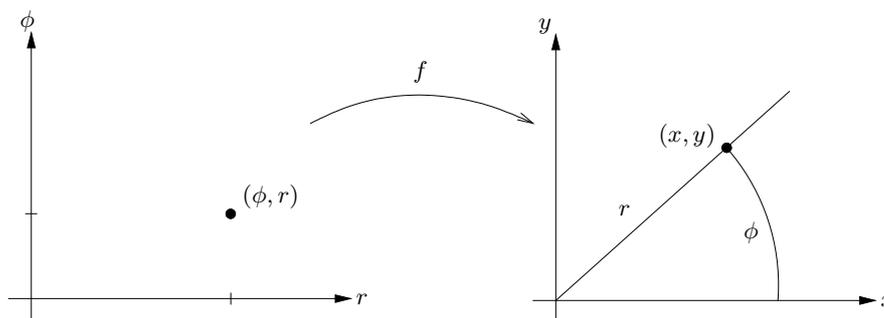
Gemäss Satz 8.7.1 kann man mittels f lokal Polarkoordinaten auf \mathbb{R}^2 einführen.

f ist injektiv zum Beispiel auf $]0, \infty[\times]-\frac{\pi}{2}, \frac{\pi}{2}[=: U$ mit

$$(f|_U)^{-1} = g: \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} r = \sqrt{x^2 + y^2} \\ \phi = \arctan(y/x) \end{pmatrix}.$$

Die **Koordinatenlinien** $\alpha(r) = f(r, \phi)$ (ϕ fest) und $\beta(\phi) = f(r, \phi)$ (r fest) schneiden sich senkrecht, da in jedem Punkt (r, ϕ) gilt

$$\frac{d\alpha}{dr} \cdot \frac{d\beta}{d\phi} = \frac{\partial f}{\partial r} \cdot \frac{\partial f}{\partial \phi} = 0.$$



Beweis von Satz 8.7.1. 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

$$\Phi_y : x \mapsto x + df(0)^{-1}(y - f(x))$$

zu erhalten. Gemäss Satz 7.5.2 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. $\exists \delta = \delta(r) > 0 \forall y \in \overline{B_\delta(0)}$:

$$\Phi_y : \overline{B_r(0)} \rightarrow \overline{B_r(0)}.$$

Beweis von Behauptung 1. Gemäss den Behauptungen 2 und 3 ist für jedes $y \in \overline{B_\delta(0)}$ die Abbildung $\Phi_y : \overline{B_r(0)} \rightarrow \overline{B_r(0)}$ kontrahierend. Mit dem Kontraktionsprinzip, Satz 7.5.2, folgt Behauptung 1 nun unmittelbar. \square

Beweis von Behauptung 3. Zu gegebenem $r > 0$ wähle $0 < \delta < \frac{r}{2\|df(0)^{-1}\|}$. Für $|y| \leq \delta$ schätze ab

$$|\Phi_y(0)| = |df(0)^{-1}y| \leq \|df(0)^{-1}\| \cdot \delta < \frac{r}{2}.$$

Mit Behauptung 2 folgt für $x \in \overline{B_r(0)}$:

$$|\Phi_y(x)| \leq |\Phi_y(x) - \Phi_y(0)| + |\Phi_y(0)| \leq \frac{1}{2} |x| + |\Phi_y(0)| < \frac{r}{2} + \frac{r}{2} = r.$$

\square

Beweis von Behauptung 2. Schreibe für $x, \tilde{x} \in \overline{B_r(0)}$

$$\begin{aligned} \Phi_y(x) - \Phi_y(\tilde{x}) &= (x - \tilde{x}) + df(0)^{-1}(f(\tilde{x}) - f(x)) \\ &= df(0)^{-1}(f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)). \end{aligned}$$

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

$$\frac{f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)}{|\tilde{x} - x|} = \int_0^1 (df(x + t(\tilde{x} - x)) - df(0)) \frac{\tilde{x} - x}{|\tilde{x} - x|} dt$$

und können daher abschätzen

$$\begin{aligned} \frac{|f(\tilde{x}) - f(x) - df(0)(\tilde{x} - x)|}{|\tilde{x} - x|} &\leq \int_0^1 |df(x + t(\tilde{x} - x)) - df(0)| dt \\ &\leq \sup_t \|df(x + t(\tilde{x} - x)) - df(0)\| \\ &\leq \sup_{z \in \overline{B_r(0)}} \|df(z) - df(0)\| \rightarrow 0 \quad (r \rightarrow 0). \end{aligned}$$

Die Behauptung folgt. \square

Bemerkung: Das in Satz 7.5.2 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 näherungsweise 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 $f|_U: U \rightarrow V$ bijektiv. Die Funktion

$$g := (f|_U)^{-1}: V \rightarrow U$$

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

$$\forall y \in V: dg(y) = (df(g(y)))^{-1}.$$

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

$$\begin{aligned} |\tilde{x} - x| &= |\Phi_0(\tilde{x}) - \Phi_0(x) + df(0)^{-1}(f(\tilde{x}) - f(x))| \\ &\leq |\Phi_0(\tilde{x}) - \Phi_0(x)| + \|df(0)^{-1}\| |\tilde{y} - y| \\ &\leq \frac{1}{2} |\tilde{x} - x| + C |\tilde{y} - y|; \end{aligned}$$

also

$$|\tilde{x} - x| \leq 2C |\tilde{y} - y|.$$

ii) Fixiere $y = f(x) \in V$ mit $x \in U$. Falls $V \ni \tilde{y} = f(\tilde{x}) \rightarrow y$, $\tilde{y} \neq y$, so folgt mit i) auch $\tilde{x} = g(\tilde{y}) \rightarrow x = g(y)$, $\tilde{x} \neq x$. Schreibe

$$\begin{aligned} R_g &:= g(\tilde{y}) - g(y) - (df(x))^{-1}(\tilde{y} - y) = \tilde{x} - x - (df(x))^{-1}(f(\tilde{x}) - f(x)) \\ &= -(df(x))^{-1}(f(\tilde{x}) - f(x) - df(x)(\tilde{x} - x)). \end{aligned}$$

Mit i) erhalten wir

$$\frac{|R_g|}{|\tilde{y} - y|} \leq 2C \|df(x)^{-1}\| \frac{|f(\tilde{x}) - f(x) - df(x)(\tilde{x} - x)|}{|\tilde{x} - x|} \rightarrow 0 \quad (\tilde{y} \rightarrow y).$$

Das heisst, g ist an der Stelle $y \in V$ differenzierbar mit

$$dg(y) = (df(x))^{-1} = (df(g(y)))^{-1},$$

und $g \in C^1(V; \mathbb{R}^n)$.

□

Mit Behauptung 4 ist nun auch Satz 8.7.1 vollständig bewiesen.

□

8.8 Implizite Funktionen

Wir beginnen mit einfachen Beispielen.

Beispiel 8.8.1. 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

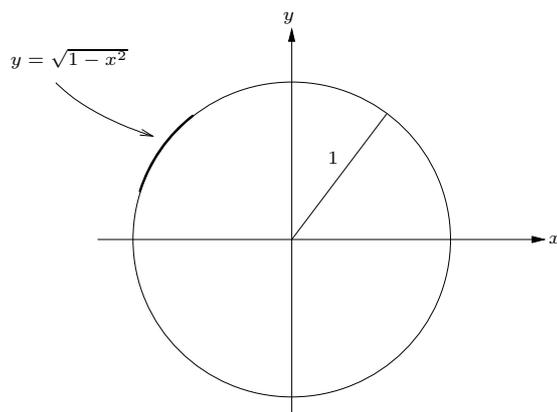
$$f(x, y) = x^2 + y^2 - 1$$

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

$$x = l(y) = \pm\sqrt{1 - y^2}, \quad -1 < y < 1.$$



Natürlich können wir M auch parametrisieren durch $\gamma(t) = (\cos(t), \sin(t))^t \in C^1(\mathbb{R}; \mathbb{R}^2)$.

ii) Sei Z der Kreiszyylinder

$$Z = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 = 1\},$$

und für festes $\mu \in \mathbb{R}$ sei M die Ellipse

$$M = \{(x, y, z) \in Z; z = \mu x\}.$$

Wie in i) können wir M lokal als Graph $(y, z) = h(x)$ schreiben mit

$$h(x) = \begin{pmatrix} \pm\sqrt{1-x^2} \\ \mu x \end{pmatrix}, \quad -1 < x < 1,$$

oder als Graph $(x, z) = \ell(y)$ mit

$$\ell(y) = \begin{pmatrix} \pm\sqrt{1-y^2} \\ \pm\mu\sqrt{1-y^2} \end{pmatrix}, \quad -1 < y < 1.$$

Falls $\mu \neq 0$, können wir M auch darstellen als Graph der Funktion $(x, y) = m(z)$ mit

$$m(z) = \begin{pmatrix} z/\mu \\ \pm\sqrt{1-(z/\mu)^2} \end{pmatrix}, \quad |z| < \mu.$$

Wir können M auch durch $\gamma(t) = (\cos(t), \sin(t), \mu \cos(t))^t \in C^1(\mathbb{R}; \mathbb{R}^3)$ parametrisieren oder **implizit** darstellen als $M = f^{-1}(\{0\})$ mit

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 - 1 \\ z - \mu x \end{pmatrix} : \mathbb{R}^3 \rightarrow \mathbb{R}^2.$$

iii) 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

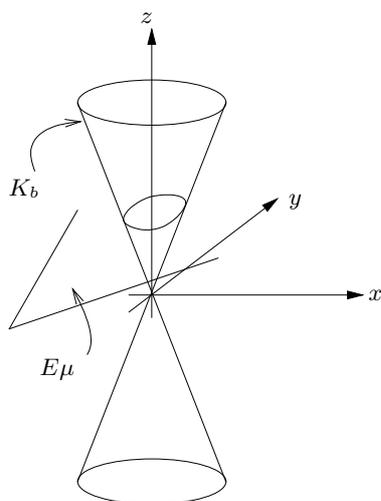
$$\Gamma : \begin{aligned} x^2 + y^2 &= b^2(1 + \mu x)^2 = b^2 + 2\mu b^2 x + b^2 \mu^2 x^2, \\ z &= 1 + \mu x. \end{aligned}$$

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.

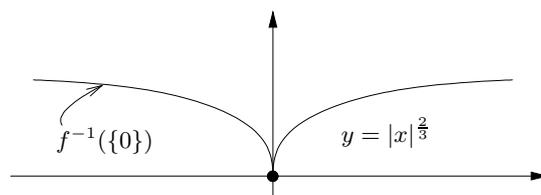
Implizit lassen sich diese Kegelschnitte wiederum bequem darstellen in der Form $\Gamma = f^{-1}(\{(0, 0)\})$, wobei $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben ist durch

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 - b^2 z^2 \\ z - (1 + \mu x) \end{pmatrix}.$$

Wie in i) lassen sich die betrachteten Schnittkurven offenbar ebenfalls lokal als Graph von Funktionen $(y, z) = h(x)$, bzw. $(x, z) = l(y)$ bzgl. x oder y schreiben. Beachte: h und l sind vektorwertig und haben ebensoviele Komponenten wie f .



iv) 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 8.8.1. 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 8.8.1. Offenbar gilt stets

$$\text{Rang}(df(p_0)) \leq \min\{n, l\},$$

und Gleichheit gilt in folgenden Fällen:

- $n \leq l$: falls $df(p_0)$ injektiv,
- $n \geq l$: falls $df(p_0)$ surjektiv,
- $n = l$: falls $df(p_0)$ bijektiv.

Definition 8.8.2. Der Punkt p_0 heisst **regulärer Punkt** von f , falls

$$\text{Rang}(df(p_0)) = \min\{n, l\},$$

das heisst, falls der Rang von $df(p_0)$ maximal ist.

Falls $n = l$, so ist f in der Nähe eines regulären Punktes invertierbar nach Satz 8.7.1. Im folgenden interessiert uns jedoch der Fall $n > l$. Wir betrachten erneut die Beispiele 8.8.1.i-iii).

Beispiel 8.8.2. i) Für $f(x, y) = x^2 + y^2 - 1$, $(x, y) \in \mathbb{R}^2$, gilt

$$\forall (x, y) \in f^{-1}(\{0\}) : df(x, y) = (2x, 2y) \neq (0, 0);$$

das heisst, jedes $p_0 = (x_0, y_0)$ mit $f(p_0) = 0$ ist regulär.

ii) Die Funktion

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 - 1 \\ z - \mu x \end{pmatrix} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

mit festem $\mu \neq 0$ hat das Differential

$$df(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ -\mu & 0 & 1 \end{pmatrix}.$$

Offenbar hat $df(x, y, z)$ den Rang 2 für alle $p_0 = (x_0, y_0, z_0) \in f^{-1}(\{(0, 0)\})$; also sind alle diese Punkte regulär.

iii) Für die Funktion

$$f(x, y, z) = \begin{pmatrix} x^2 + y^2 - b^2 z^2 \\ z - (1 + \mu x) \end{pmatrix} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$$

mit Konstanten $\mu \in \mathbb{R}$, $b > 0$ hat

$$df(x, y, z) = \begin{pmatrix} 2x & 2y & -2b^2 z \\ -\mu & 0 & 1 \end{pmatrix}$$

den Rang 2 für alle $p_0 = (x_0, y_0, z_0) \in f^{-1}(\{(0, 0)\})$; diese Punkte sind also allesamt regulär.

iv) Für $f(x, y) = x^2 - y^3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$df(x, y) = (2x, -3y^2)$$

ist der Punkt $p_0 = (0, 0)$ mit $f(p_0) = 0$ **nicht** regulär.

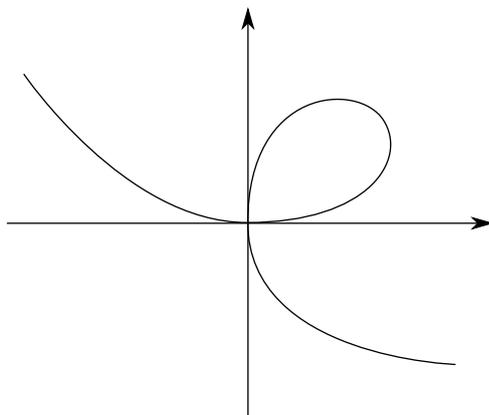
v) Sei $f(x, y) = x^3 + y^3 - 3xy : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$df(x, y) = 3(x^2 - y, y^2 - x), \quad (x, y) \in \mathbb{R}^2.$$

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 8.8.1. (Satz über implizite Funktionen) 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

$$\partial_y f(p_0) = \left(\frac{\partial f^i}{\partial y^j}(p_0) \right)_{1 \leq i, j \leq l} : \mathbb{R}^l \rightarrow \mathbb{R}^l$$

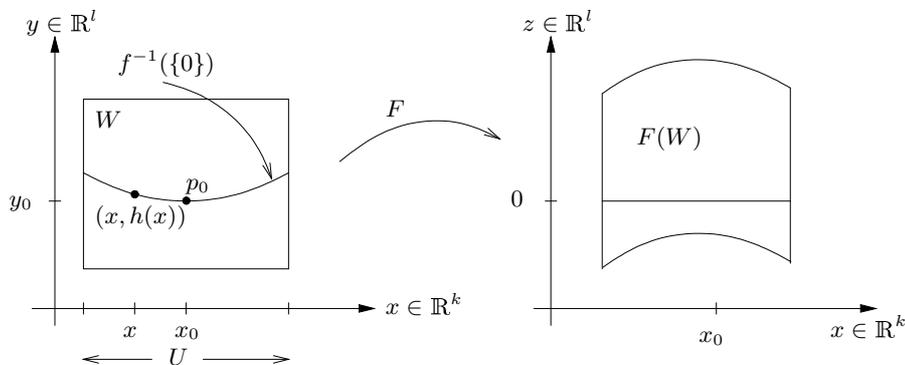
invertierbar. Sei $p_0 = (x_0, y_0)$ in diesen Koordinaten.

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:

$$\forall x \in U : f(x, h(x)) = 0, \tag{8.8.1}$$

und

$$f^{-1}(\{0\}) \cap W = \mathcal{G}(h) = \{(x, h(x)); x \in U\}. \tag{8.8.2}$$



Beispiel 8.8.3. i) Im Falle des Beispiels 8.8.2.v) wird demnach die Menge $S = f^{-1}(\{0\})$ ausser bei $p_0 = (0, 0)$ lokal um jeden Punkt als ein Graph dargestellt. Bei $p_0 = (0, 0)$ kreuzt sich S ; es kann dort keine derartige Darstellung geben.

ii) Im Falle des Beispiels 8.8.2.iv) mit $f(x, y) = x^2 - y^3$, $(x, y) \in \mathbb{R}^2$ ist offenbar $S = f^{-1}(\{0\})$ global als Graph der Funktion

$$h(x) = \sqrt[3]{x^2}, \quad x \in \mathbb{R}$$

darstellbar; jedoch ist h in $x_0 = 0$ nicht differenzierbar. ($p_0 = (0, 0)$ ist nicht regulär.)

Beweis von Satz 8.8.1. OBdA sei $p_0 = 0 \in \mathbb{R}^n$. Betrachte die Abbildung $F \in C^1(\Omega; \mathbb{R}^n)$ mit

$$F: \Omega \ni p = \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ f(x, y) \end{pmatrix} \in \mathbb{R}^k \times \mathbb{R}^l \cong \mathbb{R}^n.$$

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

$$\det(dF(p_0)) = \det(\partial_y f(p_0)) \neq 0;$$

das heisst, $dF(p_0)$ ist invertierbar. Nach dem Umkehrsatz, Satz 8.7.1, 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

$$G = (g_1, g_2) = (F|_{\tilde{U}})^{-1} \in C^1(\tilde{V}; \mathbb{R}^k \times \mathbb{R}^l)$$

von F . Mit der Darstellung von F folgt

$$(x, z) = F(G(x, z)) = (g_1(x, z), f(g_1(x, z), g_2(x, z)))$$

für alle $(x, z) \in \tilde{V}$.

Insbesondere erhalten wir für $z = 0$ die Identität

$$g_1(x, 0) = x \quad \text{für alle } x \in U := \{x; (x, 0) \in \tilde{V}\} \quad (8.8.3)$$

und somit auch

$$\forall x \in U : f(x, g_2(x, 0)) = 0.$$

Für $h(x) := g_2(x, 0) \in C^1(U; \mathbb{R}^l)$ folgt somit (8.8.1), wie gewünscht.

Mit

$$F(f^{-1}(\{0\}) \cap \tilde{U}) = \{(x, 0) \in \tilde{V}\}$$

und (8.8.3) 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 (8.8.2). \square

Bemerkung 8.8.2. i) Mittels Kettenregel kann man aus (8.8.1) eine Gleichung für das Differential dh der "impliziten Funktion" h herleiten. Sei dazu $\Phi \in C^1(U; \mathbb{R}^n)$ die Funktion $\Phi(x) = (x, h(x))^t$, $x \in U$. Dann folgt mit (8.8.1) die Identität

$$\begin{aligned} 0 &= d(f \circ \Phi)(x) = df(x, h(x))d\Phi(x) = \left(\partial_x f(x, h(x)), \partial_y f(x, h(x)) \right) \begin{pmatrix} id_{\mathbb{R}^k} \\ dh(x) \end{pmatrix} \\ &= \partial_x f(x, h(x)) + \partial_y f(x, h(x))dh(x), \end{aligned}$$

also

$$dh(x) = -(\partial_y f(x, h(x)))^{-1} \partial_x f(x, h(x)). \quad (8.8.4)$$

ii) Die obige Abbildung $\Phi(x) = (x, h(x))^t$ ist eine Parameterdarstellung der Niveaumenge $S = f^{-1}(\{0\})$ in einer Umgebung von p_0 .

iii) Falls in Satz 8.8.1 zusätzlich $f \in C^m(\Omega)$ für ein $m \geq 1$, so folgt mit (8.8.4) auch $h \in C^m(U; \mathbb{R}^l)$.

Beispiel 8.8.4. Sei $f(x, y) = x + y + x^2y$, $(x, y) \in \mathbb{R}^2$, 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 (8.8.4) gilt

$$h'(x) = \frac{dh}{dx}(x) = -\frac{\partial_x f(x, h(x))}{\partial_y f(x, h(x))} = -\frac{1 + 2xh(x)}{1 + x^2}, \quad h(0) = 0;$$

das heisst,

$$((1 + x^2)h(x))' = (1 + x^2)h'(x) + 2xh(x) = -1.$$

Als Lösung dieses Anfangswertproblems erhalten wir

$$h(x) = -\frac{x}{1 + x^2}, \quad x \in \mathbb{R}.$$

Wir verifizieren leicht

$$f(x, h(x)) = x - \frac{x}{1 + x^2}(1 + x^2) = 0.$$

8.9 Extrema mit Nebenbedingungen

Auch diesen Abschnitt beginnen wir mit einem einfach zu durchschauenden Beispiel.

Beispiel 8.9.1. Sei $f(x, y) = x(1 + y)$, $g(x, y) = x^2 + y^2 - 1$, $(x, y) \in \mathbb{R}^2$, und sei

$$S = \{(x, y) \in \mathbb{R}^2; g(x, y) = 0\} = S^1.$$

Wir wollen das Maximum von f auf $S = S^1$ bestimmen. Parametrisiere dazu 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

$$\begin{aligned} 0 &= \frac{d}{dt} f(\gamma(t)) \Big|_{t=t_0} = \frac{d}{dt} (\cos(t)(1 + \sin(t))) \Big|_{t=t_0} \\ &= \cos^2(t_0) - \sin(t_0)(1 + \sin(t_0)) = x_0^2 - y_0^2 - y_0 = 1 - 2y_0^2 - y_0; \end{aligned}$$

also

$$y_0 = \frac{1}{2}, \quad x_0 = \pm \frac{\sqrt{3}}{2}, \quad f(p_0) = \pm \frac{3\sqrt{3}}{4}$$

oder

$$y_0 = -1, \quad x_0 = 0, \quad f(p_0) = 0.$$

Offenbar ist $p_0 = (\frac{\sqrt{3}}{2}, \frac{1}{2})$ 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; das heisst, wir suchen

$$\max\{f(p); p \in \Omega, g(p) = 0\}.$$

Kann man die gewünschten Extremalstellen auch ohne eine explizite Parametrisierung der "zulässigen Menge"

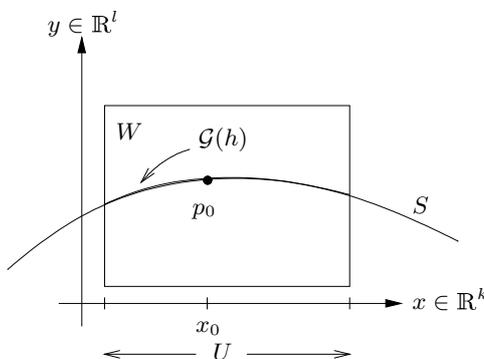
$$S = \{p \in \Omega; g(p) = 0\}$$

finden? Notwendige Bedingungen für Extrema liefert Satz 8.8.1.

8.9.1 Notwendige Bedingungen

Sei $p_0 \in S$ 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 8.8.1, 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

$$S \cap W = \mathcal{G}(h) = \{(x, h(x)); x \in U\}.$$



Wie in Bemerkung 8.8.2.ii) liefert die Abbildung

$$\Phi(x) = (x, h(x))^t \in C^1(U; \mathbb{R}^n)$$

eine Parameterdarstellung für S nahe p_0 . Nach Annahme ist p_0 ein lokales Maximum von f auf S ; also ist x_0 eine lokale Maximalstelle der Funktion $f \circ \Phi$ auf U . Mit Satz 8.5.3.i) und der Kettenregel folgt

$$0 = d(f \circ \Phi)(x_0) = df(p_0)d\Phi(x_0) = \partial_x f(p_0) + \partial_y f(p_0)dh(x_0).$$

Gemäss (8.8.4) in Bemerkung 8.8.2 gilt andererseits

$$dh(x_0) = -(\partial_y g(p_0))^{-1} \partial_x g(p_0).$$

Wir erhalten somit die Gleichung

$$0 = \partial_x f(p_0) + \lambda \partial_x g(p_0), \quad (8.9.1)$$

wobei λ die lineare Abbildung

$$\lambda = -\partial_y f(p_0)(\partial_y g(p_0))^{-1}: \mathbb{R}^l \rightarrow \mathbb{R} \quad (8.9.2)$$

bezeichnet, dargestellt durch $\lambda = (\lambda_1, \dots, \lambda_l)$. Beachte, dass mit (8.9.2) automatisch auch gilt

$$0 = \partial_y f(p_0) + \lambda \partial_y g(p_0); \quad (8.9.3)$$

das heisst, es gilt

$$0 = df(p_0) + \lambda dg(p_0).$$

Wir haben gezeigt:

Satz 8.9.1. (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 $\lambda = (\lambda_1, \dots, \lambda_l) \in \mathbb{R}^l$, so dass für $L = f + \lambda g$ gilt

$$dL(p_0) = df(p_0) + \lambda dg(p_0) = 0.$$

Satz 8.9.1 motiviert die folgenden Definitionen.

Definition 8.9.1. i) Der Vektor $\lambda = (\lambda_1, \dots, \lambda_l) = \lambda(p_0) \in \mathbb{R}^l$ mit (8.9.2) heisst **Lagrange-Multiplikator**.

ii) Die Funktion $L = f + \lambda g$ 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 (8.9.3) erhalten wir dann auch die Charakterisierung (8.9.2) von $\lambda = \lambda(p_0)$.

Beispiel 8.9.2. i) Sei $f(x, y) = x(1 + y)$ wie in Beispiel 8.9.1. Satz 8.9.1 ergibt als notwendige Bedingung für das Vorliegen einer Maximalstelle unter der Nebenbedingung

$$g(x, y) = x^2 + y^2 - 1 = 0 \quad (8.9.4)$$

in (x_0, y_0) die Gleichung

$$0 = d(f + \lambda g)(x_0, y_0) = (1 + y_0, x_0) + 2\lambda(x_0, y_0) \quad (8.9.5)$$

mit einer zu bestimmenden Zahl $\lambda \in \mathbb{R}$. Aus den Gleichungen (8.9.4), (8.9.5) können wir $(x_0, y_0) \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ wie folgt ermitteln. Die Identität (8.9.5) ist offenbar äquivalent zu dem Paar von Gleichungen

$$1 + y_0 + 2\lambda x_0 = 0, \quad x_0 + 2\lambda y_0 = 0.$$

Nehmen wir an, $y_0 = 0$, so folgt $x_0 = -2\lambda y_0 = 0$, $g(x_0, y_0) = -1$, und die Gleichung $g(x_0, y_0) = 0$ führt zu einem Widerspruch. Also gilt $y_0 \neq 0$, $\lambda = -\frac{x_0}{2y_0}$ und

$$1 + y_0 - \frac{x_0^2}{y_0} = 0.$$

Zusammen mit (8.9.4) ergibt dies die Gleichung

$$0 = y_0^2 + y_0 - x_0^2 = 2y_0^2 + y_0 - 1$$

mit den Lösungen

$$y_0 = -\frac{1}{4} \pm \sqrt{\frac{1}{16} + \frac{1}{2}} = \frac{-1 \pm 3}{4} \in \left\{ \frac{1}{2}, -1 \right\}$$

wie in Beispiel 8.9.1. Mit (8.9.4) erhalten wir wiederum

$$y_0 = \frac{1}{2}, \quad x_0 = \pm \frac{\sqrt{3}}{2}, \quad f(x_0, y_0) = \pm \frac{3\sqrt{3}}{4},$$

oder

$$y_0 = -1, \quad x_0 = 0, \quad f(x_0, y_0) = 0.$$

Die Funktion f nimmt daher auf S^1 ihr Maximum an im Punkt $(\frac{\sqrt{3}}{2}, \frac{1}{2})$, ihr Minimum im Punkt $(-\frac{\sqrt{3}}{2}, \frac{1}{2})$.

ii) Seien $f(x, y) = xy$, $g(x, y) = x + y - 1$, $(x, y) \in \mathbb{R}^2$, und sei

$$S = \{(x, y) \in \mathbb{R}^2; x + y = 1\} = g^{-1}(\{0\}).$$

Notwendig für das Vorliegen eines Maximum von f auf S am Punkt $(x, y) \in S$ sind die Bedingungen

$$x + y = 1, \quad d(f - \lambda g) = (y - \lambda, x - \lambda) = (0, 0)$$

für ein $\lambda \in \mathbb{R}$; das heisst,

$$x = \lambda = y = \frac{1}{2}.$$

iii) Sei $f(x, y) = y$, $g(x, y) = x^2 - y^3$, $(x, y) \in \mathbb{R}^2$. Offenbar gilt

$$\forall (x, y) \in S = g^{-1}(\{0\}) : f(x, y) \geq 0 = f(0, 0),$$

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 8.9.1 versagt also. Der Grund ist natürlich, dass der Punkt $p_0 = (0, 0)$ nicht regulär ist für g .

8.9.2 Geometrische Deutung

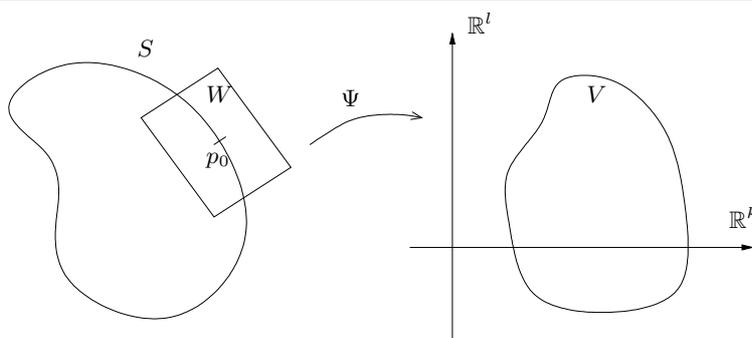
Sei $\Omega \subset \mathbb{R}^n$ offen.

Definition 8.9.2. Eine Abbildung $\Psi \in C^m(\Omega; \mathbb{R}^n)$ heisst **Diffeomorphismus** von Ω auf $V = \Psi(\Omega)$ der Klasse C^m , $m \geq 1$, falls Ψ injektiv ist und $\Psi^{-1} \in C^m(V; \mathbb{R}^n)$.

Eine geometrische Interpretation von Satz 8.8.1 ergibt sich mit dem Begriff einer k -dimensionalen Untermannigfaltigkeit im \mathbb{R}^n .

Definition 8.9.3. Eine Menge $S \subset \mathbb{R}^n$ ist eine k -dimensionale **Untermannigfaltigkeit** der Klasse C^m , $m \geq 1$, falls zu jedem $p_0 \in S$ eine Umgebung W von p_0 und ein Diffeomorphismus $\Psi \in C^m(W; \mathbb{R}^n)$ von W auf $V = \Psi(W) \subset \mathbb{R}^n$ existieren, so dass gilt:

$$\Psi(W \cap S) = V \cap (\mathbb{R}^k \times \{0\}).$$



Definition 8.9.4. Sei $g \in C^m(\Omega; \mathbb{R}^l)$, $m \geq 1$. Ein $b \in \mathbb{R}^l$ heisst **regulärer Wert** von g , falls $g^{-1}(\{b\})$ nur reguläre Punkte enthält, andernfalls heisst b **singulär**.

Beispiel 8.9.3. i) Reguläre Niveaumengen. Sei $\Omega \subset \mathbb{R}^n$ offen, $g \in C^m(\Omega; \mathbb{R}^l)$, $l < n$, $S = g^{-1}(\{0\})$, und sei $p_0 \in S$ regulär. Dann gibt es nach Satz 8.8.1 eine Umgebung W von p_0 so, dass $S \cap W$ eine k -dimensionale Untermannigfaltigkeit der Klasse C^m ist. Die im Beweis von Satz 8.8.1 konstruierte Abbildung F liefert einen Diffeomorphismus $F =: \Psi \in C^m(W; \mathbb{R}^n)$ von W auf $V = \Psi(W) \subset \mathbb{R}^n$ mit

$$\Psi(W \cap S) = V \cap (\mathbb{R}^k \times \{0\}).$$

Falls $b = 0$ regulärer Wert von g ist, also jeder Punkt $p_0 \in S$ regulär ist, so ist ganz S eine k -dimensionale Untermannigfaltigkeit der Klasse C^m .

ii) Graphen. Sei $U \subset \mathbb{R}^k$ offen, $h \in C^m(U; \mathbb{R}^l)$, $n = k + l$. Dann ist

$$G(h) = \{(x, h(x)); x \in U\}$$

eine k -dimensionale Untermannigfaltigkeit der Klasse C^m in \mathbb{R}^n .

Beweis. Wähle $W = U \times \mathbb{R}^l$. Setze $\Psi(x, y) = (x, y - h(x))$, $x \in U$. Offenbar

ist $\Psi \in C^m(W; \mathbb{R}^n)$ injektiv, und

$$d\Psi(x, y) = \begin{pmatrix} id_{\mathbb{R}^k} & 0 \\ -dh(x) & id_{\mathbb{R}^l} \end{pmatrix}$$

ist an jeder Stelle $(x, y) \in W$ invertierbar. Nach Satz 8.7.1 ist Ψ^{-1} auf $\Psi(W) = W$ differenzierbar mit

$$d\Psi^{-1} = (d\Psi \circ \Psi^{-1})^{-1} \in C^{m-1};$$

das heisst, $\Psi^{-1} \in C^m(W; \mathbb{R}^n)$ und $\Psi(S \cap W) \subset \mathbb{R}^k \times \{0\}$, wie gewünscht. \square

Sei S eine k -dimensionale Untermannigfaltigkeit in \mathbb{R}^n der Klasse C^m , $m \geq 1$. Zu $p_0 \in S$ betrachten wir Kurven $\gamma \in C^1(]-\varepsilon, \varepsilon[; S)$, wobei $\varepsilon > 0$, mit $\gamma(0) = p_0$.

Definition 8.9.5. *Der Raum*

$$T_{p_0}S = \{\dot{\gamma}(0); \gamma \in C^1(]-\varepsilon, \varepsilon[; S), \gamma(0) = p_0\}$$

heisst **Tangentialraum** an S in p_0 .

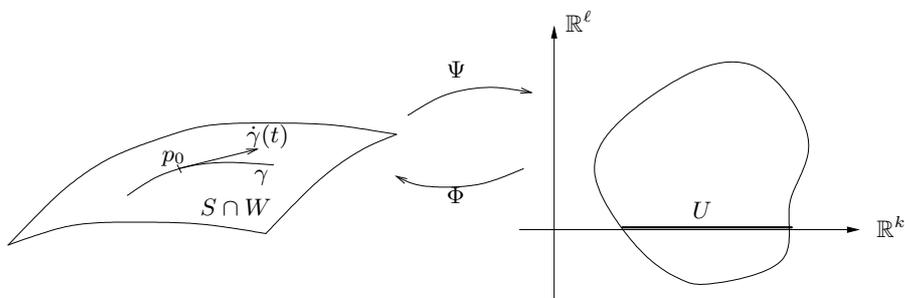
Satz 8.9.2. $T_{p_0}S$ ist ein \mathbb{R} -Vektorraum der Dimension k .

Beweis. Sei W Umgebung von p_0 , $\Psi : W \rightarrow V = \Psi(W)$ ein Diffeomorphismus mit $\Psi(p_0) = x_0 = 0$ und

$$\Psi(S \cap W) = V \cap (\mathbb{R}^k \times \{0\}) = U \subset \mathbb{R}^k \times \{0\} \cong \mathbb{R}^k.$$

Die Abbildung Ψ liefert für S die Parametrisierung

$$\Phi = \Psi^{-1}|_U : U \rightarrow S \in C^m(U; \mathbb{R}^n).$$



Die Aussage des Satzes folgt nun unmittelbar aus der Charakterisierung

$$T_{p_0}S = d\Phi(0)(\mathbb{R}^k). \quad (8.9.6)$$

Zum Beweis von (8.9.6) fixiere $\xi \in \mathbb{R}^k$. Dann ist $\gamma(t) = \Phi(t\xi) \in C^1(]-\varepsilon, \varepsilon[)$ Kurve durch $\gamma(0) = \Phi(0) = p_0$ mit $\dot{\gamma}(0) = d\Phi(0)\xi$; das heisst,

$$d\Phi(0)(\mathbb{R}^k) \subset T_{p_0}S.$$

Sei umgekehrt $\gamma \in C^1(]-\varepsilon, \varepsilon[; S)$ Kurve durch $\gamma(0) = p_0$. Betrachte den Weg $(\Psi \circ \gamma)(t) \in C^1(]-\varepsilon, \varepsilon[; U)$ mit

$$\frac{d}{dt}(\Psi \circ \gamma)(0) = d\Psi(p_0)\dot{\gamma}(0) =: \xi \in \mathbb{R}^k.$$

Da $\Psi \circ \Phi = id$, gilt insbesondere $d\Psi(p_0)d\Phi(0) = id$. Es folgt

$$d\Phi(0)\xi = (d\Psi(p_0))^{-1}d\Psi(p_0)d\Phi(0)\xi = (d\Psi(p_0))^{-1}\xi = \dot{\gamma}(0),$$

und damit (8.9.6) und der Satz. \square

Die Darstellung (8.9.6) des Tangentialraums $T_{p_0}S$ im Beweis von Satz 8.9.3 liefert sofort eine analoge Charakterisierung im Falle von Graphen.

Satz 8.9.3. Sei $S = G(h)$ mit $h \in C^m(U; \mathbb{R}^l)$, $U \subset \mathbb{R}^k$ offen, $k + l = n$ und sei $p_0 = (x_0, h(x_0)) \in S$, $x_0 \in U$. Dann gilt:

$$T_{p_0}S = \{(\xi, dh(x_0)\xi) \in \mathbb{R}^k \times \mathbb{R}^l; \xi \in \mathbb{R}^k\}.$$

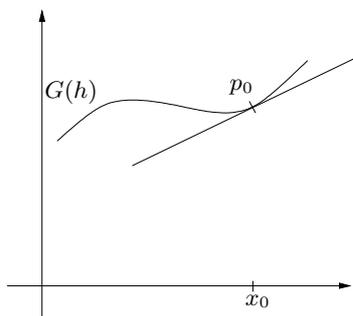
Beweis. Setze $\Phi(x) = (x, h(x))$, $x \in U$. Dann ist $\Phi \in C^m(U; \mathbb{R}^n)$ Parametrisierung von S , und die Behauptung folgt mit (8.9.6). \square

Beispiel 8.9.4. Falls $h \in C^1(\mathbb{R})$, $x_0 \in \mathbb{R}$, so hat die Tangente $t(s)$ an $G(h)$ in $p_0 = (x_0, h(x_0))$ die Darstellung

$$t(s) = p_0 + (s, h'(x_0)s), \quad s \in \mathbb{R};$$

die Tangentialvektoren sind also von der Form

$$T_{p_0}G(h) = \{(s, h'(x_0)s); s \in \mathbb{R}\}.$$



Für reguläre Niveauflächen erhalten wir die folgende Charakterisierung:

Satz 8.9.4. Sei $g \in C^m(\Omega; \mathbb{R}^l)$, $l < n$, und sei $b \in \mathbb{R}^l$ regulärer Wert. Sei weiter $S = g^{-1}(\{b\})$. Dann gilt:

$$\forall p_0 \in S : T_{p_0}S = \ker(dg(p_0)).$$

Beweis. OBdA sei $b = 0$. Sei $\Phi(x) = (x, h(x))$, $x \in U$, wobei $h \in C^m(U; \mathbb{R}^l)$ wie in Satz 8.8.1 mit $\Phi(x_0) = (x_0, h(x_0)) = p_0$. Dann ist $\Phi \in C^m(U; \mathbb{R}^n)$ Parametrisierung von $S = g^{-1}(\{0\})$ um p_0 , und mit $g \circ \Phi \equiv 0$ folgt

$$d(g \circ \Phi)(x_0) = dg(p_0)d\Phi(x_0) = 0;$$

also mit Satz 8.9.2 und (8.9.6)

$$\mathbb{R}^k \cong T_{p_0}S = d\Phi(x_0)(\mathbb{R}^k) \subset \ker(dg(p_0)).$$

Da die Dimension des Bildes von $dg(p_0)$ nach Annahme gleich l ist, ergibt die Rangformel

$$\dim(\ker(dg(p_0))) = n - l = k,$$

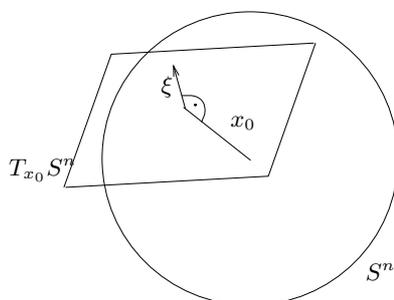
und die Behauptung folgt. \square

Beispiel 8.9.5. 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

$$\forall x \in S^n : dg(x) = 2x^t \neq 0, \quad .$$

Es folgt, S^n ist reguläre n -dimensionale Untermannigfaltigkeit in \mathbb{R}^{n+1} . Weiter liefert Satz 8.9.4 die Darstellung

$$\forall x \in S^n : T_x S^n = \ker(dg(x)) = \{\xi \in \mathbb{R}^{n+1}; x \cdot \xi = 0\}.$$



Bemerkung 8.9.1. i) Wir können Satz 8.9.1 nun auch geometrisch deuten. Seien dazu $\Omega \subset \mathbb{R}^n$, $f \in C^1(\Omega)$, $g \in C^1(\Omega; \mathbb{R}^l)$ für ein $l < n$ wie in Satz 8.9.1, dazu $p_0 = (x_0, y_0) \in S = g^{-1}(\{0\})$ regulär, $U, W, h \in C^1(U, \mathbb{R}^l)$ wie in Satz 8.8.1 mit

$$S \cap W = \mathcal{G}(h) = \{(x, h(x)); x \in U\},$$

und nimm an, $f(p_0) = \max_S f$.

Für jede Kurve $\gamma \in C^1(]-\varepsilon, \varepsilon[; S)$ durch $\gamma(0) = p_0$ gilt gemäss Korollar 6.5.1.i)

$$0 = \left. \frac{d}{dt} \right|_{t=0} (f \circ \gamma)(t) = df(p_0) \dot{\gamma}(0);$$

das heisst,

$$\forall \eta \in T_{p_0} S : df(p_0) \eta = \nabla f(p_0) \cdot \eta = 0,$$

oder

$$\nabla f(p_0) \perp T_{p_0} S.$$

Andererseits gilt nach Satz 8.9.4 für $1 \leq i \leq l$ die Bedingung

$$\nabla g^i(p_0) \perp T_{p_0} S.$$

Da nach Annahme $\text{Rang}(dg(p_0)) = l$, sind die Vektoren $\nabla g^1(p_0), \dots, \nabla g^l(p_0)$ linear unabhängig; sie spannen daher den **Normalraum**

$$T_{p_0}^\perp S = \{v \in T_{p_0} \mathbb{R}^n \cong \mathbb{R}^n; \forall \eta \in T_{p_0} S : v \cdot \eta = 0\}$$

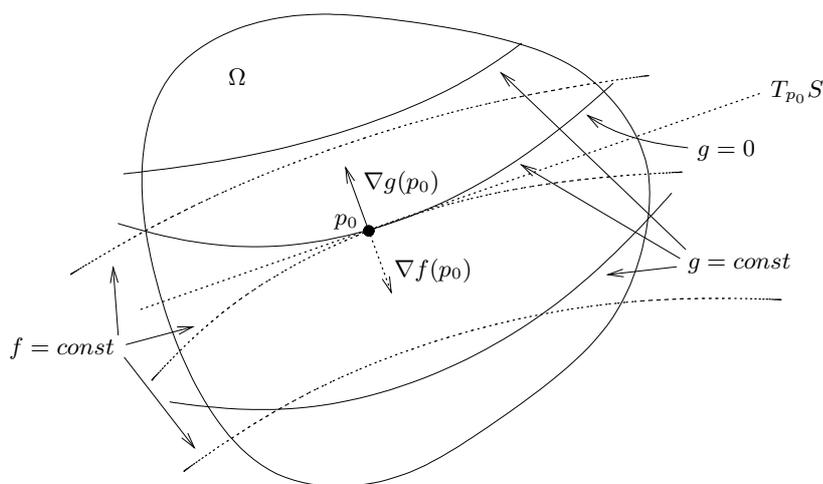
auf, und es existiert $\lambda = (\lambda_1, \dots, \lambda_l) \in \mathbb{R}^l$ mit

$$\nabla f(p_0) = \lambda_1 \nabla g^1(p_0) + \dots + \lambda_l \nabla g^l(p_0);$$

das heisst,

$$df(p_0) = \sum_{i=1}^l \lambda_i dg^i(p_0) = \lambda dg(p_0).$$

Dies ist die Aussage von Satz 8.9.1.



ii) Speziell in Fall $n = 2, l = 1$ bedeutet die Relation $\nabla f(p_0) \perp T_{p_0}S$ in einem kritischen Punkt $p_0 \in S = g^{-1}(\{0\})$, dass $\nabla f(p_0)$ auf der Niveaulinie $g^{-1}(\{0\})$ von g senkrecht steht, also proportional ist zu $\nabla g(p_0)$. Folglich existiert $\lambda \in \mathbb{R}$ mit

$$\nabla f(p_0) + \lambda \nabla g(p_0) = 0$$

oder

$$df(p_0) + \lambda dg(p_0) = 0.$$

8.9.3 Hinreichende Bedingungen

Analog zu Satz 8.5.3.ii) erhalten wir auch eine hinreichende Bedingung für das Vorliegen eines Minimums oder Maximums in einem kritischen Punkt von f in $S = g^{-1}(\{0\})$.

Satz 8.9.5. 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

$$\text{Hess}_L(p_0)(\eta, \eta) > 0$$

für alle $\eta \in T_{p_0}S \setminus \{0\}$, so ist p_0 ein striktes relatives Minimum von f auf S .

Beweis. Sei

$$\Phi(x) = (x, h(x)) \in C^1(U; \mathbb{R}^n)$$

eine lokale Parametrisierung von $S = g^{-1}(\{0\})$ in einer Umgebung W von $p_0 = \Phi(x_0)$ gemäss Satz 8.8.1. Da $g \circ \Phi \equiv 0$ auf U , folgt mit der Rechnung

$$\begin{aligned} Hess_{f \circ \Phi}(x_0)(\xi, \xi) &= \frac{d^2}{ds^2} \Big|_{s=0} (f \circ \Phi)(x_0 + s\xi) = \frac{d^2}{ds^2} \Big|_{s=0} (L \circ \Phi)(x_0 + s\xi) \\ &= \frac{d}{ds} \Big|_{s=0} ((dL \circ \Phi)(x_0 + s\xi) \frac{d}{ds} \Phi(x_0 + s\xi)) \\ &= \underbrace{dL(p_0)}_{=0} \frac{d^2}{ds^2} \Big|_{s=0} (\Phi(x_0 + s\xi)) + Hess_L(p_0) \underbrace{(d\Phi(x_0)\xi)}_{=\eta} \underbrace{d\Phi(x_0)\xi}_{=\eta} \\ &= Hess_L(p_0)(\eta, \eta) > 0 \end{aligned}$$

für alle $\xi \in \mathbb{R}^k \setminus \{0\}$ mit $\eta = d\Phi(x_0)\xi \in T_{p_0}S \setminus \{0\}$ die Behauptung aus Satz 8.5.3.ii). \square

Beispiel 8.9.6. Seien $f(x, y) = x(1+y)$, $g(x, y) = x^2 + y^2 - 1$, $(x, y) \in \mathbb{R}^2$, wie in Beispiel 8.9.1, und sei $S = g^{-1}(\{0\}) = S^1$. Wie in Beispiel 8.9.2.i) gezeigt, hat f auf S^1 die kritischen Punkte

$$p_0 = (0, -1), \quad p_{\pm} = \left(\pm \frac{\sqrt{3}}{2}, \frac{1}{2} \right) \in S^1$$

mit zugehörigen Lagrange-Multiplikatoren

$$\lambda(p_0) = 0, \quad \lambda(p_{\pm}) = \mp \frac{\sqrt{3}}{2}.$$

Mit

$$d(f - \lambda g)(x, y) = (1 + y - 2\lambda x, x - 2\lambda y)$$

folgt

$$Hess_L(p) = \begin{pmatrix} -2\lambda & 1 \\ 1 & -2\lambda \end{pmatrix}$$

mit Eigenwerten $\mu = \mu_{1,2}$, wobei

$$(2\lambda + \mu)^2 - 1 = 0, \quad \mu = \pm 1 - 2\lambda.$$

Für $p_0 = (0, -1)$ erhalten wir

$$Hess_L(p_0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

mit Eigenwerten $\mu_{1,2} = \pm 1$, für $p_+ = \left(\frac{\sqrt{3}}{2}, \frac{1}{2} \right)$ gilt:

$$Hess_L(p_+) = \begin{pmatrix} -\sqrt{3} & 1 \\ 1 & -\sqrt{3} \end{pmatrix}$$

mit Eigenwerten $\mu_{1,2} = -\sqrt{3} \pm 1 < 0$ und für $p_- = \left(-\frac{\sqrt{3}}{2}, \frac{1}{2} \right)$ folgt:

$$Hess_L(p_-) = \begin{pmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix}$$

mit Eigenwerten $\mu_{1,2} = \sqrt{3} \pm 1 > 0$. Das heisst, p_+ ist striktes lokales Maximum und p_- ist striktes lokales Minimum von f auf $S = S^1$, in Übereinstimmung mit Beispiel 8.9.2.i).

Bemerkung 8.9.2. Wie der Beweis von Satz 8.9.5 zeigt, ist die “2. Variation” der eingeschränkten Funktion $f|_S$ ist nur in den kritischen Punkten p_0 von f erklärt, und zwar durch $\text{Hess}_L(p_0)$.

8.10 Immersionen

Sei $U \subset \mathbb{R}^k$ offen, $k < n$, $\Phi \in C^1(U; \mathbb{R}^n)$.

Definition 8.10.1. Φ heisst **Immersion**, falls jeder Punkt $x \in U$ regulär ist; das heisst, falls gilt

$$\forall x \in U : \text{Rang}(d\Phi(x)) = k.$$

Beispiel 8.10.1. i) Sei $\gamma \in C^1([a, b]; \mathbb{R}^n)$ eine Kurve in \mathbb{R}^n mit $\dot{\gamma}(t) \neq 0$ für alle $t \in]a, b[$. Dann ist γ eine Immersion. *Beachte:* γ kann Selbstschnitte haben.

ii) Sei $h \in C^1(U; \mathbb{R}^l)$, $k + l = n$, und sei $\Phi(x) = (x, h(x))$, $x \in U$. Dann ist $\Phi \in C^1(U; \mathbb{R}^n)$ eine Immersion.

Wir zeigen, dass Beispiel 8.10.1.ii) typisch ist.

Satz 8.10.1. (Immersionssatz) Sei $\Phi \in C^1(U; \mathbb{R}^n)$ eine Immersion von $U \subset \mathbb{R}^k$ in \mathbb{R}^n , $n > k$. Zu $x_0 \in U$ mit $\Phi(x_0) = p_0 \in \mathbb{R}^n$ wähle Koordinaten $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$, wobei $l = n - k$, so dass

$$d\Phi(x_0)(\mathbb{R}^k) \subset \mathbb{R}^k \times \{0\}.$$

Nach Verschieben dürfen wir weiter annehmen, dass

$$p_0 = \Phi(x_0) = (x_0, y_0).$$

Dann gibt es eine Umgebung V von $x_0 \in \mathbb{R}^k$, und einen Diffeomorphismus $T \in C^1(V; \mathbb{R}^k)$ mit $T(x_0) = x_0$, sowie ein $h \in C^1(V; \mathbb{R}^l)$ mit $h(x_0) = y_0$, so dass gilt

$$\forall x \in V : (\Phi \circ T)(x) = (x, h(x)).$$

Bemerkung 8.10.1. i) Das heisst, lokal um jeden Punkt $x_0 \in U$ parametrisiert jede Immersion $\Phi \in C^1(U; \mathbb{R}^n)$ eine k -dimensionale Untermannigfaltigkeit $S = G(h)$ nach Beispiel 8.9.3.ii) mit Tangentialraum am Punkt $p_0 = \Phi(x_0) \in S$

$$T_{p_0}S = d\Phi(T(x_0))dT(x_0)(\mathbb{R}^k) = d\Phi(x_0)(\mathbb{R}^k)$$

gemäss (8.9.6) oder Satz 8.9.3. Hier benutzen wir, dass $T(x_0) = x_0$ und dass $dT(x_0): \mathbb{R}^k \rightarrow \mathbb{R}^k$ ein linearer Isomorphismus ist.

Diese Aussage gilt im allgemeinen nicht lokal um jeden Punkt $p_0 \in \Phi(U)$ in \mathbb{R}^n ; betrachte beispielsweise eine Kurve mit Selbstschnitten wie in Beispiel 8.10.1.i). Jedoch auch im Falle einer injektiven Immersion $\Phi \in C^1(U; \mathbb{R}^n)$ ist $S = \Phi(U) \subset \mathbb{R}^n$ nicht notwendig lokal als Graph darstellbar. Dies gilt im Allgemeinen nur, wenn die Umkehrabbildung $\Phi^{-1}: S \rightarrow U$ zusätzlich stetig ist bezüglich der

vom umgebenden \mathbb{R}^n induzierten Topologie auf S . In diesem Fall heisst Φ eine **Einbettung** von U in \mathbb{R}^n .

ii) Mit Satz 8.10.1 folgt, dass jede k -dimensionale Untermannigfaltigkeit S der Klasse C^1 lokal als Graph dargestellt werden kann. Sei $p_0 \in S$, dazu $W \subset \mathbb{R}^n$ und $\Psi \in C^1(W; \mathbb{R}^n)$ ein Diffeomorphismus wie in Definition 8.9.2 mit

$$\Psi(S \cap W) =: U \subset \mathbb{R}^k \times \{0\}.$$

OBdA sei $p_0 = (x_0, y_0) \in \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l$, wobei $\Psi(p_0) = x_0$, und $T_{p_0}S = \mathbb{R}^k \times \{0\}$. Dann ist $\Phi := \Psi^{-1}|_U \in C^1(U; \mathbb{R}^n)$ Immersion mit $\Phi(x_0) = p_0$ und $d\Phi(x_0)(\mathbb{R}^k) = \mathbb{R}^k \times \{0\}$. Nach Satz 8.10.1 existiert ein Diffeomorphismus $T \in C^1(V; \mathbb{R}^k)$ in einer Umgebung V von x_0 mit $x_0 = T(x_0)$ sowie ein $h \in C^1(V; \mathbb{R}^l)$, so dass

$$(\Phi \circ T)(V) = S \cap W' = G(h)$$

in einer Umgebung $W' \subset W$ von p_0 in \mathbb{R}^n .

Beweis von Satz 8.10.1. Seien $\pi_1: \mathbb{R}^n \rightarrow \mathbb{R}^k$, $\pi_2: \mathbb{R}^n \rightarrow \mathbb{R}^l$ die Projektionen

$$\pi_1(x, y) \mapsto x, \quad \pi_2(x, y) \mapsto y, \quad (x, y) \in \mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^l.$$

Betrachte die Abbildung $f = \pi_1 \circ \Phi \in C^1(U; \mathbb{R}^k)$. Nach Annahme hat

$$df(x_0) = \pi_1 \circ d\Phi(x_0) = d\Phi(x_0)$$

Rang k , ist also invertierbar. Gemäss Satz 8.7.1 gibt es Umgebungen U_0 von x_0 , V von $f(x_0) = x_0$ in \mathbb{R}^k , so dass $f|_{U_0}: U_0 \rightarrow V$ ein Diffeomorphismus ist. Setze

$$T = (f|_{U_0})^{-1} \in C^1(V; \mathbb{R}^k), \quad h = \pi_2 \circ \Phi \circ T \in C^1(V; \mathbb{R}^l).$$

Dann folgt mit $\pi_1 \circ \Phi \circ T = f \circ T = id$ für alle $x \in V$ die Gleichung

$$(\Phi \circ T)(x) = (\pi_1 \circ \Phi, \pi_2 \circ \Phi)(T(x)) = (x, h(x)),$$

wie gewünscht. □

Beispiel 8.10.2. Sei $\Phi \in C^1(\mathbb{R}^2, \mathbb{R}^3)$ gegeben mit

$$\Phi(x, y) = \frac{1}{1 + x^2 + y^2} \begin{pmatrix} 2x \\ 2y \\ 1 - x^2 - y^2 \end{pmatrix}, \quad (x, y) \in \mathbb{R}^2.$$

Mit der Identität

$$(2x)^2 + (2y)^2 + (1 - x^2 - y^2)^2 = (1 + x^2 + y^2)^2$$

erhalten wir

$$\Phi: \mathbb{R}^2 \rightarrow S^2 = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 + z^2 = 1\}.$$

Behauptung 1. Φ ist eine Immersion.

Beweis. Betrachte für festes $(x, y) \in \mathbb{R}^2$ die Matrix

$$A = \left(\Phi, \frac{\partial \Phi}{\partial x}, \frac{\partial \Phi}{\partial y} \right) (x, y).$$

Wegen Antisymmetrie der Determinante gilt:

$$\begin{aligned} \det(A) &= \frac{1}{(1+x^2+y^2)^3} \det \begin{pmatrix} 2x & 2 & 0 \\ 2y & 0 & 2 \\ 1-x^2-y^2 & -2x & -2y \end{pmatrix} \\ &= \frac{4(1-x^2-y^2) + 8x^2 + 8y^2}{(1+x^2+y^2)^3} = \frac{4}{(1+x^2+y^2)^2} \neq 0 \end{aligned}$$

Es folgt:

$$\forall (x, y) \in \mathbb{R}^2 : \text{Rang}(d\Phi(x, y)) = 2.$$

□

Behauptung 2. Φ ist injektiv, $\Phi(\mathbb{R}^2) = S^2 \setminus \{(0, 0, -1)\}$.

Beweis. Für festes $r > 0$ bildet Φ den Kreis $\partial B_r(0) \subset \mathbb{R}^2$ bijektiv ab auf den Breitenkreis auf S^2 mit der e_3 -Komponente

$$\zeta = \frac{1-r^2}{1+r^2}.$$

Die Funktion

$$\zeta : r \mapsto \frac{1-r^2}{1+r^2}$$

ist streng monoton fallend mit

$$\zeta(0) = 1, \quad \lim_{r \rightarrow \infty} \zeta(r) = -1.$$

□

Behauptung 3. Die Inverse $\pi = (\Phi)^{-1}$ von Φ liefert die stereografische Projektion vom Südpol mit der Darstellung

$$\pi : S^2 \setminus \{0, 0, -1\} \ni (\xi, \eta, \zeta) \mapsto \frac{(\xi, \eta)}{1+\zeta} \in \mathbb{R}^2.$$

Beweis. Betrachte einen beliebigen Punkt $(\xi, \eta, \zeta) = \Phi(x, y) \in S^2 \setminus \{0, 0, -1\}$. OBdA sei $\eta = 0$, also auch $y = 0$, und mit den Darstellungen

$$\xi = \frac{2x}{1+x^2}, \quad \zeta = \frac{1-x^2}{1+x^2}$$

folgt $1 + \zeta = \frac{2}{1+x^2}$, also

$$x = \frac{\xi}{1+\zeta},$$

wie gewünscht.

□

Kapitel 9

Integration im \mathbb{R}^n

9.1 Riemannsches Integral über einem Quader

Zur Definition des R-Integrals über einem n -dimensionalen Quader gehen wir vollkommen analog vor wie im Fall $n = 1$; vergleiche Abschnitt 7.2.

9.1.1 Zerlegungen und Treppenfunktionen

Definition 9.1.1. i) *Ein n -dimensionaler Quader ist ein Produkt*

$$Q = \prod_{i=1}^n I_i = \{x = (x^i)_{1 \leq i \leq n}; x^i \in I_i, 1 \leq i \leq n\}$$

von (offenen, abgeschlossenen, oder halb-offenen) Intervallen I_1, \dots, I_n . Solch ein Q hat den **Elementarinhalt**

$$\mu(Q) = \mu_n(Q) = \prod_{i=1}^n |I_i|.$$

ii) *Eine **Zerlegung** $P = \{Q_k; 1 \leq k \leq K\}$ eines Quaders $Q = \bigcup_{k=1}^K Q_k$ in disjunkte Teilquader $Q_k \subset Q, 1 \leq k \leq K$, hat die **Feinheit***

$$\delta_P = \max_{1 \leq k \leq K} \text{diam } Q_k,$$

wobei

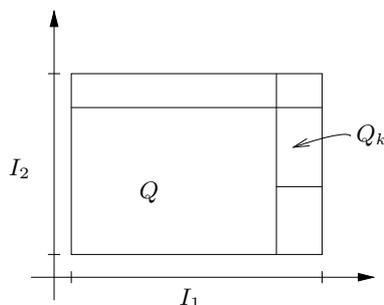
$$\text{diam } Q_k = \sup_{x, y \in Q_k} |x - y|, \quad 1 \leq k \leq K,$$

den **Durchmesser** von Q_k bezeichnet.

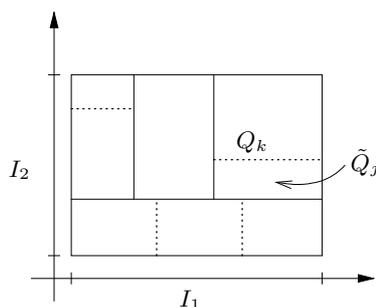
iii) *Eine Funktion $f: Q \rightarrow \mathbb{R}$ auf einem Quader Q heisst **Treppenfunktion**, falls f eine Darstellung der Form*

$$f = \sum_{k=1}^K c_k \chi_{Q_k}$$

besitzt mit einer Zerlegung $P = \{Q_k; 1 \leq k \leq K\}$ von Q und Konstanten $c_k \in \mathbb{R}$, $1 \leq k \leq K$.



Definition 9.1.2. Eine Zerlegung $\tilde{P} = \{\tilde{Q}_j; 1 \leq j \leq J\}$ ist eine **Verfeinerung** der Zerlegung $P = \{Q_k; 1 \leq k \leq K\}$ des Quaders Q , falls jedes \tilde{Q}_j in einem Quader Q_k enthalten ist.



Beispiel 9.1.1. i) Seien $P = \{Q_k; 1 \leq k \leq K\}$, $R = \{S_l; 1 \leq l \leq L\}$ Zerlegungen von Q . Dann ist

$$\tilde{P} = \{Q_k \cap S_l; 1 \leq k \leq K, 1 \leq l \leq L\}$$

Zerlegung von Q , welche sowohl P als auch R verfeinert. (Vgl. den Beweis von Lemma 7.2.1.)

ii) Sei $Q = \prod_{i=1}^n I_i$ mit Intervallen I_i , und seien $P_i = \{I_{ik}; 1 \leq k \leq K_i\}$ Zerlegungen von I_i , $1 \leq i \leq n$. Die von $(P_i)_{1 \leq i \leq n}$ induzierte **Produktzerlegung**

$$P = \prod_{i=1}^n P_i = \left\{ \prod_{i=1}^n I_{ik_i}; 1 \leq k_i \leq K_i \right\}$$

hat offenbar die Feinheit

$$\delta_P \leq \sqrt{n} \max_{1 \leq i \leq n} \delta_{P_i}.$$

iii) Jede Zerlegung P eines Quaders $Q = \prod_{i=1}^n I_i$ lässt sich zu einer Produktzerlegung verfeinern.

9.1.2 Das Riemann-Integral

Sei $Q \subset \mathbb{R}^n$ ein Quader.

Definition 9.1.3. Das Riemann-Integral (*R-Integral*) einer Treppenfunktion

$f = \sum_{k=1}^K c_k \chi_{Q_k}$ wie in Definition 9.1.1.iii) ist

$$\int_Q f \, d\mu = \int_Q \left(\sum_{k=1}^K c_k \chi_{Q_k} \right) d\mu = \sum_{k=1}^K c_k \mu(Q_k). \quad (9.1.1)$$

Bemerkung 9.1.1. Analog zu Bemerkung 7.2.1 ist die Definition des R-Integrals einer Treppenfunktion $f: Q \rightarrow \mathbb{R}$ unabhängig von der gewählten Darstellung

$f = \sum_{k=1}^K c_k \chi_{Q_k}$; insbesondere ändert sich der Wert der Summe (9.1.1) nicht bei Verfeinerungen der Zerlegung.

Sei $f: Q \rightarrow \mathbb{R}$ beschränkt. Wie in Definition 7.2.2 können wir nun das R-Integral von f definieren.

Definition 9.1.4. i) Das untere, bzw. obere R-Integral von f sind erklärt durch

$$\underline{\int}_Q f \, d\mu = \sup \left\{ \int_Q f e \, d\mu; e \text{ Treppenfunktion}; e \leq f \right\},$$

bzw.

$$\overline{\int}_Q f \, d\mu = \inf \left\{ \int_Q f g \, d\mu; g \text{ Treppenfunktion}; f \leq g \right\}.$$

ii) Die Funktion f heißt **R-integrierbar** über Q , falls

$$\underline{\int}_Q f \, d\mu = \overline{\int}_Q f \, d\mu =: \int_Q f \, d\mu.$$

Bemerkung 9.1.2. i) Analog zu Bemerkung 7.2.2 gilt (unter Verwendung von Beispiel 9.1.1) für jedes beschränkte f die Ungleichung

$$\underline{\int}_Q f \, d\mu \leq \overline{\int}_Q f \, d\mu.$$

ii) Weiter ist f R-integrierbar genau dann, wenn für jedes $\epsilon > 0$ Treppenfunktionen $e, g: Q \rightarrow \mathbb{R}$ existieren mit $e \leq f \leq g$ und

$$\int_Q g \, d\mu - \int_Q e \, d\mu < \epsilon.$$

Vollkommen analog zu den Sätzen 7.2.2 und 7.2.3 erhalten wir sodann die folgende Aussage.

Satz 9.1.1. Sei $f \in C^0(\overline{Q})$. Dann ist f über Q R-integabel, und für jede Folge $(P^{(l)})_{l \in \mathbb{N}}$ von Zerlegungen $P^{(l)} = \{Q_k^{(l)}; 1 \leq k \leq K^{(l)}\}$ von Q mit Feinheit $\delta_{P^{(l)}} \rightarrow 0$ ($l \rightarrow \infty$) gilt für eine beliebige Auswahl von Punkten $x_k^{(l)} \in Q_k^{(l)}$, $1 \leq k \leq K^{(l)}$, $l \in \mathbb{N}$, stets

$$\int_Q \left(\sum_{k=1}^{K^{(l)}} f(x_k^{(l)}) \chi_{Q_k^{(l)}} \right) d\mu = \sum_{k=1}^{K^{(l)}} f(x_k^{(l)}) \mu(Q_k^{(l)}) \rightarrow \int_Q f d\mu \quad (l \rightarrow \infty).$$

Bemerkung 9.1.3. Falls $f: Q \rightarrow \mathbb{R}$ stückweise stetig ist in dem Sinne, dass für eine Zerlegung $P = \{Q_k; 1 \leq k \leq K\}$ des Quaders Q gilt, $f \chi_{Q_k} \in C^0(\overline{Q_k})$, $1 \leq k \leq K$, so ist f ebenfalls R-integabel. Es genügt, in Definition 9.1.3 Treppenfunktionen auf Verfeinerungen der Zerlegung P zu betrachten.

Monotonie des R-Integrals folgt ebenfalls analog zu Satz 7.3.1.

Satz 9.1.2. Seien $f, g: Q \rightarrow \mathbb{R}$ beschränkt und R-integabel, und sei $f \leq g$. Dann gilt

$$\int_Q f d\mu \leq \int_Q g d\mu.$$

Weiter gilt die Linearität des R-Integrals entsprechend Satz 7.3.2.

Satz 9.1.3. Seien $f, f_1, f_2: Q \rightarrow \mathbb{R}$ beschränkt und über Q R-integabel, und sei $\alpha \in \mathbb{R}$. Dann sind die Funktionen $\alpha f, f_1 + f_2: Q \rightarrow \mathbb{R}$ über Q R-integabel, und

$$\int_Q (\alpha f) d\mu = \alpha \int_Q f d\mu,$$

bzw.

$$\int_Q (f_1 + f_2) d\mu = \int_Q f_1 d\mu + \int_Q f_2 d\mu.$$

Auch die folgende Aussage können wir ohne Mühe zeigen.

Satz 9.1.4. Seien $f, f_1, f_2: Q \rightarrow \mathbb{R}$ beschränkt und R-integabel. Dann sind die Funktionen $|f|, \min\{f_1, f_2\}$ sowie $\max\{f_1, f_2\}$ ebenfalls über Q R-integabel.

Beweis. i) Seien $e_i \leq f_i \leq g_i$ Treppenfunktionen, $i = 1, 2$. OBdA seien die zugehörigen Partitionen von Q gleich. (Betrachte sonst ihre gemeinsame Verfeinerung.) Dann sind $\min\{e_1, e_2\}, \min\{g_1, g_2\}$ Treppenfunktionen mit

$$\min\{e_1, e_2\} \leq \min\{f_1, f_2\} \leq \min\{g_1, g_2\},$$

und

$$\begin{aligned} \int_Q \min\{g_1, g_2\} d\mu - \int_Q \min\{e_1, e_2\} d\mu &= \int_Q (\min\{g_1, g_2\} - \min\{e_1, e_2\}) d\mu \\ &\leq \int_Q (g_1 - e_1) d\mu + \int_Q (g_2 - e_2) d\mu, \end{aligned}$$

da

$$0 \leq \min\{g_1, g_2\} - \min\{e_1, e_2\} \leq (g_1 - e_1) + (g_2 - e_2).$$

Mit f_1 und f_2 ist somit gemäss Bemerkung 9.1.2.ii) auch $\min\{f_1, f_2\}$ R-integrierbar; analog auch $\max\{f_1, f_2\}$.

ii) Sei f beschränkt und R-integrierbar. Aus der Darstellung $|f| = \max\{f, -f\}$ folgt mit i) die R-Integrierbarkeit von $|f|$. \square

Kombination von Satz 9.1.2 mit den Sätzen 9.1.3 und 9.1.4 ergibt die den Korollaren 7.3.1-7.3.3 analogen Aussagen.

Korollar 9.1.1. Sei $f: Q \rightarrow \mathbb{R}$ beschränkt und R-integrierbar. Dann gilt

$$\left| \int_Q f \, d\mu \right| \leq \int_Q |f| \, d\mu \leq \sup_Q |f| \cdot \mu(Q).$$

Beachte, dass gemäss Satz 9.1.4 mit f auch $|f|$ über Q R-integrierbar ist.

Korollar 9.1.2. Seien $f, f_k \in C^0(\overline{Q})$ mit $f_k \xrightarrow{glm.} f$ ($k \rightarrow \infty$). Dann gilt

$$\left| \int_Q f_k \, d\mu - \int_Q f \, d\mu \right| \leq \int_Q |f_k - f| \, d\mu \leq \|f_k - f\|_{C^0} \cdot \mu(Q) \xrightarrow{(k \rightarrow \infty)} 0.$$

Schliesslich gilt auch Satz 7.3.3 analog.

Satz 9.1.5. (Gebietsadditivität) Sei $f: Q \rightarrow \mathbb{R}$ beschränkt und über Q R-integrierbar, 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

$$\int_Q f \, d\mu = \sum_{k=1}^K \int_{Q_k} f \, d\mu.$$

9.2 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 9.1 allenfalls approximativ; vergleiche Satz 9.1.1.

Der folgende Satz hilft uns weiter.

Satz 9.2.1. (Fubini) Sei $Q = [a, b] \times [c, d] \subset \mathbb{R}^2$, und sei $f \in C^0(Q)$. Dann gilt

$$\int_Q f \, d\mu = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy.$$

Das heisst, das Integral von f über Q kann iterativ durch 1-dimensionale Integration bestimmt werden!

Beispiel 9.2.1. Sei $Q = [0, 1] \times [0, 2\pi]$, $f(x, y) = \sin(y - x) \in C^0(Q)$. Nach Satz 9.2.1 gilt

$$\int_Q f \, d\mu = \int_0^1 \underbrace{\left(\int_0^{2\pi} \sin(y - x) \, dy \right)}_{=0} dx = 0.$$

Bemerkung 9.2.1. i) Die Voraussetzung $f \in C^0(Q)$ in Satz 9.2.1 ist wichtig. Insbesondere kann man für allgemeine beschränkte Funktionen $f: Q \rightarrow \mathbb{R}$ aus der Existenz eines der iterierten Integrale **nicht** auf die Existenz der R-Integrale $\int_Q f \, d\mu$ schliessen, wie auch das folgende Beispiel 9.2.2 zeigt.

ii) In der Lebesgueschen Mass- und Integrationstheorie erhält man die Aussage von Satz 9.2.1 allgemein für integrierbare Funktionen f .

Beispiel 9.2.2. Sei $Q = [0, 1] \times [0, 2\pi] \subset \mathbb{R}^2$,

$$f(x, y) = \chi_Q(x) \cdot \sin(y - x): Q \rightarrow \mathbb{R}.$$

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-integrierbar, und auch $\int_0^1 f(x, y) \, dx$ existiert für kein $y \in [0, 2\pi]$.

Beweis von Satz 9.2.1. Seien

$$P_1 = \{I_{1j}; 1 \leq j \leq J\}, \quad P_2 = \{I_{2k}; 1 \leq k \leq K\}$$

Zerlegungen von $I_1 = [a, b]$, bzw. $I_2 = [c, d]$,

$$P = P_1 \times P_2 = \{Q_{jk} = I_{1j} \times I_{2k}; 1 \leq j \leq J, 1 \leq k \leq K\}$$

die zugehörige Produktzerlegung gemäss Beispiel 9.1.1.ii) mit Feinheit

$$\delta_P \leq \sqrt{2} \max\{\delta_{P_1}, \delta_{P_2}\}.$$

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 $x_k \rightarrow x_0$ ($k \rightarrow \infty$) gilt gemäss Korollar 7.3.2

$$\begin{aligned} |g(x_k) - g(x_0)| &= \left| \int_c^d (f(x_k, y) - f(x_0, y)) \, dy \right| \\ &\leq \sup_{c \leq y \leq d} |f(x_k, y) - f(x_0, y)| \cdot |d - c| \rightarrow 0 \quad (k \rightarrow \infty), \end{aligned}$$

da f auf Q gemäss Satz 5.5.2 gleichmässig stetig ist. □

Für beliebig gewählte Punkte $x_j \in I_{1j}$, $1 \leq j \leq J$, $y_k \in I_{2k}$, $1 \leq k \leq K$, gilt nun gemäss Satz 9.1.1, bzw. Satz 7.2.3

$$\begin{aligned} \int_Q f \, d\mu &\stackrel{(\text{Satz 9.1.1})}{=} \lim_{\delta_{P_1}, \delta_{P_2} \rightarrow 0} \left(\sum_{j,k} f(x_j, y_k) \underbrace{\mu(Q_{jk})}_{=|I_{1j}| \cdot |I_{2k}|} \right) \\ &= \lim_{\delta_{P_1} \rightarrow 0} \sum_{j=1}^J \underbrace{\lim_{\delta_{P_2} \rightarrow 0} \left(\sum_{k=1}^K f(x_j, y_k) |I_{2k}| \right)}_{\stackrel{(\text{Satz 7.2.3})}{=} \int_c^d f(x_j, y) \, dy = g(x_j)} \cdot |I_{1j}| \\ &= \lim_{\delta_{P_1} \rightarrow 0} \sum_{j=1}^J g(x_j) |I_{1j}| \stackrel{(\text{Satz 7.2.3})}{=} \int_a^b g(x) \, dx = \int_a^b \left(\int_c^d f(x, y) \, dx \right) dx. \end{aligned}$$

Die 2. Identität erhält man analog nach Vertauschen von x und y . □

Beispiel 9.2.3. 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 9.2.1 erhalten wir

$$\begin{aligned} \int_Q e^{x+y} \, d\mu &= \int_0^1 \left(\int_0^1 e^{x+y} \, dy \right) dx \\ &= \int_0^1 e^x \cdot \left(\int_0^1 e^y \, dy \right) dx = \left(\int_0^1 e^x \right)^2 = (e - 1)^2. \end{aligned}$$

ii) Ebenso erhalten wir bei geschickter Wahl der Integrationsreihenfolge und Substitution

$$\begin{aligned} \int_{[0,1]^2} ye^{xy} \, d\mu &= \int_0^1 \left(\int_0^1 ye^{xy} \, dx \right) dy \\ &\stackrel{(z=xy)}{=} \int_0^1 \left(\int_0^y e^z \, dz \right) dy = \int_0^1 (e^y - 1) \, dy = e - 2. \end{aligned}$$

Satz 9.2.1 gilt analog auch in höheren Dimensionen

Satz 9.2.2. Sei $Q = \prod_{i=1}^n [a_i, b_i]$, $f \in C^0(Q)$. Dann gilt

$$\int_Q f \, d\mu = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\dots \left(\int_{a_n}^{b_n} f(x^1, \dots, x^n) \, dx^n \right) \dots \right) dx^2 \right) dx^1,$$

und die Reihenfolge der Integration darf beliebig vertauscht werden.

Beispiel 9.2.4. i) Das folgende Integral kann elementar berechnet werden

$$\begin{aligned} \int_{[0,1]^3} xy^2z^3 \, d\mu &= \int_0^1 \left(\int_0^1 \left(\int_0^1 xy^2z^3 \, dz \right) dy \right) dx \\ &= \int_0^1 x \, dx \cdot \int_0^1 y^2 \, dy \cdot \int_0^1 z^3 \, dz = \frac{1}{2} \cdot \frac{1}{3} \cdot \frac{1}{4} = \frac{1}{24}. \end{aligned}$$

ii) Im nächsten Beispiel kommt es wieder auf die geschickte Wahl der Integrationsreihenfolge an.

$$\begin{aligned} \int_{[0,\pi] \times [0,1]^2} x^2 y \cos(xyz) \, d\mu &= \int_0^\pi \left(\int_0^1 \underbrace{x^2 y \cos(xyz) \, dz}_{\stackrel{(t=xyz)}{=} \int_0^{xy} x \cos t \, dt = x \sin(xy)} \right) dy \, dx \\ &= \int_0^\pi \left(\underbrace{\int_0^1 x \sin(xy) \, dy}_{\stackrel{(s=xy)}{=} \int_0^x \sin s \, ds = 1 - \cos x} \right) dx = \int_0^\pi (1 - \cos x) \, dx = \pi - \sin \pi = \pi. \end{aligned}$$

9.3 Jordan-Bereiche

Mit dem in Abschnitt 9.1 eingeführten Integralbegriff können wir nun auch gewisse krummlinig berandete Gebiete $\Omega \subset \mathbb{R}^n$ "ausmessen".

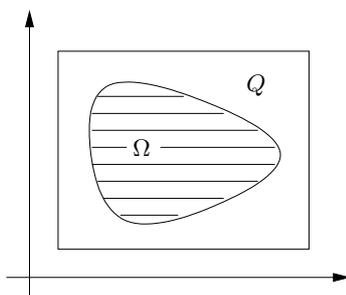
9.3.1 Das Jordansche Mass

Sei $\Omega \subset \mathbb{R}^n$ beschränkt, $Q \subset \mathbb{R}^n$ ein Quader mit $\Omega \subset Q$, χ_Ω die charakteristische Funktion von Ω .

Definition 9.3.1. Die Menge Ω heisst **Jordan-messbar**, falls χ_Ω über Q R-integrierbar ist. In diesem Fall ist

$$\mu(\Omega) = \int_Q \chi_\Omega \, d\mu$$

das n -dimensionale **Jordansche Mass** von Ω .



Bemerkung 9.3.1. i) Wegen Satz 9.1.5 ist die Definition 9.3.1 unabhängig von der Wahl von Q .

ii) Mit Satz 9.1.2 folgt für beschränkte, Jordan-messbare $\Omega_1 \subset \Omega_2 \subset \mathbb{R}^n$ die Ungleichung $\mu(\Omega_1) \leq \mu(\Omega_2)$. Das Jordansche Mass ist also monoton.

Beispiel 9.3.1. i) Jeder Quader $Q_0 \subset \mathbb{R}^n$ ist Jordan-messbar, und das Jordansche Mass von Q_0 stimmt mit dem Elementarinhalt überein.

Beweis. χ_{Q_0} ist eine Treppenfunktion, und Definition 9.1.3 ergibt für jeden beliebigen Quader $Q \supset Q_0$ den Wert $\int_Q \chi_{Q_0} d\mu = \mu(Q_0)$. \square

ii) Die Vereinigung $\Omega = \bigcup_{k=1}^K Q_k$ disjunkter Quader Q_k , $1 \leq k \leq K$, ist Jordan-messbar, und $\mu(\Omega) = \sum_{1 \leq k \leq K} \mu(Q_k)$.

Beweis. Die Funktion $\chi_\Omega = \sum_{i=1}^K \chi_{Q_k}$ ist als Treppenfunktion R-integrierbar, und $\int_Q \chi_\Omega d\mu = \sum_{1 \leq k \leq K} \mu(Q_k)$ nach Definition 9.1.3. \square

Definition 9.3.2. Ein $\Omega = \bigcup_{k=1}^K Q_k$ mit disjunkten Quadern Q_k , $1 \leq k \leq K$, heisst **Elementarfigur**.

Gemäss Beispiel 9.3.1.ii) sind Elementarfiguren Jordan-messbar.

Satz 9.3.1. Sei $\Omega \subset \mathbb{R}^n$ beschränkt. Es sind äquivalent:

i) Ω ist Jordan-messbar;

ii) Für jedes $\epsilon > 0$ gibt es Elementarfiguren $E, G \subset \mathbb{R}^n$ mit $E \subset \Omega \subset G$ und

$$\mu(G \setminus E) = \mu(G) - \mu(E) < \epsilon;$$

iii) $\partial\Omega$ ist Jordan-messbar, und $\mu(\partial\Omega) = 0$.

In jedem dieser Fälle gilt

$$\mu(\Omega) = \inf\{\mu(G); G \supset \Omega \text{ El.Fig.}\} = \sup\{\mu(E); E \subset \Omega \text{ El.Fig.}\} \quad (9.3.1)$$

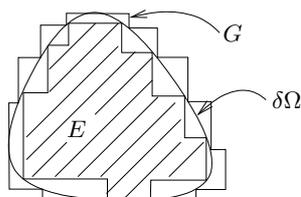
Beweis. “i) \Leftrightarrow ii)”: Zum Beweis der Messbarkeit von Ω genügt es offenbar, in Bemerkung 9.1.2.ii) Treppenfunktionen $e \leq \chi_\Omega \leq g$ mit Werten 0 oder 1 zu betrachten, also $e = \chi_E$, $g = \chi_G$ für Elementarfiguren $E \subset \Omega \subset G$. Weiter gilt in diesem Fall

$$\chi_{G \setminus E} = \chi_G - \chi_E,$$

also

$$\mu(G \setminus E) = \mu(G) - \mu(E),$$

und die Behauptung folgt aus Bemerkung 9.1.2.ii). Die Identität (9.3.1) ergibt sich analog aus Definition 9.1.3.



“ii) \Leftrightarrow iii)”: Die Menge $\partial\Omega$ ist messbar mit $\mu(\partial\Omega) = 0$ 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

$$E = \Omega \setminus U, \quad G = \Omega \cup U$$

Elementarfiguren mit $E \subset \Omega \subset G$, und $U = G \setminus E$, also

$$\mu(G \setminus E) = \mu(U) < \epsilon,$$

und umgekehrt. □

Beispiel 9.3.2. 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

$$\mu(\Omega_\psi) = \int_a^b \psi(x) dx = \int_a^b \left(\int_0^m \chi_{\Omega_\psi}(x, y) dy \right) dx,$$

wobei $m \in \mathbb{R}$ eine beliebige Zahl ist mit $m \geq \sup_{a \leq x \leq b} \psi(x)$. (Es genügt anzunehmen, dass $\psi \in C_{pw}^0([a, b])$ stückweise stetig ist.) Wir können also den Flächeninhalt von Ω berechnen, indem wir die die Längen der Schnittmengen $\Omega(x) = \{y; (x, y) \in \Omega\}$ bestimmen und bzgl. x integrieren (**“Cavalierisches Prinzip”**). Analog in höheren Dimensionen; vergleiche die Beispiele iv) und v).

Beweis. Die Funktion ψ ist gemäss Satz 7.2.2 R-integrierbar; also existieren zu vorgegebenem $\epsilon > 0$ gemäss Bemerkung 7.2.2.ii) Treppenfunktionen $e, g: [a, b] \rightarrow \mathbb{R}$ 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 Satz 9.3.1, und

$$\mu(\Omega) = \inf\{\mu(G); G \supset \Omega \text{ El.fig}\} = \overline{\int_a^b \psi dx} = \int_a^b \psi dx.$$

□

ii) Der Rhombus

$$\Omega = \{(x, y) \in \mathbb{R}^2; |x| + |y| \leq 1\}$$

ist gemäss i) Jordan-messbar, und es gilt

$$\int_{[-1,1]^2} \chi_\Omega d\mu = \int_{-1}^1 \left(\int_{-1}^1 \chi_\Omega(x, y) dy \right) dx = 4 \int_0^1 (1-x) dx = 2.$$

iii) 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

$$\mu(\Omega_\psi) = \int_{-a}^a b\sqrt{1 - \frac{x^2}{a^2}} dx \stackrel{(x=a \sin \varphi)}{=} ab \int_{-\pi/2}^{\pi/2} \cos^2 \varphi d\varphi = \frac{\pi}{2} ab.$$

Ein zu i) analoges Resultat gilt auch in höheren Dimensionen.

iv) Sei $Q' \subset \mathbb{R}^{n-1}$ ein $(n-1)$ -Quader, $\psi \in C_{pw}^0(\overline{Q'})$, $\psi \geq 0$. 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 analog zu i) folgt unter Verwendung von Satz 9.1.1

$$\mu_n(\Omega_\psi) = \int_{Q'} \psi(x') d\mu_{n-1}(x'),$$

wobei das letztere Integral mittels Satz 9.2.2 weiter umgeformt werden kann. Der Deutlichkeit halber bezeichnet hier μ_n das n -dimensionale Jordansche Mass.

v) Insbesondere erhalten wir für $Q' = [-1, 1]^2$ und mit der Funktion

$$\psi(x, y) = \sqrt{\max\{0, 1 - x^2 - y^2\}} \in C^0(Q'),$$

die obere Halbkugel

$$\Omega_\psi = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 + z^2 \leq 1, z \geq 0\}.$$

Mit iv) sowie Satz 9.2.1 folgt

$$\begin{aligned}\mu_3(\Omega_\psi) &= \int_{[-1, 1]^2} \psi(x, y) d\mu_2(x, y) = \int_{-1}^1 \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sqrt{1-x^2-y^2} dy \right) dx \\ &\stackrel{(y=\sqrt{1-x^2} \sin \varphi)}{=} \int_{-1}^1 (1-x^2) \underbrace{\left(\int_{-\pi/2}^{\pi/2} \cos^2 \varphi d\varphi \right)}_{=\pi/2} dx = \frac{2\pi}{3}.\end{aligned}$$

Beachte, dass $\frac{\pi}{2}(1-x^2)$ für jedes x wiederum dem Flächeninhalt des Schnittes $\Omega(x) = \{(y, z); (x, y, z) \in \Omega_\psi\}$ entspricht analog zu i).

vi) Die Menge $\Omega = \mathbb{Q}^n \cap [0, 1]^n$ ist nicht Jordan-messbar, da $\partial\Omega = [0, 1]^n$ positives Mass $\mu(\partial\Omega) = 1$ besitzt.

Definition 9.3.3. Eine beschränkte, Jordan-messbare Menge Ω heisst eine **Jordan-Nullmenge**, falls $\mu(\Omega) = 0$.

Bemerkung 9.3.2. Eine beschränkte Menge $\Omega \subset \mathbb{R}^n$ ist messbar mit $\mu(\Omega) = 0$ genau dann, wenn zu jedem $\epsilon > 0$ eine beschränkte Jordan-messbare Menge $\tilde{\Omega}$ existiert mit $\Omega \subset \tilde{\Omega}$ und $\mu(\tilde{\Omega}) < \epsilon$. Gemäss Satz 9.3.1 können wir nämlich zu $\epsilon > 0$ und $\tilde{\Omega}$ eine Elementarfigur $G \subset \mathbb{R}^n$ finden mit $\Omega \subset \tilde{\Omega} \subset G$ und $\mu(G) < \mu(\tilde{\Omega}) + \epsilon < 2\epsilon$, und die Behauptung folgt mit Satz 9.3.1.ii) (bei Wahl von $E = \emptyset$) und (9.3.1).

Wir stellen einige weitere Eigenschaften des Jordanschen Masses zusammen.

Satz 9.3.2. i) Sei $\Omega \subset \mathbb{R}^n$ beschränkt und Jordan-messbar, Q ein Quader mit $\Omega \subset Q$. Dann ist auch $\Omega^c := Q \setminus \Omega$ Jordan-messbar.

ii) Seien $\Omega_k \subset \mathbb{R}^n$ beschränkt und Jordan-messbar, $1 \leq k \leq K$. Dann ist $\Omega = \bigcup_{k=1}^K \Omega_k$ Jordan-messbar, und

$$\mu(\Omega) \leq \sum_{k=1}^K \mu(\Omega_k),$$

mit Gleichheit, falls $\Omega_k \cap \Omega_l = \emptyset$ für $k \neq l$.

iii) Sei Ω beschränkt und Jordan-messbar, und sei $a \in \mathbb{R}^n$. Dann ist die Menge

$$\Omega + a = \{x + a; x \in \Omega\}$$

ebenfalls Jordan-messbar, und $\mu(\Omega + a) = \mu(\Omega)$.

iv) Sei Ω beschränkt und Jordan-messbar, $R \in SO(n)$. Dann ist die Menge

$$R\Omega = \{Rx; x \in \Omega\}$$

ebenfalls Jordan-messbar, und $\mu(R\Omega) = \mu(\Omega)$.

Das heisst, das Jordansche Mass ist subadditiv und translationsinvariant sowie invariant unter Rotationen.

Beweis von Satz 9.3.2. i) Die Funktion $\chi_{\Omega^c} = \chi_Q - \chi_{\Omega} : Q \rightarrow \mathbb{R}$ ist nach Satz 9.1.3 R-integrierbar; also ist Ω^c Jordan-messbar.

ii) (Induktion nach K) $K = 2$. Mit der Darstellung

$$\chi_{\Omega_1 \cup \Omega_2} = \max\{\chi_{\Omega_1}, \chi_{\Omega_2}\} \leq \chi_{\Omega_1} + \chi_{\Omega_2} \quad (9.3.2)$$

und den Sätzen 9.1.2 - 9.1.4 folgt, dass $\Omega_1 \cup \Omega_2$ Jordan-messbar ist und

$$\mu(\Omega_1 \cup \Omega_2) \leq \mu(\Omega_1) + \mu(\Omega_2), \quad (9.3.3)$$

mit Gleichheit in (9.3.2) (und dann auch in (9.3.3)), falls $\Omega_1 \cap \Omega_2 = \emptyset$.

$K \mapsto K + 1$: Setze

$$\tilde{\Omega}_1 = \bigcup_{k=1}^K \Omega_k, \quad \tilde{\Omega}_2 = \Omega_{K+1}$$

und verfähre wie oben.

iii) Die Behauptung gilt für Elementarfiguren. Mit Satz 9.3.1 erhalten wir die Aussage dann auch für beliebige beschränkte Jordan-messbare Mengen.

iv) Zum Beweis der Rotationsinvarianz des Masses benötigen wir ein Lemma. □

Lemma 9.3.1. *Sei $A \in Gl(n)$ ein linearer Isomorphismus. Es gibt eine Zahl $m_A > 0$ mit folgende Eigenschaft: Für jedes beschränkte und Jordan-messbare $\Omega \subset \mathbb{R}^n$ ist die Menge $A\Omega = \{Ax; x \in \Omega\}$ Jordan-messbar, und $\mu(A\Omega) = m_A\mu(\Omega)$.*

Beweis. Sei Ω beschränkt und Jordan-messbar. Zu gegebenem $\epsilon > 0$ können wir Ω durch Elementarfiguren $E \subset \Omega^\circ \subset \Omega \subset \overline{\Omega} \subset G$ mit $\mu(G \setminus E) < \epsilon$ einschachteln. Indem wir \mathbb{R}^n mittels eines Gitters der Kantenlänge $\delta_0 = \delta_0(\epsilon)$ in Würfel zerlegen, können wir annehmen, dass $E = \cup_{1 \leq k \leq K_E} Q_k$, $G = \cup_{1 \leq k \leq K_G} Q_k$ disjunkte Vereinigungen von achsenparallelen, zu einem Würfel Q_0 der Kantenlänge δ_0 kongruenten Würfeln Q_k sind.

Das Bild AQ_k eines Würfels kann offenbar durch achsenparallele Hyperebenen durch die Eckpunkte von AQ_k in endlich viele Mengen vom Typ Ω_ψ mit stetigem ψ zerlegt werden, ist also messbar. Zudem gehen die Bilder AQ_k durch Verschiebung aus dem Bild AQ_0 hervor, haben nach Satz 9.3.2.iii) also dasselbe Mass

$$\mu(AQ_k) = \mu(AQ_0) = m_A\mu(Q_0)$$

mit der Zahl $m_A := m_A(Q_0) := \mu(AQ_0)/\mu(Q_0) > 0$.

Da A ein Isomorphismus, ist AE die disjunkte Vereinigung der Bilder AQ_k , $1 \leq k \leq K_E$; analog AG und $A(G \setminus E) = AG \setminus AE$. Es folgt dann mit Satz 9.3.2.ii) auch

$$\mu(AE) = m_A\mu(E), \quad \mu(AG) = m_A\mu(G), \tag{9.3.4}$$

sowie

$$\mu(AG \setminus AE) = \mu(A(G \setminus E)) = m_A\mu(G \setminus E) < m_A\epsilon.$$

Angewandt auf den Würfel $E = NQ_0 =: R_0$, $N \in \mathbb{N}$, ergibt (9.3.4) die Gleichung $m_A(R_0) := \mu(AR_0)/\mu(R_0) = m_A = m_A(Q_0)$. Die Zahl m_A ist also für jede Kantenlänge $N\delta_0$ dieselbe, nach Ersetzen von δ_0 durch δ_0/N dann auch für jede Kantenlänge δ_0/N , $N \in \mathbb{N}$. Insbesondere ist die Zahl m_A von ϵ unabhängig.

Da $\epsilon > 0$ beliebig und da

$$\partial(A\Omega) = A(\partial\Omega) \subset A(G \setminus E) = AG \setminus AE,$$

ist $\partial(A\Omega)$ somit gemäss Bemerkung 9.3.2 eine Nullmenge, und die Menge $A\Omega$ ist nach Satz 9.3.1 Jordan-messbar.

Schliesslich folgt mit $AE \subset A\Omega \subset AG$ auch

$$m_A\mu(E) = \mu(AE) \leq \mu(A\Omega) \leq \mu(AG) = m_A\mu(G),$$

und mit

$$(\mu(G) - \mu(\Omega)) + (\mu(\Omega) - \mu(E)) = \mu(G \setminus E) < \epsilon$$

erhalten wir

$$|\mu(A\Omega) - m_A\mu(\Omega)| < m_A\epsilon;$$

da $\epsilon > 0$ beliebig, also $\mu(A\Omega) = m_A\mu(\Omega)$. □

Beweis von Satz 9.3.2, vollendet. Falls $A = R \in SO(n)$ in Lemma 9.3.1, so ergibt die Wahl $\Omega = B_1(0)$ mit $RB_1(0) = B_1(0)$ den Wert $m_R = 1$, und damit $\mu(R\Omega) = \mu(\Omega)$ für jedes beschränkte, Jordan-messbare $\Omega \subset \mathbb{R}^n$. \square

Definition 9.3.4. Ein Gebiet $\Omega \subset \mathbb{R}^n$ ist von der Klasse C^1 , (bzw. von der Klasse C_{pw}^1, C^k), falls zu jedem Punkt $x_0 \in \partial\Omega$ Koordinaten $(x', x^n) \in \mathbb{R}^{n-1} \times \mathbb{R}$ um $x_0 = (0, x_0^n)$, eine Zahl $d > 0$, ein offener Quader $Q' \subset \mathbb{R}^{n-1}$ um $x'_0 = 0$ und eine Funktion $\psi \in C^1(\overline{Q}')$, (bzw. $\psi \in C_{pw}^1(\overline{Q}')$, $\psi \in C^k(\overline{Q}')$), existieren mit $0 \leq \psi \leq 2d$, wobei $\psi(0) = d = x_0^n$, so dass

$$\Omega \cap (Q' \times [0, 2d]) = \{(x', x^n) \in \mathbb{R}^n; x' \in Q', 0 \leq x^n < \psi(x')\} = \Omega_\psi.$$

Beispiel 9.3.3. i) $B_1(0) \subset \mathbb{R}^2$ ist von der Klasse C^k für beliebiges $k \geq 0$.

ii) Ein n -Quader Q ist von der Klasse C_{pw}^1 .

iii) Falls $\partial\Omega$ eine $(n-1)$ -dimensionale Untermannigfaltigkeit der Klasse C^k ist, $k \geq 1$, so ist Ω gemäss Bemerkung 8.10.1.ii) von der Klasse C^k .

Es folgt:

Satz 9.3.3. Sei $\Omega \subset \mathbb{R}^n$ beschränkt und von der Klasse C^k mit $k \geq 0$. Dann ist Ω Jordan-messbar.

Beweis. Gemäss Beispiel 9.3.3 kann man Ω als Vereinigung von Mengen der Gestalt Ω_ψ bzgl. geeignet gewählter Achsen mit $\psi \in C^0$ darstellen. Somit folgt die Behauptung aus Satz 9.3.2. \square

9.3.2 Das Riemann-Integral über Jordan-Bereiche

Sei $\Omega \subset \mathbb{R}^n$ beschränkt und Jordan-messbar, $f: \Omega \rightarrow \mathbb{R}$ beschränkt. Weiter sei $Q \subset \mathbb{R}^n$ ein Quader mit $\Omega \subset Q$, und sei $\overline{f}: Q \rightarrow \mathbb{R}$ mit $\overline{f}|_\Omega = f$.

Definition 9.3.5. f heisst **R-integrabel über Ω** , falls die Funktion $\overline{f}\chi_\Omega$ über Q R-integrabel ist, und

$$\int_\Omega f \, d\mu := \int_Q \overline{f}\chi_\Omega \, d\mu.$$

Bemerkung 9.3.3. i) Die Definition ist unabhängig von der Fortsetzung \overline{f} und wegen Satz 9.1.5 auch unabhängig von Q . Insbesondere können wir die Fortsetzung $\overline{f}(x) = 0$ für $x \in Q \setminus \Omega$ wählen.

ii) Korollar 9.1.1 liefert für beschränktes, Jordan-messbares $\Omega \subset Q \subset \mathbb{R}^n$ und beschränktes, R-integrables $f: \Omega \rightarrow \mathbb{R}$ mit der Fortsetzung $\overline{f}(x) = 0$ für $x \in Q \setminus \Omega$ die zu Korollar 9.1.1 analoge Abschätzung

$$\left| \int_\Omega f \, d\mu \right| = \left| \int_Q \overline{f} \, d\mu \right| \leq \int_Q |\overline{f}| \, d\mu = \int_\Omega |f| \, d\mu \leq \int_Q m\chi_\Omega \, d\mu = m\mu(\Omega),$$

wobei $m = \sup_\Omega |f| = \sup_Q |\overline{f}| < \infty$.

iii) Sei Q ein Quader in \mathbb{R}^n , und sei $f: Q \rightarrow \mathbb{R}$ beschränkt und R-integrierbar. Weiter sei $\Omega \subset Q$ Jordan-messbar. Dann ist f über Ω R-integrierbar.

Beweis. Sei $m = \sup_Q |f| < \infty$, $s_{\pm} = \max\{\pm s, 0\}$, $s \in \mathbb{R}$. Gemäss Satz 9.1.4 ist $(f\chi_{\Omega})_+ = \min\{f_+, m\chi_{\Omega}\}$ R-integrierbar über Q und ebenso $(f\chi_{\Omega})_-$, also auch $f\chi_{\Omega} = (f\chi_{\Omega})_+ - (f\chi_{\Omega})_-$. \square

Satz 9.3.4. (Gebietsadditivität) Sei $f: \Omega \rightarrow \mathbb{R}$ beschränkt und über Ω R-integrierbar, und sei $\Omega = \cup_{1 \leq k \leq K} \Omega_k$ eine Zerlegung von Ω in disjunkte, messbare Ω_k , $1 \leq k \leq K$. Dann gilt

$$\int_{\Omega} f \, d\mu = \sum_{k=1}^K \int_{\Omega_k} f \, d\mu.$$

Beweis. Zerlege $f = \sum_{1 \leq k \leq K} f\chi_{\Omega_k}$ und benutze Bemerkung 9.3.3.iii) sowie Satz 9.1.3. \square

Satz 9.3.5. Falls Ω eine Jordan-Nullmenge ist, $f: \Omega \rightarrow \mathbb{R}$ beschränkt, so ist f über Ω R-integrierbar, und $\int_{\Omega} f \, d\mu = 0$.

Beweis. Für R-integrierbares f folgt die Aussage direkt aus Bemerkung 9.3.3.ii).

Für den allgemeinen Fall sei Q ein Quader mit $\Omega \subset Q$, und sei $\bar{f}: Q \rightarrow \mathbb{R}$ die Fortsetzung von f durch $\bar{f}(x) = 0$ für $x \in Q \setminus \Omega$. Sei weiter

$$m = \sup_{\Omega} |f| = \sup_Q |\bar{f}| < \infty.$$

Zu $\epsilon > 0$ sei $G \subset Q$ eine Elementarfigur mit $\Omega \subset G$ und $\mu(G) < \epsilon$ gemäss Satz 9.3.1. Setze $e = -m\chi_G$, $g = m\chi_G$. Dann sind e, g Treppenfunktionen auf Q mit $e \leq \bar{f} \leq g$, und

$$-m\epsilon \leq \int_Q e \, d\mu \leq \int_Q g \, d\mu \leq m\epsilon.$$

Grenzübergang $\epsilon \downarrow 0$ ergibt die Behauptung. \square

Zusammen mit Satz 9.3.4 folgt aus Satz 9.3.5, dass Jordan-Nullmengen $\Omega \subset \mathbb{R}^n$ bei der Integration vernachlässigt werden können.

Mit Satz 9.2.1, bzw. Satz 9.2.2 können wir für Mengen $\Omega = \Omega_{\psi}$ wie in Beispiel 9.3.2.i) und iv) das Integral aus Definition 9.3.4 auf iterierte 1-dimensionale Integrale zurückführen.

Beispiel 9.3.4. i) Sei $0 \leq \psi \in C^0([a, b])$, $\Omega_{\psi} \subset \mathbb{R}^2$ die Menge

$$\Omega_{\psi} = \{(x, y) \in \mathbb{R}^2; a \leq x \leq b, 0 \leq y \leq \psi(x)\},$$

und sei $f \in C^0(\overline{\Omega_{\psi}})$. Dann ist f über Ω_{ψ} R-integrierbar, und es gilt

$$\int_{\Omega_{\psi}} f \, d\mu = \int_a^b \left(\int_0^{\psi(x)} f(x, y) \, dy \right) dx.$$

Beweis. Die Aussage gilt für Treppenfunktionen ψ gemäss Satz 9.2.1.

Für allgemeines $\psi \in C^0([a, b])$ betrachte eine Folge von Zerlegungen

$$P^{(l)} = \{I_k^{(l)}; 1 \leq k \leq K^{(l)}\}$$

von $I = [a, b]$ mit Feinheit $\delta_{P^{(l)}} \rightarrow 0$ ($l \rightarrow \infty$) und setze

$$\psi^{(l)} = \sum_{k=1}^{K^{(l)}} c_k^{(l)} \chi_{I_k^{(l)}}, \text{ wobei } c_k^{(l)} = \inf_{I_k^{(l)}} \psi \geq 0, 1 \leq k \leq K^{(l)}, l \in \mathbb{N}.$$

Dann gilt offenbar $0 \leq \psi^{(l)} \leq \psi$, und $\Omega_{\psi^{(l)}} \subset \Omega_\psi$, $l \in \mathbb{N}$, mit $\mu(\Omega_\psi \setminus \Omega_{\psi^{(l)}}) \rightarrow 0$ ($l \rightarrow \infty$). Mit ähnlichen Argumenten wie im Beweis von Satz 9.3.5 folgt, dass

$$\left| \int_{\Omega_\psi} f \, d\mu - \int_{\Omega_{\psi^{(l)}}} f \, d\mu \right| \leq m \mu(\Omega_\psi \setminus \Omega_{\psi^{(l)}}) \rightarrow 0 \quad (l \rightarrow \infty),$$

wobei $m = \sup_{\Omega_\psi} |f|$, und analog für das untere Riemann-Integral. Andererseits erhalten wir gleichmässig in $x \in [a, b]$ die Abschätzung

$$\left| \int_0^{\psi^{(l)}(x)} f(x, y) \, dy - \int_0^{\psi(x)} f(x, y) \, dy \right| \leq m \sup_{a \leq x \leq b} |\psi(x) - \psi^{(l)}(x)| \xrightarrow{(l \rightarrow \infty)} 0,$$

und

$$\int_{\Omega_{\psi^{(l)}}} f \, d\mu = \int_a^b \left(\int_0^{\psi^{(l)}(x)} f(x, y) \, dy \right) dx \xrightarrow{(l \rightarrow \infty)} \int_a^b \left(\int_0^{\psi(x)} f(x, y) \, dy \right) dx.$$

Die gewünschte Aussage folgt. \square

ii) Speziell für $\psi(x) = b\sqrt{1 - \frac{x^2}{a^2}} \in C^0([-a, a])$ wie in Beispiel 9.3.2.iii) und die Funktion $f(x, y) = y$ erhalten wir

$$\int_{\Omega_\psi} f \, d\mu = \int_{-a}^a \int_0^{b\sqrt{1 - \frac{x^2}{a^2}}} y \, dy = \frac{1}{2} \int_{-a}^a b^2 \left(1 - \frac{x^2}{a^2}\right) dx = \frac{2ab^2}{3}.$$

9.4 Der Satz von Green

Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C_{pw}^1 . 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 9.4.1. i) Sei $Q = [a, b] \times [c, d]$, $g \in C^1(Q)$. Mit Satz 9.2.1 folgt

$$-\int_Q \frac{\partial g}{\partial y} \, d\mu = \int_a^b \left(-\int_c^d \frac{\partial g}{\partial y}(x, y) \, dy \right) dx = \int_a^b (g(x, c) - g(x, d)) \, dx.$$

Analog erhalten wir für $h \in C^1(Q)$ die Gleichung

$$\int_Q \frac{\partial h}{\partial x} d\mu = \int_c^d \left(\int_a^b \frac{\partial h}{\partial x}(x, y) dx \right) dy = \int_c^d (h(b, y) - h(a, y)) dy;$$

also

$$\int_Q \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu = \int_{\gamma_1 + \gamma_2 - \gamma_3 - \gamma_4} (gdx + hdy) = \int_{\partial Q} (gdx + hdy),$$

wobei die Kurven

$$\gamma_1(x) = (x, c), \quad \gamma_3(x) = (x, d), \quad a \leq x \leq b,$$

sowie

$$\gamma_2(y) = (b, y), \quad \gamma_4(y) = (a, y), \quad c \leq y \leq d,$$

den Rand von ∂Q parametrisieren. Die Teilstücke $\gamma_1, \dots, \gamma_4$ 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

$$\mu(Q) = - \int_Q \frac{\partial g}{\partial y} d\mu = \int_a^b (g(x, c) - g(x, d)) dx = (b - a)(d - c)$$

Analog können wir für eine grosse Klasse von Gebieten argumentieren.

Definition 9.4.1. i) $\Omega \subset \mathbb{R}^2$ heisst **Normalbereich bzgl. y** der Klasse C^1 , falls

$$\Omega = \{(x, y) \in \mathbb{R}^2; a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\}$$

für geeignete $-\infty < a < b < \infty$ und mit Funktionen $\varphi \leq \psi \in C^1([a, b])$. Analog definieren wir einen Normalbereich bzgl. x , oder von der Klasse C_{pw}^1, C^k , etc.

ii) $\Omega \subset \mathbb{R}^2$ heisst **Normalbereich**, falls Ω sowohl bzgl. x als auch bzgl. y ein Normalbereich ist.

Beispiel 9.4.2. i) Ein bzgl. der Achsen gedrehter Quader ist ein Normalbereich der Klasse C_{pw}^1 .

ii) $\overline{B_1(0)} = \{(x, y) \in \mathbb{R}^2; |y| \leq \sqrt{1 - x^2}, |x| \leq 1\} \subset \mathbb{R}^2$ ist ein Normalbereich der Klasse C^0 .

iii) Der Kreisring $B_2 \setminus B_1(0)$ ist kein Normalbereich.

Die folgende elementare Beobachtung wird später entscheidend benutzt.

Bemerkung 9.4.1. i) Mit Beispiel 9.3.4 erhalten wir für jeden Normalbereich

$$\Omega = \{(x, y) \in \mathbb{R}^2; a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\}$$

und $f \in C^0(\overline{\Omega})$ die Identität

$$\int_{\Omega} f \, d\mu = \int_a^b \left(\int_{\varphi(x)}^{\psi(x)} f(x, y) \, dy \right) dx.$$

ii) Jedes beschränkte Gebiet $\Omega \subset \mathbb{R}^2$ der Klasse C_{pw}^1 kann man in endlich viele disjunkte Gebiete $\Omega_1, \dots, \Omega_L \in C_{pw}^1$ zerlegen, wobei jedes Ω_l ein Normalbereich ist bzgl. geeigneter Achsen.

iii) Selbst für $\Omega \in C^1$ sind die Gebiete Ω_l in der Regel nur von der Klasse C_{pw}^1 .

Beispiel 9.4.3. i) $\Omega = B_1(0)$ und

ii) $\Omega = B_2 \setminus B_1(0)$ sind in endlich viele disjunkte Normalbereiche zerlegbar.

Satz 9.4.1. (Green) Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C_{pw}^1 , und seien $g, h \in C^1(\overline{\Omega})$. Dann gilt

$$\int_{\Omega} \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu = \int_{\partial\Omega} (g dx + h dy),$$

wobei der Rand von Ω so parametrisiert wird, dass Ω zur Linken liegt.

Beweis. i) Sei zunächst Ω ein Normalbereich, insbesondere also von der Form

$$\Omega = \{(x, y); a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\},$$

wobei $a < b$, $\varphi \leq \psi \in C_{pw}^1([a, b])$, und sei $h = 0$. Mit Bemerkung 9.4.1.i) folgt

$$\begin{aligned} \int_{\Omega} \left(-\frac{\partial g}{\partial y} \right) d\mu &= \int_a^b \left(- \int_{\varphi(x)}^{\psi(x)} \frac{\partial g}{\partial y}(x, y) \, dy \right) dx \\ &= \int_a^b (g(x, \varphi(x)) - g(x, \psi(x))) \, dx. \end{aligned}$$

Parametrisiere $\partial\Omega = \gamma_1 + \gamma_2 - \gamma_3 - \gamma_4$, wobei

$$\begin{aligned} \gamma_1(x) &= (x, \varphi(x)), \quad a \leq x \leq b, \quad \gamma_2(y) = (b, y), \quad \varphi(b) \leq y \leq \psi(b), \\ \gamma_3(x) &= (x, \psi(x)), \quad a \leq x \leq b, \quad \gamma_4(y) = (a, y), \quad \varphi(a) \leq y \leq \psi(a). \end{aligned}$$

Dann gilt

$$\int_{\gamma_1 + \gamma_2 - \gamma_3 - \gamma_4} g \, dx = \int_{\gamma_1} g \, dx - \int_{\gamma_3} g \, dx = \int_a^b g(x, \varphi(x)) \, dx - \int_a^b g(x, \psi(x)) \, dx;$$

das heisst,

$$\int_{\Omega} \left(-\frac{\partial g}{\partial y} \right) d\mu = \int_{\partial\Omega} g \, dx.$$

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

$$\int_{\Omega} = \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu = \int_{\partial\Omega} (gdx + hdy),$$

wie gewünscht.

iii) Zerlege $\Omega = \Omega_1 \cup \dots \cup \Omega_L$ 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

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu &= \sum_{i=1}^L \int_{\Omega_i} \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu \\ &= \sum_{i=1}^L \int_{\partial\Omega_i} (gdx + hdy) = \int_{\partial\Omega} (gdx + hdy). \end{aligned}$$

□

Beispiel 9.4.4. i) Mit $g(x, y) = -y, h(x, y) = 0$ erhalten wir wie in Beispiel 9.4.1.ii) für ein beliebiges beschränktes Gebiet $\Omega \subset \mathbb{R}^2$ von der Klasse C_{pw}^1 den Flächeninhalt

$$\mu(\Omega) = \int_{\Omega} \left(-\frac{\partial g}{\partial y} \right) d\mu = \int_{\partial\Omega} g dx = - \int_{\partial\Omega} y dx.$$

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

$$\mu(\Omega) = \frac{1}{2} \int_{\partial\Omega} (x dy - y dx).$$

ii) Für $\Omega = B_1(0)$ mit der Parametrisierung $\gamma(t) = (\cos t, \sin t)^t, 0 \leq t \leq 2\pi$, des Randes ergibt i) den Wert

$$\mu(B_1(0)) = \frac{1}{2} \int_{\partial\Omega} (x dy - y dx) = \frac{1}{2} \int_0^{2\pi} (-\sin t, \cos t) \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = \pi.$$

Definition 9.4.2. Sei $\Omega \subset \mathbb{R}^2$ offen, und sei $\lambda = gdx + hdy$ eine 1-Form auf Ω der Klasse C^1 . Dann heisst

$$d\lambda := \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) d\mu$$

die **äussere Ableitung** von λ , mit der **Volumenform** $d\mu = dx dy =: dx \wedge dy$ von \mathbb{R}^2 .

Beispiel 9.4.5. i) Sei $\lambda = \frac{1}{2}(x dy - y dx)$. Dann gilt $d\lambda = d\mu$.

ii) Sei $\Omega \subset \mathbb{R}^2$ offen, und sei $f \in C^2(\Omega)$, $\lambda = df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$. Dann gilt

$$d\lambda = d^2f = \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) d\mu = 0.$$

Mit Definition 9.4.2 können wir Satz 9.4.1 nun auch wie folgt formulieren.

Satz 9.4.2. Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C_{pw}^1 , und sei $\lambda = gdx + hdy$ eine 1-Form auf Ω der Klasse $C^1(\overline{\Omega})$. Dann gilt

$$\int_{\Omega} d\lambda = \int_{\partial\Omega} \lambda,$$

wobei $\partial\Omega$ so orientiert durchlaufen wird, dass Ω zur Linken liegt.

Statt als Koeffizienten einer 1-Form können wir die Funktionen g und h in Satz 9.4.1 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

$$\text{rot } v := \frac{\partial h}{\partial x} - \frac{\partial g}{\partial y},$$

so nimmt Satz 9.4.1 die folgende Gestalt an.

Satz 9.4.3. (Stokes) Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C_{pw}^1 , und sei $v \in C^1(\overline{\Omega}; \mathbb{R}^2)$. Dann gilt

$$\int_{\Omega} \text{rot } v \, d\mu = \int_{\partial\Omega} v \cdot \vec{ds},$$

wobei $\partial\Omega$ so orientiert durchlaufen wird, dass Ω zur Linken liegt.

Bemerkung 9.4.2. Der Ausdruck $\text{rot } v$ heisst **Wirbelstärke** von v , das Randintegral

$$\int_{\partial\Omega} v \cdot \vec{ds}$$

die **Zirkulation** von v längs $\partial\Omega$.

Die Orientierung von $\partial\Omega$ kann man wie folgt präzisieren. Sei \vec{n} der nach aussen zeigende Normalenvektor längs $\partial\Omega$ und $\tau = \dot{\gamma}/|\dot{\gamma}|$ der durch eine Parametrisierung γ von $\partial\Omega$ mit $\dot{\gamma} \neq 0$ induzierte Einheitstangentenvektor.

Definition 9.4.3. i) Eine Orthonormalbasis (v, w) von \mathbb{R}^2 heisst **positiv orientiert** oder ein **Rechtssystem**, falls (v, w) durch eine Drehung aus der Standardbasis (e_1, e_2) hervorgeht, sonst heisst (v, w) **negativ orientiert**.

ii) Eine Parametrisierung γ von $\partial\Omega$ heisst **positiv orientiert**, falls in jedem Punkt $p \in \partial\Omega$ das Paar $(\vec{n}(p), \tau(p))$ eine positiv orientierte Orthonormalbasis von \mathbb{R}^2 bildet.

Der Rand in Satz 9.4.2 oder 9.4.3 ist also positiv orientiert zu verstehen.

Satz 9.4.3 liefert insbesondere ein Kriterium für konservative Kraftfelder.

Definition 9.4.4. 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 9.4.6. i) $B_1(0)$ ist einfach zusammenhängend.

ii) $B_2(0) \setminus B_1(0)$ ist nicht einfach zusammenhängend.

Bemerkung 9.4.3. i) 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; also $\gamma(s) \neq \gamma(t)$ für alle $0 \leq s < t < 1$. Nach dem "Jordanschen Kurvensatz" berandet γ ein beschränktes Gebiet $\Omega_\gamma \subset \mathbb{R}^2$, und Ω_γ ist eindeutig bestimmt. Die Menge Ω ist nun genau dann einfach zusammenhängend, wenn die Relation $\Omega_\gamma \subset \Omega$ für alle derartigen Kurven gilt. Beispiel 9.4.6 illustriert diesen Zusammenhang auf einfache Weise.

ii) Auch in höheren Dimensionen $n \geq 3$ ist der Begriff einer einfach zusammenhängenden Menge von grosser Bedeutung, und es gibt eine zu i) analoge Charakterisierung. Die Definition 9.4.4 gilt jedoch nur im Fall $n = 2$.

Satz 9.4.4. (Poincaré) Sei $\Omega \subset \mathbb{R}^2$ von der Klasse C^1_{pw} , beschränkt, zusammenhängend sowie einfach zusammenhängend, und sei $v \in C^1(\bar{\Omega}; \mathbb{R}^2)$. Dann sind äquivalent

i) v ist konservativ,

ii) $\text{rot } v = 0$.

Beweis. $i) \Rightarrow ii)$ Nach Satz 8.4.3 besitzt jedes konservative Vektorfeld $v \in C^1(\bar{\Omega}; \mathbb{R}^2)$ ein Potential $f \in C^2(\bar{\Omega})$ mit $v = \nabla f$, und

$$\text{rot } \nabla f = \frac{\partial(\partial f / \partial y)}{\partial x} - \frac{\partial(\partial f / \partial x)}{\partial y} = \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} = 0$$

nach Satz 8.5.1.

$ii) \Rightarrow i)$ Sei $\gamma \in C^1([0, 1]; \Omega)$ ein geschlossener Weg in Ω . OBdA sei γ ohne Selbstschnitte, und sei Ω_γ das von γ berandete Gebiet. Nach Bemerkung 9.4.3.i) gilt $\Omega_\gamma \subset \Omega$, und die Annahme $\text{rot } v = 0$ zusammen mit Satz 9.4.3 ergibt

$$\int_{\gamma} v \cdot d\vec{s} = \int_{\Omega_\gamma} \text{rot } v \, d\mu = 0,$$

wie gewünscht. □

Beispiel 9.4.7. i) Sei $v(x, y) = (y + \sin x, x + \cos y)^t \in C^1(\mathbb{R}^2; \mathbb{R}^2)$ mit

$$\text{rot } v = -\frac{\partial(y + \sin x)}{\partial y} + \frac{\partial(x + \cos y)}{\partial x} = 0.$$

Jede Kugel $B_R(0) \subset \mathbb{R}^2$ ist einfach zusammenhängend. Also ist v konservativ nach Satz 9.4.4. Die Funktion

$$f(x, y) = xy + \sin y - \cos x, \quad (x, y) \in \mathbb{R}^2,$$

ist ein Potential für v .

ii) Sei $v(x, y) = \frac{1}{x^2+y^2} \cdot (-y, x)^t \in C^1(\mathbb{R}^2 \setminus \{0\}; \mathbb{R}^2)$. Es gilt

$$\operatorname{rot} v = \frac{2}{x^2+y^2} - \frac{2(x^2+y^2)}{(x^2+y^2)^2} = 0,$$

jedoch ist $B_R(0) \setminus \{0\}$ nicht einfach zusammenhängend für $R > 0$.

In der Tat gilt für $\gamma(t) = (\cos t, \sin t)^t$, $0 \leq t \leq 2\pi$:

$$\int_{\gamma} v \cdot \vec{ds} = \int_0^{2\pi} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} dt = 2\pi \neq 0;$$

das heisst, v ist **nicht** konservativ.

Auf der einfach zusammenhängenden Menge $\mathbb{R}^2 \setminus \{(x, 0); x \leq 0\}$ besitzt jedoch v die **Argumentfunktion**

$$f(x, y) = \arctan\left(\frac{y}{x}\right)$$

als Potential. Weiter gilt auf jedem einfach geschlossenen stückweise C^1 -Weg $\tilde{\gamma}: [0, 1] \rightarrow \mathbb{R}^2 \setminus \{0\}$ mit $0 \notin \Omega_{\tilde{\gamma}}$ nach Satz 9.4.3

$$\int_{\tilde{\gamma}} v \cdot \vec{ds} = \int_{\Omega_{\tilde{\gamma}}} \operatorname{rot} v \, d\mu = 0.$$

Wir beschliessen den Abschnitt mit einer weiteren Interpretation von Satz 9.4.1. Sei $\Omega \subset \mathbb{R}^2$ offen, $v = (g, h)^t \in C^1(\Omega; \mathbb{R}^2)$, und setze

$$\operatorname{div} v := \left(\frac{\partial g}{\partial x} + \frac{\partial h}{\partial y} \right).$$

Die 1-Form $\lambda = -hdx + gdy$ ist dann von der Klasse C^1 mit

$$d\lambda = \operatorname{div} v \, d\mu,$$

und Satz 9.4.1 ergibt:

Satz 9.4.5. (Gauss) Sei $\Omega \subset \mathbb{R}^2$ beschränkt und von der Klasse C_{pw}^1 , und sei $v \in C^1(\overline{\Omega}; \mathbb{R}^2)$. Dann gilt

$$\int_{\Omega} \operatorname{div} v \, d\mu = \int_{\partial\Omega} v \cdot \vec{n} \, ds,$$

wobei ds das skalare Längenelement längs $\partial\Omega$ bezeichnet und \vec{n} die äussere Normale.

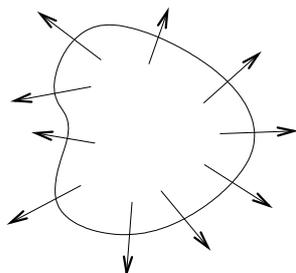
Bemerkung 9.4.4. i) Für $v = (v^1, v^2)^t \in C^1(\overline{\Omega}; \mathbb{R}^2)$ gilt:

$$\operatorname{div} v = \sum_{i=1}^2 \frac{\partial v^i}{\partial x^i}.$$

ii) Der Ausdruck $\operatorname{div} v$ heisst **Quellstärke** des Vektorfeldes v , das Integral

$$\int_{\partial\Omega} v \cdot \vec{n} \, ds$$

der **Fluss** von v durch $\partial\Omega$.



Beispiel 9.4.8. Sei $\Omega = B_1(0)$, $v(x, y) = (x, y)$ das Ortsvektorfeld mit $\operatorname{div} v = 2$. Da

$$\vec{n}(p) = p = v(p) \text{ für alle } p \in S^1 = \partial B_1(0),$$

erhalten wir mit der positiv orientierten Parametrisierung

$$\gamma(t) = (\cos(t), \sin(t))^t \in C^1([0, 2\pi]; \mathbb{R}^2)$$

von $\partial B_1(0)$ die Identität

$$2\mu(B_1(0)) = \int_{\Omega} \operatorname{div} v \, d\mu = \int_{\partial\Omega} v \cdot \vec{n} \, ds = \int_{\partial\Omega} ds = \int_0^{2\pi} |\dot{\gamma}(t)| dt = 2\pi;$$

also $\mu(B_1(0)) = \pi$; vergleiche Beispiel 9.4.4.ii).

9.5 Substitutionsregel

Gibt es eine zur Substitutionsregel Satz 7.1.5 analoge Regel in \mathbb{R}^n ? Wird insbesondere eine messbare Menge $\Omega \subset \mathbb{R}^n$ bei "Transformation" mit einer Abbildung $\Phi \in C^1(\overline{\Omega}; \mathbb{R}^n)$ wieder in eine messbare Menge $\Phi(\Omega)$ überführt?

9.5.1 Lineare Transformationen

Wir betrachten zunächst lineare Abbildungen.

Lemma 9.5.1. Sei $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ linear, und sei $\Omega \subset \mathbb{R}^n$ beschränkt und Jordan-messbar. Dann ist die Menge $A\Omega = \{Ax; x \in \Omega\}$ Jordan-messbar, und

$$\mu(A\Omega) = |\det(A)|\mu(\Omega).$$

Beweis. i) Betrachte zunächst den Fall $\det(A) = 0$. Dann gibt es eine Drehung $R \in SO(n)$, so dass $RA(\mathbb{R}^n) \subset \mathbb{R}^{n-1} \times \{0\}$. Falls $\Omega \subset \mathbb{R}^n$ beschränkt, so ist

$RA(\Omega) \subset \mathbb{R}^{n-1} \times \{0\}$ beschränkt und Jordan-messbar mit $\mu(RA(\Omega)) = 0$; also ist nach Satz 9.3.2.iv) auch $A\Omega = R^{-1}(RA(\Omega))$ Jordan-messbar mit

$$\mu(A\Omega) = \mu(RA(\Omega)) = 0 = |\det(A)|\mu(\Omega).$$

ii) Sei nun $A = \text{diag}(\lambda_1, \dots, \lambda_n)$ diagonal mit Eigenwerten $\lambda_i \neq 0$. Nach Lemma 9.3.1 ist für jede beschränkte Jordan-messbare Menge $\Omega \subset \mathbb{R}^n$ die Menge $A\Omega$ Jordan-messbar, und $\mu(A\Omega) = m_A\mu(\Omega)$ mit einer von Ω unabhängigen Konstanten $m_A > 0$.

Wählen wir insbesondere $\Omega = Q = [0, 1]^n$, so erhalten wir $AQ = \prod_{i=1}^n [0, \lambda_i]$ mit

$$\mu(AQ) = \prod_{i=1}^n |\lambda_i| = |\det(A)| = m_A\mu(Q) = m_A,$$

und $m_A = |\det(A)|$.

iii) Schliesslich sei $A \in Gl(n)$ eine beliebige invertierbare Matrix. Die symmetrische Matrix $M = A^t A$ ist wegen

$$\xi^t M \xi = |A\xi|^2 > 0 \text{ für jedes } 0 \neq \xi \in \mathbb{R}^n$$

positiv definit. Also existieren eine Matrix $R \in SO(n)$ und eine Diagonalmatrix $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ mit Eigenwerten $\lambda_i = \omega_i^2 > 0$, wobei $\omega_i > 0$, $1 \leq i \leq n$, so dass

$$M = R\Lambda R^t = RD(RD)^t, \text{ wobei } D = \text{diag}(\omega_1, \dots, \omega_n) = D^t.$$

Dann erfüllt die Matrix $B = ARD^{-1}$ die Bedingung $B^t B = id_{\mathbb{R}^n}$; das heisst, $B \in O(n)$, und $A = BDR^t$. Nachdem wir allenfalls B weiter zerlegen in das Produkt von Rotationen und einer Spiegelung $S = \text{diag}(1, \dots, 1, -1)$, folgt $|\det(D)| = |\det(A)|$, und mit ii) und Satz 9.3.2.iv) erhalten wir die Behauptung.

□

9.5.2 Transformation Jordan-messbarer Mengen

Ähnliches gilt auch für geeignete Transformationen mit $\Phi \in C^1(\bar{\Omega}; \mathbb{R}^n)$. Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^1(U; \mathbb{R}^n)$ ein Diffeomorphismus von U auf $\Phi(U) = V$ gemäss Definition 8.9.2; das heisst, Φ ist injektiv mit Umkehrabbildung $\Psi = \Phi^{-1} \in C^1(V; \mathbb{R}^n)$.

Bemerkung 9.5.1. Nach dem Umkehrsatz, Satz 8.7.1, ist eine injektive Abbildung $\Phi \in C^1(U; \mathbb{R}^n)$ genau dann ein Diffeomorphismus, wenn gilt

$$\forall x_0 \in U : \det(d\Phi(x_0)) \neq 0.$$

Beispiel 9.5.1. i) Eine lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ 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 6.2.2 ein Diffeomorphismus auf $g(]a, b[)$.

iii) Die Abbildung $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ aus Beispiel 8.6.1 mit

$$f(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix} = z^2, \quad (x, y) = z = x + iy \in \mathbb{R}^2,$$

liefert einen Diffeomorphismus $\Phi = f|_{\mathbb{R}_+^2}: \mathbb{R}_+^2 \rightarrow \mathbb{R}^2 \setminus \{(x, 0); x \leq 0\}$, wobei $\mathbb{R}_+^2 = \{(x, y); x > 0\}$.

Satz 9.5.1. (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 Satz 9.3.1.iii) und Bemerkung 9.3.2 genügt es zu zeigen, dass man zu jedem $\epsilon > 0$ eine Elementarfigur $F \subset \mathbb{R}^n$ finden kann mit $\partial(\Phi(\Omega)) \subset F$ und $\mu(F) < \epsilon$.

Da Φ ein Diffeomorphismus ist, bildet Φ innere Punkte von Ω ab auf innere Punkte von $\Phi(\Omega)$, ebenso innere Punkte von $U \setminus \Omega$ auf innere Punkte des Komplements, und umgekehrt; also folgt

$$\partial(\Phi(\Omega)) = \Phi(\partial\Omega).$$

Da Ω beschränkt und Jordan messbar, gibt es wie im Beweis von Lemma 9.3.1 zu jedem $\epsilon > 0$ Elementarfiguren $E \subset \Omega^\circ \subset \Omega \subset \overline{\Omega} \subset G$ mit $\mu(G \setminus E) < \epsilon$, welche als disjunkte Vereinigungen $E = \cup_{1 \leq l \leq L_E} Q_l$, $G = \cup_{1 \leq l \leq L_G} Q_l$ von achsenparallelen, zu einem Würfel Q_0 der Kantenlänge δ kongruenten Würfeln Q_l darstellbar sind. Es folgt

$$\partial\Omega \subset G \setminus E = \cup_{L_E < l \leq L_G} Q_l, \quad \mu(G \setminus E) = L\delta^n < \epsilon,$$

wobei $L = L_G - L_E$. Weiter können wir annehmen, dass für ein $\epsilon_0 > 0$ und eine kompakte Menge $K \subset U$ gilt

$$\forall 0 < \epsilon < \epsilon_0 : G \subset K = \overline{K} \subset U.$$

Nach Korollar 5.3.1 existiert $C > 0$ mit

$$\sup_{x \in K} |d\Phi(x)| \leq C < \infty.$$

Für $1 \leq l \leq L_G$ folgt

$$\sup_{x, y \in Q_l} |\Phi(x) - \Phi(y)| \leq \int_0^1 \left| d\Phi \left(\underbrace{x + t(y-x)}_{\in Q_l \subset K} \right) \right| \underbrace{|x-y|}_{\leq \sqrt{n}\delta} dt \leq C\sqrt{n}\delta; \quad (9.5.1)$$

das heisst, $\Phi(Q_l)$ ist enthalten in einem Würfel $R_l \subset \mathbb{R}^n$ der Kantenlänge $C_1\delta$ mit einer von ϵ unabhängigen Konstanten C_1 . Somit gilt

$$\partial(\Phi(\Omega)) = \Phi(\partial\Omega) \subset \bigcup_{l=L_E+1}^{L_G} \Phi(Q_l) \subset \bigcup_{l=L_E+1}^{L_G} R_l =: F$$

mit

$$\mu(F) \leq \sum_{l=L_E+1}^{L_G} \mu(R_l) = LC_1^n \delta^n = C_1^n \mu(G \setminus E) < C_1^n \epsilon.$$

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

ii) Für einen beliebigen Würfel $Q_l \subset G \subset K$ sei x_l der Mittelpunkt von Q_l und sei $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ die Abbildung

$$Ax = \Phi(x_l) + d\Phi(x_l)(x - x_l).$$

Da $d\Phi$ auf K nach Satz 5.5.2 gleichmässig stetig, erhalten wir analog zu (9.5.1) für $x \in Q_l$ die Abschätzung

$$\frac{|\Phi(x) - A(x)|}{|x - x_l|} \leq \int_0^1 |d\Phi(x_l + t(x - x_l)) - d\Phi(x_l)| dt \rightarrow 0 \quad (\delta \rightarrow 0),$$

gleichmässig in $x \in Q_l$, $1 \leq l \leq L_G$; insbesondere folgt

$$\sup_{1 \leq l \leq L_G} \sup_{x \in Q_l} \frac{|\Phi(x) - A(x)|}{\delta} \rightarrow 0 \quad (\delta \rightarrow 0).$$

Verschieben wir den Koordinatenursprung in den Punkt x_l , so erhalten wir für jedes $\gamma > 0$ und hinreichend kleines $\delta > 0$ somit

$$A((1 - \gamma)Q_l) \subset \Phi(Q_l) \subset A((1 + \gamma)Q_l),$$

also mit $\mu(A((1 \pm \gamma)Q_l)) = (1 \pm \gamma)^n \mu(AQ_l)$ auch

$$(1 - \gamma)^n \mu(AQ_l) \leq \mu(\Phi(Q_l)) \leq (1 + \gamma)^n \mu(AQ_l).$$

Gemäss Lemma 9.5.1 gilt $\mu(AQ_l) = |\det(d\Phi(x_l))| \mu(Q_l)$. Für den Fehlerterm

$$I_l := \mu(\Phi(Q_l)) - |\det(d\Phi(x_l))| \mu(Q_l)$$

folgt so die Abschätzung

$$\frac{|I_l|}{\delta^n} \leq C \frac{|\mu(\Phi(Q_l)) - \mu(AQ_l)|}{\mu(AQ_l)} \rightarrow 0 \quad (\delta \rightarrow 0),$$

gleichmässig in l .

Zu vorgegebenem $\epsilon > 0$ seien $E \subset \Omega^\circ \subset \Omega \subset \bar{\Omega} \subset G \subset K$ Elementarfiguren wie in i) mit $\mu(G \setminus E) < \epsilon$, bestehend aus Würfeln der Kantenlänge δ . Bei festem E und G können wir δ durch $2^{-k}\delta$ mit beliebigem $k \in \mathbb{N}$ ersetzen und erhalten (mit $L_E = L_E(k) \leq C(2^{-k}\delta)^{-n}$, $I_l = I_l(k) = o((2^{-k}\delta)^n)$)

$$\mu(\Phi(E)) = \sum_{l=1}^{L_E} \mu(\Phi(Q_l)) = \sum_{l=1}^{L_E} \left(|\det(d\Phi(x_l))| \mu(Q_l) + I_l \right)$$

mit dem Gesamtfehler

$$I := \sum_{l=1}^{L_E} |I_l| \leq L_E \max_{1 \leq l \leq L_E} |I_l| \rightarrow 0 \quad (k \rightarrow \infty).$$

Im Limes $k \rightarrow \infty$ ergibt Satz 9.1.1 dann die Beziehung

$$\mu(\Phi(E)) = \int_E |\det(d\Phi(x))| d\mu(x).$$

Mit

$$\begin{aligned} \mu(\Phi(\Omega)) - \int_{\Omega} |\det(d\Phi(x))| d\mu(x) \\ = \mu(\Phi(\Omega)) - \mu(\Phi(E)) - \int_{\Omega \setminus E} |\det(d\Phi(x))| d\mu(x) \end{aligned}$$

und den Abschätzungen

$$\begin{aligned} 0 < \mu(\Phi(\Omega)) - \mu(\Phi(E)) \\ \leq \mu(\Phi(G)) - \mu(\Phi(E)) = \mu(\Phi(G \setminus E)) \leq \mu(F) < C\epsilon \end{aligned}$$

sowie

$$\int_{\Omega \setminus E} |\det(d\Phi(x))| d\mu(x) \leq \int_{G \setminus E} \underbrace{|\det(d\Phi(x))|}_{\leq C} d\mu(x) \leq C\mu(G \setminus E) \leq C\epsilon$$

folgt die Behauptung nach Übergang zum Limes $\epsilon \downarrow 0$.

□

Beispiel 9.5.2. i) Sei $g \in C^1(]a, b[)$ mit $g' > 0$ ein Diffeomorphismus auf $g(]a, b[) =]c, d[$ gemäss Satz 6.2.2, 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 $g(x) = \log(x) : \mathbb{R}_+ \rightarrow \mathbb{R}$, so erhalten wir beispielsweise " $\mu(g(]0, 1[)) = \infty$ ".

ii) Polarkoordinaten. Die Abbildung

$$\Phi(r, \theta) = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}, \quad r > 0, \quad 0 < \theta < 2\pi$$

erfüllt gemäss Beispiel 8.7.1.iii) die Gleichung

$$\det(d\Phi(r, \theta)) = r$$

und ist für festes $R > 0$ ein Diffeomorphismus von $]0, R[\times]0, 2\pi[$ auf die Menge

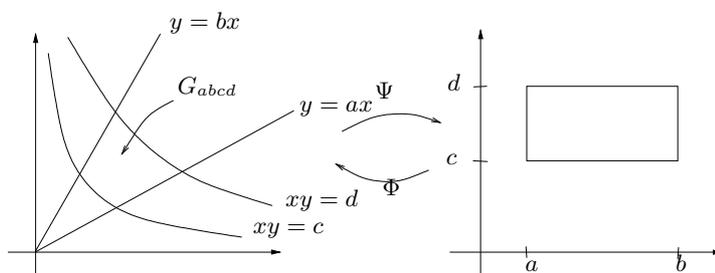
$$\Phi(]0, R[\times]0, 2\pi[) = B_R(0) \setminus \{(x, 0); 0 \leq x < R\}.$$

Da $\mu(\{(x, 0); 0 \leq x < R\}) = 0$, folgt

$$\mu(B_R(0)) = \lim_{\epsilon \downarrow 0} \int_{] \epsilon, R[\times] \epsilon, 2\pi - \epsilon[} r d\mu(r, \theta) = \int_0^{2\pi} \underbrace{\left(\int_0^R r dr \right)}_{=R^2/2} d\theta = \pi R^2.$$

iii) Für $0 < a < b$, $0 < c < d$ betrachte die Menge

$$G_{abcd} = \{(x, y) \in \mathbb{R}^2; 0 < x, a < y/x < b, c < xy < d\}.$$



Stelle G_{abcd} dar als $\Phi(]a, b[\times]c, d[)$, wobei $\Phi = \Psi^{-1}$ mit

$$\Psi : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} y/x \\ xy \end{pmatrix} =: \begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

Auflösen der Gleichungen

$$\frac{y}{x} = \xi, \quad xy = \eta$$

für $x, y, \xi, \eta > 0$ ergibt

$$x = \sqrt{\frac{\eta}{\xi}}, \quad y = \sqrt{\xi\eta}.$$

Die Abbildung

$$\Phi : \begin{pmatrix} \xi \\ \eta \end{pmatrix} \mapsto \begin{pmatrix} \sqrt{\eta/\xi} \\ \sqrt{\xi\eta} \end{pmatrix}$$

ist offenbar auf $]0, \infty[\times]0, \infty[$ bijektiv mit

$$d\Phi(\xi, \eta) = \frac{1}{2} \begin{pmatrix} -x/\xi & x/\eta \\ y/\xi & y/\eta \end{pmatrix},$$

also

$$\det(d\Phi(\xi, \eta)) = -\frac{xy}{2\xi\eta} = -\frac{1}{2\xi} < 0$$

und

$$\mu(G_{abcd}) = \int_a^b \int_c^d |\det(d\Phi)| d\eta d\xi = \frac{d-c}{2} \int_a^b \frac{d\xi}{\xi} = \frac{d-c}{2} \log\left(\frac{b}{a}\right).$$

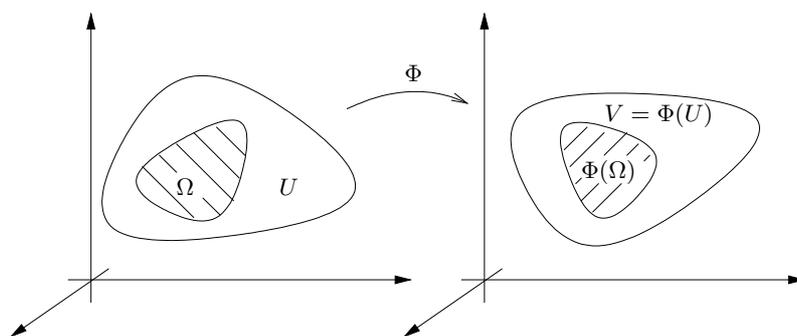
9.5.3 Substitutionsregel

Analog zu Satz 9.5.1 erhalten wir auch eine Satz 7.1.5 verallgemeinernde Regel für Integrale.

Satz 9.5.2. (Substitutionsregel) Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^1(U; \mathbb{R}^n)$ ein Diffeomorphismus von U auf $V = \Phi(U) \subset \mathbb{R}^n$, $\Omega \subset \bar{\Omega} \subset U$ beschränkt und Jordan messbar, und sei $f: \Phi(\Omega) \rightarrow \mathbb{R}$ beschränkt und R -integrabel.

Dann ist die Funktion $(f \circ \Phi) \cdot |\det(d\Phi)| : \Omega \rightarrow \mathbb{R}$ ebenfalls \mathbb{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 9.5.1, 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 $\Omega_1, \dots, \Omega_J$ disjunkt und überdecken Ω . Mit Satz 9.5.1 erhalten wir nun

$$\begin{aligned} \int_{\Phi(\Omega)} f \, d\mu &= \sum_{j=1}^J c_j \mu(\underbrace{Q_j \cap \Phi(\Omega)}_{=\Phi(\Omega_j)}) \\ &= \sum_{j=1}^J c_j \int_{\Omega_j} |\det(d\Phi)| \, d\mu = \int_{\Omega} (f \circ \Phi) |\det(d\Phi)| \, d\mu. \end{aligned}$$

ii) Für eine beliebige beschränkte, \mathbb{R} -integrable Funktion $f: \Phi(\Omega) \rightarrow \mathbb{R}$ folgt nun

$$\begin{aligned} \int_{\Phi(\Omega)} f \, d\mu &= \sup \left\{ \int_{\Phi(\Omega)} e \, d\mu; e \leq f, e \text{ TF} \right\} \\ &\stackrel{i)}{=} \sup \left\{ \int_{\Omega} (e \circ \Phi) |\det(d\Phi)| \, d\mu; e \leq f, e \text{ TF} \right\} \\ &\leq \int_{\underline{\Omega}} (f \circ \Phi) |\det(d\Phi)| \, d\mu, \end{aligned}$$

wobei wir benutzen, dass nach Definition des R-Integrals gilt

$$\begin{aligned} \int_{\Omega} (e \circ \Phi) |det(d\Phi)| d\mu &= \sup \left\{ \int_{\Omega} h d\mu; h \text{ TF}, h \leq (e \circ \Phi) |det(d\Phi)| \right\} \\ &\leq \sup \left\{ \int_{\Omega} h d\mu; h \text{ TF}, h \leq (f \circ \Phi) |det(d\Phi)| \right\} \\ &= \int_{\underline{\Omega}} (f \circ \Phi) |det(d\Phi)| d\mu. \end{aligned}$$

Analog erhalten wir

$$\begin{aligned} \overline{\int_{\Omega} (f \circ \Phi) |det(d\Phi)| d\mu} &\leq \inf \left\{ \int_{\Omega} (g \circ \Phi) |det(d\Phi)| d\mu; f \leq g, g \text{ TF} \right\} \\ &\stackrel{i)}{=} \inf \left\{ \int_{\Phi(\Omega)} g d\mu; f \leq g, g \text{ TF} \right\} \\ &= \int_{\Phi(\Omega)} f d\mu. \end{aligned}$$

Die Behauptung folgt. □

Beispiel 9.5.3. Die Funktion $f(x) = e^{-x^2} : \mathbb{R} \rightarrow \mathbb{R}$ hat keine elementar berechenbare Stammfunktion. Das Integral $\int_{-\infty}^{\infty} e^{-x^2} dx$ können wir jedoch mit Satz 9.5.2 explizit bestimmen.

Nach dem Satz 9.2.1 von Fubini gilt

$$\begin{aligned} \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy = \int_{\mathbb{R}^2} e^{-r^2} d\mu, \end{aligned}$$

wobei $r^2 = x^2 + y^2$. Gemäss Satz 9.5.2 erhalten wir nach Einführung von Polarkoordinaten entsprechend Beispiel 9.5.3.ii) für den letzten Term den Ausdruck

$$\int_{\mathbb{R}^2} e^{-r^2} d\mu = \int_0^{2\pi} \underbrace{\left(\int_0^{\infty} r e^{-r^2} dr \right)}_{=1/2} d\theta = \pi;$$

also gilt

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

9.6 Oberflächenmass und Fluss-Integral

Die Sätze 9.5.1 und 9.5.2 zeigen einen natürlichen Weg auf zur Definition des Inhalts von regulären Flächenstücken im $S \subset \mathbb{R}^3$ und von Oberflächenintegralen. Der im Abschnitt 8.10 definierte Begriff einer Immersion liefert hierfür das geeignete Hilfsmittel.

9.6.1 Der Inhalt eines parametrischen Flächenstücks

Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U; \mathbb{R}^3)$ eine injektive Immersion gemäss Definition 8.10.1. Weiter sei $\Omega \subset \bar{\Omega} \subset U$ beschränkt und Jordan messbar, und sei $S = \Phi(\Omega) \subset \mathbb{R}^3$ 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 9.6.1. Der 2-dimensionale **Flächeninhalt** von S (bzgl. Φ) ist

$$\mu_2(\Omega) := \int_S do := \int_{\Omega} |\Phi_u \times \Phi_v| \, d\mu(u, v),$$

mit $\Phi_u = \frac{\partial \Phi}{\partial u}$, etc, und mit dem **skalaren Flächenelement**

$$do := |\Phi_u \times \Phi_v| \, d\mu(u, v).$$

Bemerkung 9.6.1. 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 9.6.1 ist somit konsistent mit Satz 9.5.1.

iii) Allgemein gilt

$$|\Phi_u \times \Phi_v| = \sqrt{\det(g)},$$

wobei

$$g = d\Phi^t d\Phi = \begin{pmatrix} |\Phi_u|^2 & \Phi_u \cdot \Phi_v \\ \Phi_u \cdot \Phi_v & |\Phi_v|^2 \end{pmatrix}$$

die von Φ induzierte **Gramsche Matrix** (Metrik) bezeichnet.

Beweis. In geeigneten Koordinaten gilt $d\Phi(w): \mathbb{R}^2 \rightarrow \mathbb{R}^2 \times \{0\} \cong \mathbb{R}^2$ und

$$\begin{aligned} |\Phi_u \times \Phi_v|(w) &= |\det(d\Phi(w))| = |\det(d\Phi^t(w))| \\ &= \sqrt{\det(d\Phi^t(w)) \cdot \det(d\Phi(w))} = \sqrt{\det(g)}. \end{aligned}$$

□

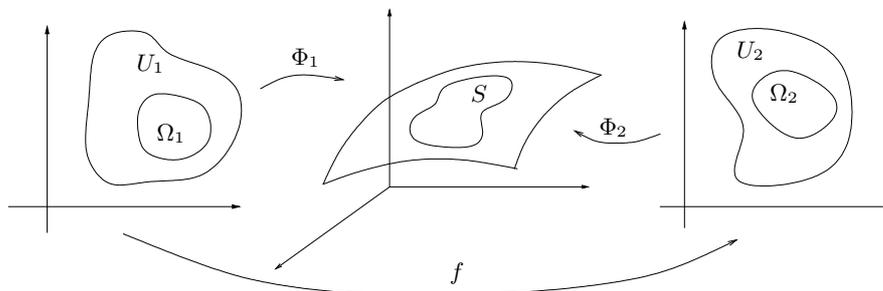
iv) Mit Satz 9.5.2 folgt aus iii), dass der Flächeninhalt $\mu_2(S)$ von der Parametrisierung unabhängig ist.

Beweis. Seien $\Phi_i \in C^1(U_i; \mathbb{R}^3)$ injektive Immersionen, Ω_i beschränkt und Jordan-messbar mit $\Omega_i \subset \bar{\Omega}_i \subset U_i$ und $\Phi_i(\Omega_i) = S$, $i = 1, 2$.

Behauptung 1. Die Abbildung

$$f = \Phi_2^{-1} \circ \Phi_1 \in C^1(\bar{\Omega}_1; \mathbb{R}^2)$$

ist ein Diffeomorphismus von Ω_1 auf Ω_2 .



Behauptung 1 liefert die gewünschte Aussage: Mit $\Phi_1 = \Phi_2 \circ f$, der Kettenregel

$$d\Phi_1 = (d\Phi_2 \circ f)df$$

und Koordinaten wie in ii) folgt

$$\begin{aligned} \det(g_1) &= \det(d\Phi_1^t d\Phi_1) \\ &= \det(df^t) \det((d\Phi_2^t d\Phi_2) \circ f) \det(df) = \det(g_2 \circ f) |\det(df)|^2. \end{aligned}$$

Mit Bemerkung iii) und Satz 9.5.2 erhalten wir so

$$\int_S do = \int_{\Omega_1} \sqrt{\det(g_1)} d\mu = \int_{\Omega_1} \sqrt{\det(g_2) \circ f} |\det(df)| d\mu = \int_{\Omega_2} \sqrt{\det(g_2)} d\mu;$$

$\mu_2(S)$ ist also unabhängig von Φ erklärt. \square

Beweis von Behauptung 1. Beachte, dass $\Phi_2|_{\overline{\Omega_2}} : \overline{\Omega_2} \rightarrow \overline{S}$ stetig und bijektiv ist. Da $\overline{\Omega_2}$ kompakt, ist $\Phi_2^{-1} : \overline{S} \rightarrow \overline{\Omega_2}$ stetig nach Satz 5.3.5, und $f \in C^0(\overline{\Omega_1}; \mathbb{R}^2)$.

Sei $w_1 \in \Omega_1$, und sei $w_2 = f(w_1) \in \Omega_2$. Sei weiter $\Phi_2(w_2) = \Phi_1(w_1) = p \in S$ mit

$$d\Phi_2(w_2)(\mathbb{R}^2) = d\Phi_1(w_1)(\mathbb{R}^2) = T_p S \cong \mathbb{R}^2$$

gemäss Bemerkung 8.10.1.i). In geeigneten Koordinaten gilt $T_p S = \mathbb{R}^2 \times \{0\}$, und wir können $d\Phi_2(w_2)$ als linearen Isomorphismus $d\Phi_2(w_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ auffassen. Wir zeigen, f ist an der Stelle w_1 differenzierbar mit

$$df(w_1) = d\Phi_2^{-1}(w_2) d\Phi_1(w_1). \quad (9.6.1)$$

Für $\tilde{w}_1 \in \Omega_1$ nahe w_1 gilt

$$\begin{aligned} \tilde{p} &= \Phi_1(\tilde{w}_1) = \Phi_1(w_1) + d\Phi_1(w_1)(\tilde{w}_1 - w_1) + o(|\tilde{w}_1 - w_1|) \\ &= \Phi_2(\tilde{w}_2) = \Phi_2(w_2) + d\Phi_2(w_2)(\tilde{w}_2 - w_2) + o(|\tilde{w}_2 - w_2|), \end{aligned}$$

wobei $\tilde{w}_2 = f(\tilde{w}_1) \rightarrow f(w_1) = w_2$ für $\tilde{w}_1 \rightarrow w_1$ wegen der Stetigkeit von f . Da $\Phi_2(w_2) = p = \Phi_1(w_1)$, folgt

$$\begin{aligned} \tilde{w}_2 - w_2 &= f(\tilde{w}_1) - f(w_1) = d\Phi_2(w_2)^{-1} d\Phi_1(w_1)(\tilde{w}_1 - w_1) \\ &\quad + o(|\tilde{w}_1 - w_1|) + o(|\tilde{w}_2 - w_2|) \quad (\tilde{w}_1 \rightarrow w_1) \end{aligned}$$

und damit zunächst $\tilde{w}_2 - w_2 = O(|\tilde{w}_1 - w_1|)$ für $\tilde{w}_1 \rightarrow w_1$ und dann auch (9.6.1). Nach Vertauschen von Φ_1 und Φ_2 folgt dasselbe für f^{-1} ; das heisst, $f \in C^1$ ist ein Diffeomorphismus, wie gewünscht. \square

Beispiel 9.6.1. 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 8.10.2 mit

$$|\Phi_u \times \Phi_v| = |\det(\Phi, \Phi_u, \Phi_v)| \equiv \frac{4}{(1 + u^2 + v^2)^2}.$$

Nach Übergang zu Polarkoordinaten folgt

$$\begin{aligned} \mu_2(S^2) &= \int_{\mathbb{R}^2} |\Phi_u \times \Phi_v| \, d\mu(u, v) = \int_{\mathbb{R}^2} \frac{4}{(1 + u^2 + v^2)^2} \, d\mu(u, v) \\ &= \int_0^{2\pi} \left(\int_0^\infty \frac{4r \, dr}{(1 + r^2)^2} \right) d\theta = 4\pi, \end{aligned}$$

wobei wir benutzen, dass Substitution $s = r^2$ ergibt

$$\int_0^\infty \frac{4r \, dr}{(1 + r^2)^2} = 2 \int_0^\infty \frac{ds}{(1 + s)^2} = -\frac{2}{1 + s} \Big|_{s=0}^\infty = 2.$$

Wir können nun auch stetige Funktionen über reguläre Flächenstücke integrieren.

Definition 9.6.2. Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U; \mathbb{R}^3)$ injektive Immersion, $\Omega \subset \bar{\Omega} \subset U$ beschränkt und Jordan-messbar, und sei $S = \Phi(\Omega)$ das zugehörige Flächenstück im \mathbb{R}^3 , $f: \bar{S} \rightarrow \mathbb{R}$ 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 9.6.2. Für Φ , $S = S^2$ wie in Beispiel 8.10.2 und $f(x, y, z) = z^2$ erhalten wir

$$\begin{aligned} \int_{S^2} f \, d\mu &= \int_{\mathbb{R}^2} (f \circ \Phi) \cdot |\Phi_u \times \Phi_v| \, d\mu \\ &= \int_{\mathbb{R}^2} \left(\frac{1 - u^2 - v^2}{1 + u^2 + v^2} \right)^2 \frac{4}{(1 + u^2 + v^2)^2} \, d\mu \\ &= 2\pi \int_0^\infty \left(\frac{1 - r^2}{1 + r^2} \right)^2 \frac{4r}{(1 + r^2)^2} \, dr. \end{aligned}$$

Mit

$$\left(\frac{1 - r^2}{1 + r^2} \right)^2 = \left(\frac{2}{1 + r^2} - 1 \right)^2 = \frac{4}{(1 + r^2)^2} - \frac{4}{1 + r^2} + 1$$

folgt nach der Substitution $s = r^2$ die Gleichung

$$\int_{S^2} f \, d\mu = 2\pi \int_0^\infty \left(\frac{8 \, ds}{(1 + s)^4} - \frac{8 \, ds}{(1 + s)^3} + \frac{2 \, ds}{(1 + s)^2} \right) = \frac{4\pi}{3}.$$

9.6.2 Der Fluss eines Vektorfeldes

Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U; \mathbb{R}^3)$ injektive Immersion, $\Omega \subset \bar{\Omega} \subset U$ beschränkt und Jordan-messbar, und sei $S = \Phi(\Omega)$ das zugehörige Flächenstück im \mathbb{R}^3 . In jedem Punkt $p = \Phi(w)$ mit $w \in \Omega$ spannen die Vektoren $\Phi_u(w)$ und $\Phi_v(w)$ den Tangentialraum auf an S ; das heisst,

$$T_p S = d\Phi(w)(\mathbb{R}^2) = \text{span}\{\Phi_u(w), \Phi_v(w)\}.$$

Durch Auswahl eines **Einheitsnormalenvektorfeldes** $n: S \rightarrow \mathbb{R}^3$ mit $n(p) \perp T_p S$, $|n(p)| = 1$ an jedem Punkt $p \in S$ können wir S **orientieren**, also “oben” und “unten” definieren. Die Parametrisierung Φ liefert

$$n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$$

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 9.6.3. *Das Integral*

$$\int_S K \cdot n \, do = \int_{\Omega} (K \circ \Phi) \cdot \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|} |\Phi_u \times \Phi_v| \, d\mu = \int_{\Omega} (K \circ \Phi) \cdot \Phi_u \times \Phi_v \, d\mu$$

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 9.6.3. i) Sei $\Phi \in C^1(\mathbb{R}^2; S^2)$ wie in Beispiel 8.10.2, $S = \Phi(\mathbb{R}^2) = S^2$, und sei K das Ortsvektorfeld mit $K(p) = p$, $p \in S^2$. Für die durch Φ definierte Orientierung $n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$ gilt offenbar $n(p) = p$; also

$$\int_{S^2} K \cdot n \, do = \int_{S^2} do = \mu(S^2) = 4\pi.$$

ii) Mit Φ , $S = S^2$ mit $n(p) = p$ wie in i) und $K(x, y, z) = (0, 0, z)^t$ erhalten wir

$$\int_{S^2} K \cdot n \, do = \int_{S^2} z^2 \, do = \frac{4\pi}{3};$$

vergleiche Beispiel 9.6.2.

9.7 Der Satz von Stokes im \mathbb{R}^3

Analog zu Definition 8.9.3 und unter Berücksichtigung von Bemerkung 8.10.1 können wir allgemein den Begriff einer Untermannigfaltigkeit mit Rand im \mathbb{R}^n erklären. Die folgende Definition gilt für allgemeine Dimensionen $1 \leq k < n$.

Definition 9.7.1. i) Eine Menge $S \subset \mathbb{R}^n$ ist eine eingebettete k -dimensionale C^m -Untermannigfaltigkeit mit Rand der Klasse C^1_{pw} , $m \geq 1$, falls zu jedem $p_0 \in S$ eine Umgebung W von p_0 und eine injektive Immersion $\Phi \in C^m(U; \mathbb{R}^n)$ einer offenen Menge $U \subset \mathbb{R}^k$ sowie eine beschränkte Menge $\Omega \subset \bar{\Omega} \subset U$ von der Klasse C^1_{pw} existieren, so dass $S \cap W = \Phi(\Omega)$.

ii) Ein Punkt $p_0 \in S$ gehört zum **relativen (intrinsischen) Rand** ∂S von S , falls für jedes Φ wie in i) gilt $p_0 \in \Phi(\partial\Omega)$.

Im folgenden Abschnitt betrachten wir speziell den Fall $k = 2$, $n = 3$.

Sei $S \subset \mathbb{R}^3$ ein Flächenstück mit Rand der Klasse C^1_{pw} , und sei $p = \Phi(w) \in S$ mit einer injektiven Immersion $\Phi \in C^1(U; \mathbb{R}^3)$, wobei $U \subset \mathbb{R}^2$ offen, $w \in \Omega \subset \bar{\Omega} \subset U$ beschränkt und von der Klasse C^1_{pw} .

Wie in Abschnitt 9.6 können wir S durch Auswahl eines Einheitsnormalenvektorfeldes $n: S \rightarrow \mathbb{R}^3$ lokal um p herum orientieren.

Definition 9.7.2. Ein stetiges Einheitsnormalenvektorfeld $n \in C^0(S; \mathbb{R}^3)$ definiert eine **Orientierung** von S .

Bemerkung 9.7.1. i) Falls $S = \Phi(\Omega)$ mit einer injektiven Immersion $\Phi \in C^m(U; \mathbb{R}^3)$ einer offenen Menge $U \subset \mathbb{R}^2$ und einer beschränkten Menge $\Omega \subset \bar{\Omega} \subset U$, so ergibt

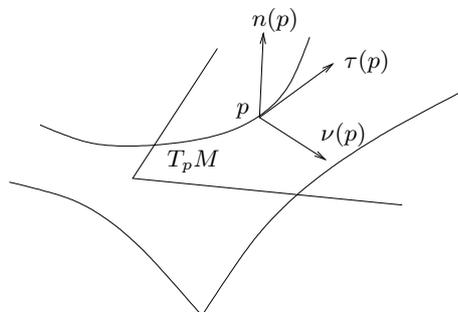
$$n = \Phi_u \times \Phi_v / |\Phi_u \times \Phi_v|$$

eine Orientierung von S .

ii) Es gibt nicht orientierbare Flächen, zum Beispiel das Möbiusband. Nach Bemerkung i) lassen sich diese Flächen nicht als $S = \Phi(\Omega)$ mit einer injektiven Immersion $\Phi \in C^1(U; \mathbb{R}^3)$ und $\Omega \subset \bar{\Omega} \subset U \subset \mathbb{R}^2$ darstellen.

iii) Falls S zusammenhängend ist, so gibt es höchstens zwei Möglichkeiten, ein stetiges Einheitsnormalenvektorfeld n auf S zu definieren.

Sei S ein orientierbares Flächenstück mit Rand der Klasse C^1_{pw} , und sei $n \in C^0(S; \mathbb{R}^3)$ eine Orientierung für S . Wir wollen auch den Rand ∂S von S orientieren. Sei dazu Φ eine lokale Parametrisierung von S wie in Definition 9.7.1,i) um $p = \Phi(w) \in \partial S \subset M := \Phi(U)$ mit $w \in \partial\Omega$.



Wähle eine Orthonormalbasis $\nu(p), \tau(p)$ für $T_p M$ so, dass

$$\tau(p) \in T_p \partial S = d\Phi(w)(T_w \partial\Omega),$$

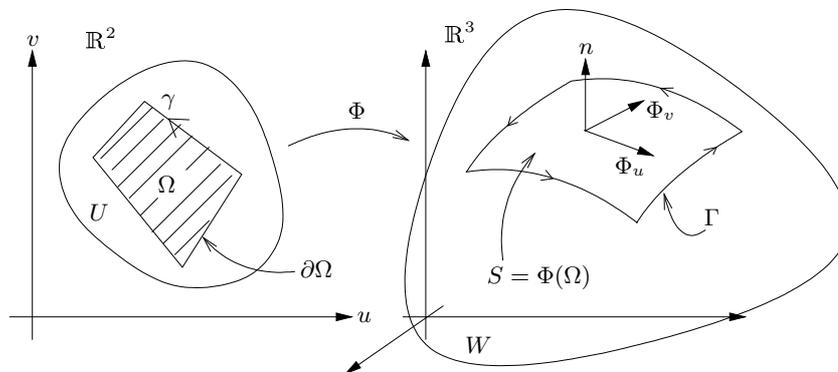
und so, dass $\nu(p)$ aus S heraus zeigt; das heisst, $\nu(p)$ ist die äussere Normale von S in M . Dann ist das Tripel $(\nu(p), \tau(p), n(p))$ eine Orthonormalbasis für \mathbb{R}^3 . Dabei ist $\nu(p)$ durch S eindeutig bestimmt, $\tau(p)$ bis auf das Vorzeichen.

Analog zu Definition 9.4.3 erklären wir:

Definition 9.7.3. i) Die Orthonormalbasis $(\nu(p), \tau(p), n(p))$ von \mathbb{R}^3 ist **positiv orientiert** oder ein **Rechtssystem**, falls $(\nu(p), \tau(p), n(p))$ durch eine Drehung aus der Standardbasis (e_1, e_2, e_3) hervorgeht; sonst heisst $(\nu(p), \tau(p), n(p))$ **negativ orientiert**.

ii) Eine Parametrisierung Γ von ∂S mit $\dot{\Gamma} \neq 0$ heisst **positiv orientiert**, falls in jedem Punkt $p = \Gamma(t) \in \partial S$ das Tripel $(\nu(p), \tau(p), n(p))$ mit $\tau(p) = \dot{\Gamma}(t)/|\dot{\Gamma}(t)|$ eine positiv orientierte Orthonormalbasis von \mathbb{R}^3 bildet.

Beispiel 9.7.1. i) Sei $U \subset \mathbb{R}^2$ offen, $\Phi \in C^1(U; \mathbb{R}^3)$ injektive Immersion, $\Omega \subset \bar{\Omega} \subset U$ beschränkt mit Rand der Klasse C^1_{pw} , und sei $S = \Phi(\Omega)$ das zugehörige Flächenstück im \mathbb{R}^3 . Sei $\gamma \in C^1_{pw}([0, 1]; \mathbb{R}^2)$ eine positiv orientierte Parametrisierung von $\partial\Omega$ in dem Sinne, dass Ω beim Durchlaufen von γ zur Linken liegt, und sei $\Gamma := \Phi \circ \gamma \in C^1_{pw}([0, 1]; \mathbb{R}^3)$ die davon induzierte Parametrisierung des intrinsischen Randes $\partial S = \Phi(\partial\Omega)$ von S . Versehen wir S mit der durch Φ induzierten Orientierung $n = \frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|}$, so ist Γ ebenfalls positiv orientiert: Wenn wir “auf” Γ die Fläche S umfahren, liegt S stets zur Linken.



Genauer sei $p = \Phi(w) \in \partial S$ mit $w = \gamma(t) \in \partial\Omega$, und sei $\nu_\Omega \in T_w\mathbb{R}^2$ die äussere Normale an $\partial\Omega$ im Punkt w . OBdA dürfen wir annehmen, dass $\nu_\Omega = e_u$, $\dot{\gamma}(t) = e_v$. Setze

$$\tilde{\nu} = d\Phi(w)\nu_\Omega = \Phi_u(w), \quad \tilde{\tau} = d\Phi(w)\dot{\gamma}(t) = \Phi_v(w)$$

und normiere

$$\tau = \tilde{\tau}/|\tilde{\tau}|, \quad \nu = (\tilde{\nu} - \tau(\tau \cdot \tilde{\nu}))/|\tilde{\nu} - \tau(\tau \cdot \tilde{\nu})|.$$

Dann ist das Tripel (ν, τ, n) positiv orientiert, da mit einem Faktor $c_p > 0$ gilt

$$c_p \det(\nu, \tau, n) = \det(\tilde{\nu}, \tilde{\tau}, n) = (\tilde{\nu} \times \tilde{\tau}) \cdot n = |\tilde{\nu} \times \tilde{\tau}| > 0.$$

$\Gamma = \Phi \circ \gamma$ ist somit positiv orientierte Parametrisierung von ∂S um p .

ii) Sei insbesondere $\Phi \in C^1(\mathbb{R}^2; \mathbb{R}^3)$ wie in Beispiel 8.10.2, und sei $\Omega = B_1(0) \subset \overline{\Omega} \subset U = \mathbb{R}^2$ mit $\partial\Omega = \partial B_1(0) = S^1$ parametrisiert durch $\gamma(t) = (\cos(t), \sin(t))^t$, $0 \leq t \leq 2\pi$. Dann ist $S = \Phi(\Omega)$ die obere Halbsphäre, orientiert durch $n(p) = p$, mit Rand $\partial S = S^1$, parametrisiert durch $\Gamma = \gamma$ und mit der äusseren Normalen $\nu = -e_3$.

Sei $K = (P, Q, R)^t \in C^1(W; \mathbb{R}^3)$ in einer Umgebung W von S in \mathbb{R}^3 .

Definition 9.7.4. Das Vektorfeld

$$\vec{\text{rot}} K = \begin{pmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{pmatrix} = \nabla \times K = \det \begin{pmatrix} \partial_x & P & e_1 \\ \partial_y & Q & e_2 \\ \partial_z & R & e_3 \end{pmatrix}$$

heisst **Rotation** ("Wirbelstärke", engl. curl) von K .

Es gilt das folgende Analogon zu Satz 9.4.3.

Satz 9.7.1. (Stokes) Sei $S \subset \mathbb{R}^3$ ein kompaktes, orientiertes Flächenstück der Klasse C^2 mit Rand der Klasse C_{pw}^1 , und sei ∂S positiv orientiert. Dann gilt

$$\int_S \vec{\text{rot}} K \cdot n \, do = \int_{\partial S} K \cdot \vec{ds}.$$

Bemerkung 9.7.2. 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

$$\vec{\text{rot}} K \cdot n = Q_x - P_y$$

sowie

$$\int_{\partial S} K \cdot \vec{ds} = \int_{\partial\Omega} P \, dx + Q \, dy,$$

und wir erhalten Satz 9.4.1.

Beispiel 9.7.2. 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

$$\vec{\text{rot}} K = \begin{pmatrix} \cdots \\ \cdots \\ \frac{\partial(y(x^2+z^2))}{\partial x} - \frac{\partial(x(y^2+z^2))}{\partial y} \end{pmatrix} = 0;$$

also

$$\int_{\Gamma} K \cdot d\vec{s} = \int_S \vec{\text{rot}} K \cdot n \, do = 0.$$

Auf direktem Wege ist die Berechnung der Zirkulation offenbar viel aufwendiger.

Beweis von Satz 9.7.1. i) Sei zunächst $S = \Phi(\Omega)$ für eine injektive Immersion $\Phi \in C^2(U; \mathbb{R}^3)$, wobei $U \subset \mathbb{R}^2$ offen, $\Omega \subset \bar{\Omega} \subset U$ beschränkt mit Rand der Klasse C_{pw}^1 , und sei $\gamma \in C_{pw}^1([0, 1]; \mathbb{R}^2)$ eine positiv orientierte Parametrisierung von $\partial\Omega$, $\Gamma = \Phi \circ \gamma \in C_{pw}^1([0, 1]; \mathbb{R}^2)$ die davon induzierte, positiv orientierte Parametrisierung von ∂S .

Wir führen Satz 9.7.1 zurück auf Satz 9.4.1. 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 (mit $\dot{\gamma} = d\gamma/dt$)

$$\begin{aligned} \int_{\partial S} K \cdot d\vec{s} &= \int_0^1 \underbrace{(K \circ \Gamma)(t)}_{=(K \circ \Phi)(\gamma(t))} \cdot \underbrace{\frac{d\Gamma}{dt}(t)}_{=d\Phi(\gamma(t))\dot{\gamma}(t)} \, dt \\ &= \int_0^1 \lambda(\gamma(t))\dot{\gamma}(t) \, dt = \int_{\partial\Omega} g \, du + h \, dv. \end{aligned}$$

Behauptung. Es gilt

$$(\vec{\text{rot}} K \circ \Phi) \cdot \Phi_u \times \Phi_v = h_u - g_v.$$

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

$$\begin{aligned} (h_u - g_v)(w) &= aK_x(p) \cdot (b \, e_1 + c \, e_2) - (bK_x(p) + cK_y(p)) \cdot (a \, e_1) \\ &= (abP_x + acQ_x - abP_x - acP_y)(p) = ac(Q_x - P_y)(p). \end{aligned}$$

Die linke Seite ergibt andererseits, mit $(\Phi_u \times \Phi_v)(w) = ac \, e_3$,

$$\left((\vec{\text{rot}} K \circ \Phi) \cdot \Phi_u \times \Phi_v \right)(w) = \vec{\text{rot}} K(p) \cdot ac \, e_3 = ac (Q_x - P_y)(p),$$

und die Behauptung folgt.

ii) Falls S nicht wie oben darstellbar ist als $S = \Phi(\Omega)$, so zerlege $S = \bigcup_{1 \leq j \leq J} S_j$ in endlich viele disjunkte, kompatibel orientierte derartige Teile S_j , $1 \leq j \leq J$. Gemeinsame Randkurven von aneinander angrenzenden Flächenstücken S_j und S_k sind dann entgegengesetzt orientiert. Mit i) folgt

$$\int_{\partial S} K \cdot \vec{ds} = \sum_{j=1}^J \int_{\partial S_j} K \cdot \vec{ds} = \sum_{j=1}^J \int_{S_j} \operatorname{rot} K \cdot n \, do = \int_S \operatorname{rot} K \cdot n \, do,$$

wie gewünscht. □

Definition 9.7.5. Sei $W \subset \mathbb{R}^3$ offen und zusammenhängend. Dann heisst W **einfach zusammenhängend**, falls jede einfach geschlossene Kurve $\Gamma \in C_{pw}^1([0, 1]; W)$ in W ein orientiertes Flächenstück $S \subset W$ berandet.

Bemerkung 9.7.3. Die obige Definition ist für unsere Zwecke leicht vereinfacht. Genauer heisst W einfach zusammenhängend, falls jede einfach geschlossene Kurve $\Gamma \in C_{pw}^1([0, 1]; W)$ in W **null-homotop** ist, das heisst, sich in W "zusammenziehen" lässt.

Beispiel 9.7.3. i) Der Ball $B_1(0; \mathbb{R}^3)$ ist einfach zusammenhängend;

ii) Die Kugelschale $B_2(0) \setminus B_1(0) \subset \mathbb{R}^3$ ist einfach zusammenhängend;

iii) $B_2(0; \mathbb{R}^3) \setminus (B_1(0; \mathbb{R}^2) \times \mathbb{R})$ ist nicht einfach zusammenhängend.

iv) Das Möbiusband hat nur eine einzige Randkomponente, ist aber kein orientiertes Flächenstück.

Analog zu Satz 9.4.3 erhalten wir:

Satz 9.7.2. (Poincaré) Sei $W \subset \mathbb{R}^3$ einfach zusammenhängend, $K \in C^1(W; \mathbb{R}^3)$. Es sind äquivalent:

i) K ist konservativ;

ii) $\operatorname{rot} K = 0$.

Beweis. *i) \Rightarrow ii)* Sei K konservativ. Nach Satz 8.4.3 gibt es $f \in C^2(W)$ mit $K = \nabla f$; also

$$\operatorname{rot} K = \begin{pmatrix} \cdots & \cdots \\ \frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \end{pmatrix} = 0.$$

ii) \Rightarrow i) Sei $\Gamma \subset W$ einfach geschlossen, $S \subset W$ orientiertes Flächenstück mit $\partial S = \Gamma$ gemäss Definition 9.7.5. Dann gilt nach Satz 9.7.1

$$\int_{\Gamma} K \cdot \vec{ds} = \pm \int_S \operatorname{rot} K \cdot n \, do = 0;$$

das heisst, K ist konservativ. □

Beispiel 9.7.4. Sei K wie in Beispiel 9.7.2. Nach Satz 9.7.2 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 9.7.4. Ein Möbiusband S lässt sich so in $W = \mathbb{R}^3 \setminus (\{0, 0\} \times \mathbb{R})$ 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

$$\operatorname{rot} \vec{K} = \begin{pmatrix} R_y - Q_z \\ P_z - R_x \\ Q_x - P_y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

jedoch erhalten wir

$$\int_{\partial S} K \cdot \vec{ds} = 4\pi;$$

vergleiche Beispiel 9.4.7.ii).

Beachte, dass ∂S nur eine Zusammenhangskomponente hat; jedoch ist S nicht orientierbar. In der Tat ist W nicht einfach zusammenhängend.

9.8 Der Satz von Gauss im \mathbb{R}^3

Sei $\Omega \subset \mathbb{R}^3$ offen, $K = (K^i)_{1 \leq i \leq 3} \in C^1(\bar{\Omega}; \mathbb{R}^3)$, $x = (x^i)_{1 \leq i \leq 3} \in \Omega$.

Definition 9.8.1. Der Ausdruck

$$\operatorname{div}(K) = \sum_{i=1}^3 \frac{\partial K^i}{\partial x^i}$$

heißt **Divergenz** (Quellstärke) von K .

Allgemein bezeichnen wir für Vektorfelder $\varphi = (\varphi^i)_{1 \leq i \leq n} \in C^1(\Omega; \mathbb{R}^n)$ auf einer offenen Menge $\Omega \subset \mathbb{R}^n$ mit $\operatorname{div}(\varphi) = \sum_{i=1}^n \partial \varphi^i / \partial x^i$ die **Divergenz** von φ .

Satz 9.8.1. (Gauss) Sei $\Omega \subset \bar{\Omega} \subset W \subset \mathbb{R}^3$ beschränkt und von der Klasse C_{pw}^1 . Dann gilt

$$\int_{\Omega} \operatorname{div}(K) \, d\mu = \int_{\partial\Omega} K \cdot n \, d\sigma,$$

wobei n die äussere Normale ist.

Bemerkung 9.8.1. Nach Satz 9.8.1 ist also der Fluss des Vektorfeldes K durch $\partial\Omega$ gleich dem Integral über die Quellstärke von K .

Beweis von Satz 9.8.1. i) Sei zunächst Ω ein Normalbereich bzgl. aller Achsen x^1, \dots, x^3 . Vereinfachend gelte mit einem Quader $Q \subset \mathbb{R}^2$, $0 \leq \psi \in C^1(\overline{Q})$:

$$\Omega = \{x = (x^i); (x^1, x^2) \in Q, 0 \leq x^3 \leq \psi(x^1, x^2)\}.$$

Es folgt

$$\begin{aligned} \int_{\Omega} \frac{\partial K^3}{\partial x^3} d\mu_3 &= \int_Q \left(\int_0^{\psi(x^1, x^2)} \frac{\partial K^3}{\partial x^3} dx^3 \right) d\mu_2(x^1, x^2) \\ &= \int_Q (K^3(x^1, x^2, \psi(x^1, x^2)) - K^3(x^1, x^2, 0)) d\mu_2(x^1, x^2). \end{aligned} \quad (9.8.1)$$

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

$$\Phi_{x^1} \times \Phi_{x^2} = \begin{pmatrix} -\partial\psi/\partial x^1 \\ -\partial\psi/\partial x^2 \\ 1 \end{pmatrix}, \quad (9.8.2)$$

und wir erhalten

$$\begin{aligned} \int_S \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \cdot n \, do &= \int_Q \begin{pmatrix} 0 \\ 0 \\ K^3 \circ \Phi \end{pmatrix} \cdot \Phi_{x^1} \times \Phi_{x^2} \, d\mu_2(x^1, x^2) \\ &= \int_Q K^3(x^1, x^2, \psi(x^1, x^2)) \, d\mu_2(x^1, x^2). \end{aligned}$$

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

$$\int_{\Omega} \operatorname{div} \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} d\mu = \int_{\partial\Omega} \begin{pmatrix} 0 \\ 0 \\ K^3 \end{pmatrix} \cdot n \, do;$$

analog für die übrigen Komponenten.

ii) Zerlege ein beliebiges beschränktes Gebiet $\Omega \subset \overline{\Omega} \subset W$ der Klasse C_{pw}^1 gemäss Bemerkung 9.4.1.ii) in endlich viele disjunkte Normalbereiche $\Omega_1, \dots, \Omega_L$. Entlang gemeinsamer Begrenzungsflächen von Ω_k und Ω_l heben sich die Flüsse

auf. Es folgt

$$\begin{aligned} \int_{\Omega} \operatorname{div}(K) \, d\mu &= \sum_{l=1}^L \int_{\Omega_l} \operatorname{div}(K) \, d\mu \\ &= \sum_{l=1}^L \int_{\partial\Omega_l} K \cdot n_l \, do = \int_{\partial\Omega} K \cdot n \, do, \end{aligned}$$

wie gewünscht. □

Bemerkung 9.8.2. Mit einer zu (9.8.1) analogen Rechnung erhalten wir für alle $n \in \mathbb{N}$ auf jeder beschränkten Menge $\Omega \subset \mathbb{R}^n$ von der Klasse C_{pw}^1 für jedes $K = (K^i)_{1 \leq i \leq n} \in C^1(\overline{\Omega}; \mathbb{R}^n)$ mit $K = 0$ auf $\partial\Omega$ die Identität

$$\int_{\Omega} \operatorname{div}(K) \, d\mu = 0.$$

Man kann Satz 9.8.1 insbesondere zur Volumenberechnung verwenden.

Beispiel 9.8.1. i) Sei $\Omega = B_1(0; \mathbb{R}^3)$. Für $K(x) = x$ mit $\operatorname{div}(K) = 3$ folgt mit Satz 9.8.1

$$3\mu_3(B_1(0; \mathbb{R}^3)) = \int_{B_1(0; \mathbb{R}^3)} \operatorname{div} x \, dx = \int_{S^2} \underbrace{x \cdot n(x)}_{\equiv 1} \, do = \mu_2(S^2) = 4\pi;$$

also

$$\mu_3(B_1(0; \mathbb{R}^3)) = \frac{4\pi}{3}.$$

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 $\operatorname{div} K \equiv 1$ ergibt Satz 9.8.1 und die in (9.8.1) durchgeführte Rechnung

$$\begin{aligned} \mu_3(\Omega) &= \int_{\Omega} \operatorname{div} K \, d\mu_3 = \int_{\partial\Omega} K \cdot n \, do \\ &= 2c \int_{\{(x, y); \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1\}} \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \, d\mu_2(x, y) = \frac{4\pi}{3} abc; \end{aligned}$$

vergleiche Beispiel 9.3.2. Beachte, dass $z = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$ auf $\partial\Omega$.

Andererseits kann man auch gewisse Oberflächeninhalte bestimmen.

Beispiel 9.8.2. i) Sei Ω der Kegelstumpf

$$\Omega = \{(x, y, z) \in \mathbb{R}^3; \sqrt{x^2 + y^2} < 1 - z, z > 0\}$$

mit Mantel

$$M = \{(x, y, z); \sqrt{x^2 + y^2} = 1 - z \leq 1\} = \mathcal{G}(\psi),$$

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 $\operatorname{div}(K) = 0$. Mit Satz 9.8.1 erhalten wir

$$0 = \int_{\Omega} \operatorname{div} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} d\mu = \int_{\partial\Omega} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \cdot n \, do = \int_M n^3 \, do - \mu_2(B_1(0; \mathbb{R}^2)).$$

Gemäss (9.8.2) 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

$$\int_M n^3 \, do = \frac{1}{\sqrt{2}} \mu_2(M) = \mu_2(B_1(0; \mathbb{R}^2)) = \pi;$$

das heisst,

$$\mu_2(M) = \sqrt{2} \cdot \pi.$$

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

$$K(p) = \frac{p}{|p|^3}, \quad p \in \mathbb{R}^3 \setminus \{0\}.$$

Beachte die Gleichung

$$\operatorname{div}(K) = \frac{\partial}{\partial x} \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}^3} \right) + \dots = 0 \quad \text{in } \mathbb{R}^3 \setminus \{0\}.$$

Mit Satz 9.8.1 folgt

$$\begin{aligned} \mu_2(S_{z_0}) &= \int_{S_{z_0}} K \cdot n \, do = \int_{D_{z_0}} K \cdot e_3 \, d\mu(x, y) \\ &= z_0 \int_{D_{z_0}} \frac{d\mu(x, y)}{\sqrt{x^2 + y^2 + z_0^2}^3} = 2\pi z_0 \int_0^{\sqrt{1-z_0^2}} \frac{r \, dr}{\sqrt{r^2 + z_0^2}^3} \\ &= \pi z_0 \int_0^{1-z_0^2} \frac{ds}{\sqrt{s + z_0^2}^3} = \pi z_0 \cdot \left(-\frac{2}{\sqrt{s + z_0^2}} \right) \Big|_{s=0}^{1-z_0^2} = 2\pi (1 - z_0). \end{aligned}$$

iii) Sei $0 \in \Omega \subset \mathbb{R}^3$ offen und beschränkt mit Rand $\partial\Omega$ von der Klasse C^1 , und sei $q \in \mathbb{R}$. Betrachte das Vektorfeld

$$v(x) = q \frac{x}{|x|^3} \in C^\infty(\mathbb{R}^3 \setminus \{0\}; \mathbb{R}^3).$$

Wir wollen den Fluss von v durch $\partial\Omega$ bestimmen.

Mit Satz 9.8.1 folgt für genügend kleine $r > 0$ mit $B_r(0) \subset \Omega$ die Identität

$$\int_{\Omega \setminus B_r(0)} \operatorname{div}(v) d\mu = \int_{\partial\Omega} v \cdot n \, do - \int_{\partial B_r(0)} v \cdot \frac{x}{|x|} ds,$$

wobei n die äussere Normale bezeichnet. Beachte, dass gilt

$$\operatorname{div}(v) = \sum_{i=1}^3 q \frac{\partial(x^i/|x|^3)}{\partial x^i} = 3q \left(\frac{1}{|x|^3} - \frac{|x|^2}{|x|^5} \right) = 0$$

für $x \neq 0$; das heisst,

$$\int_{\partial\Omega} v \cdot n \, do = \int_{\partial B_r(0)} v \cdot \frac{x}{|x|} do = \frac{q}{r^2} \underbrace{\int_{\partial B_r(0)} do}_{=4\pi r^2} = 4\pi q.$$

9.8.1 Partielle Integration

Sei $\Omega \subset \mathbb{R}^3$ beschränkt mit $\partial\Omega \in C_{pw}^1$, $K = (K^i)_{1 \leq i \leq 3} \in C^1(\overline{\Omega}; \mathbb{R}^3)$, $f \in C^1(\overline{\Omega})$. Die Produktregel ergibt

$$\operatorname{div}(fK) = \sum_{i=1}^3 \left(\frac{\partial f}{\partial x^i} K^i + f \frac{\partial K^i}{\partial x^i} \right) = \nabla f \cdot K + f \operatorname{div}(K).$$

Mit Satz 9.8.1 folgt:

Satz 9.8.2. (Partielle Integration) Für Ω, K und f wie oben gilt:

$$\int_{\Omega} \nabla f \cdot K \, d\mu = \int_{\partial\Omega} (fK) \cdot n \, do - \int_{\Omega} f \operatorname{div}(K) \, d\mu.$$

Bemerkung 9.8.3. Nach Bemerkung 9.8.2 gilt Satz 9.8.2 in allen Dimensionen $n \in \mathbb{N}$ auf jeder beschränkten Menge $\Omega \subset \mathbb{R}^n$ von der Klasse C_{pw}^1 , falls $K = (K^i)_{1 \leq i \leq n} \in C^1(\overline{\Omega}; \mathbb{R}^n)$, $f \in C^1(\overline{\Omega})$ mit $fK = 0$ auf $\partial\Omega$.

Besonders interessante Anwendungen ermöglicht Satz 9.8.2, wenn $K = \nabla u$ das Gradientenfeld einer Funktion $u \in C^2(\overline{\Omega})$ ist. In diesem Fall erhalten wir

$$\operatorname{div}(K) = \operatorname{div}(\nabla u) = \sum_{i=1}^n \frac{\partial^2 u}{(\partial x^i)^2} =: \Delta u$$

mit dem **Laplace-Operator** $\Delta = \sum_{i=1}^n \frac{\partial^2}{(\partial x^i)^2}$.

9.8.2 Harmonische Funktionen

Sei $\Omega \subset \mathbb{R}^n$ beschränkt mit $\partial\Omega \in C^1_{pw}$, $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$.

Definition 9.8.2. Die Funktion u heisst **harmonisch**, falls $\Delta u = 0$.

Beispiel 9.8.3. i) Sei

$$u(x) = \frac{1}{|x|} \in C^\infty(\mathbb{R}^3 \setminus \{0\})$$

mit

$$\nabla u = -\frac{x}{|x|^3}, \quad x \neq 0.$$

Wie in Beispiel 9.8.2.iii) gilt

$$\Delta u = \operatorname{div}(\nabla u) = 0 \text{ in } \mathbb{R}^3 \setminus \{0\};$$

das heisst, u ist harmonisch.

ii) Sei $n = 2$, $u \in C^\infty(\mathbb{R}^2)$ die Funktion

$$u(x, y) = e^x \sin(y), \quad (x, y) \in \mathbb{R}^2.$$

Es gilt:

$$\frac{\partial^2 u}{\partial x^2} = u, \quad \frac{\partial^2 u}{\partial y^2} = -u, \quad \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0;$$

also ist u harmonisch.

Mit Satz 9.8.2 erhalten wir die folgende variationelle Charakterisierung harmonischer Funktionen. Für $v \in C^1(\overline{\Omega})$ sei

$$E(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 \, d\mu$$

die **Dirichlet-Energie** von v .

Satz 9.8.3. (Variationelle Charakterisierung harmonischer Funktionen) Sei $\Omega \subset \mathbb{R}^n$ beschränkt mit $\partial\Omega \in C^1_{pw}$, $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$. Es sind äquivalent:

i) $\Delta u = 0$;

ii) $E(u) = \min\{E(v); v \in C^1(\overline{\Omega}), v = u \text{ auf } \partial\Omega\}$.

Beweis. i) \Rightarrow ii): Sei u harmonisch und sei $v \in C^1(\overline{\Omega})$ mit $v = u$ auf $\partial\Omega$. Setze $\varphi = v - u \in C^1(\overline{\Omega})$ mit $\varphi = 0$ auf $\partial\Omega$. Mit

$$|\nabla(u + \varphi)|^2 = |\nabla u|^2 + 2\nabla u \nabla \varphi + |\nabla \varphi|^2$$

folgt

$$E(v) = E(u + \varphi) = \frac{1}{2} \int_{\Omega} |\nabla(u + \varphi)|^2 \, d\mu = E(u) + E(\varphi) + \int_{\Omega} \nabla u \nabla \varphi \, d\mu. \quad (9.8.3)$$

Setzen wir $\nabla u =: K$ und beachten wir, dass $\varphi = 0$ auf $\partial\Omega$ sowie $\operatorname{div}(\nabla u) = \Delta u = 0$ in Ω , so folgt mit Satz 9.8.2

$$\int_{\Omega} \nabla u \nabla \varphi \, d\mu = \int_{\partial\Omega} \varphi \nabla u \cdot n \, d\sigma - \int_{\Omega} \varphi \operatorname{div}(\nabla u) \, d\mu = 0;$$

also

$$E(v) = E(u) + E(\varphi) \geq E(u).$$

ii) \Rightarrow i): Betrachte $v = u + t\varphi$, wobei $\varphi \in C^1(\overline{\Omega})$ mit $\varphi = 0$ auf $\partial\Omega$, $t \in \mathbb{R}$. Nach Annahme gilt

$$\forall t \in \mathbb{R} : E(u + t\varphi) \geq E(u).$$

Mit der Gleichung

$$E(u + t\varphi) = E(u) + t \int_{\Omega} \nabla u \nabla \varphi \, d\mu + t^2 E(\varphi)$$

gemäss (9.8.3) und Satz 9.8.2 folgt

$$\begin{aligned} 0 &= \left. \frac{d}{dt} \right|_{t=0} E(u + t\varphi) = \int_{\Omega} \nabla u \nabla \varphi \, d\mu \\ &= \underbrace{\int_{\partial\Omega} \varphi \nabla u \cdot n \, d\sigma}_{=0} - \int_{\Omega} \varphi \Delta u \, d\mu = - \int_{\Omega} \varphi \Delta u \, d\mu. \end{aligned} \quad (9.8.4)$$

Nimm widerspruchswise an, dass $\Delta u(x_0) \neq 0$ für ein $x_0 \in \Omega$. OBdA $\Delta u(x_0) > 0$. Da $u \in C^2(\Omega)$, gibt es $r > 0$ mit

$$\Delta u > 0 \text{ in } B_{2r}(x_0) \subset \Omega.$$

Wähle $\varphi \in C^\infty(\Omega)$ mit

$$\varphi(x) = \begin{cases} \exp\left(-\frac{1}{1-|x-x_0|^2/r^2}\right), & x \in B_r(x_0), \\ 0, & \text{sonst;} \end{cases}$$

vergleiche Beispiel 8.5.7.ii). Offenbar gilt $\varphi \in C^1(\overline{\Omega})$ und $\varphi = 0$ auf $\partial\Omega$ und

$$\int_{\Omega} \varphi \Delta u \, d\mu > 0$$

im Widerspruch zu (9.8.4). □

Setze

$$C_0^1(\Omega) = \{\varphi \in C^1(\overline{\Omega}); \varphi = 0 \text{ in einer Umgebung von } \partial\Omega\}.$$

Der Beweis von Satz 9.8.3 zeigt:

Satz 9.8.4. Für $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$ sind äquivalent:

i) u ist harmonisch, $\Delta u = 0$;

ii) $\forall \varphi \in C_0^1(\Omega) : \int_{\Omega} \nabla u \nabla \varphi \, d\mu = 0$.

Definition 9.8.3. Bedingung ii) in Satz 9.8.4 heisst **schwache Form** der Gleichung $\Delta u = 0$.

9.8.3 Transformation des Dirichlet-Integrals

Sei $U \subset \mathbb{R}^n$ offen, $\Phi \in C^2(U; \mathbb{R}^n)$ ein Diffeomorphismus von U auf $V = \Phi(U) \subset \mathbb{R}^n$ mit $\Psi = \Phi^{-1} \in C^2(V; \mathbb{R}^n)$. Sei weiter $\Omega \subset \bar{\Omega} \subset U$ beschränkt mit $\partial\Omega$ von der Klasse C^1_{pw} , und sei $\Omega' = \Phi(\Omega) \subset V$. Betrachte $u \in C^2(\Omega) \cap C^1(\bar{\Omega})$ und dazu $v = u \circ \Psi \in C^2(\Omega') \cap C^1(\bar{\Omega}')$. Berechne

$$\begin{aligned} |\nabla v|^2 &= \sum_{i=1}^n \left| \frac{\partial(u \circ \Psi)}{\partial y^i} \right|^2 \\ &= \sum_{i,j,k=1}^n \left(\frac{\partial u}{\partial x^j} \circ \Psi \right) \frac{\partial \Psi^j}{\partial y^i} \left(\frac{\partial u}{\partial x^k} \circ \Psi \right) \frac{\partial \Psi^k}{\partial y^i} = \sum_{j,k=1}^n (h^{jk} \frac{\partial u}{\partial x^j} \frac{\partial u}{\partial x^k}) \circ \Psi, \end{aligned}$$

wobei

$$\begin{aligned} h &= (h^{jk})_{1 \leq j,k \leq n} = \left(\sum_{i=1}^n \left(\frac{\partial \Psi^j}{\partial y^i} \frac{\partial \Psi^k}{\partial y^i} \right) \circ \Phi \right)_{1 \leq j,k \leq n} \\ &= (d\Psi d\Psi^t) \circ \Phi = (d\Phi)^{-1} (d\Phi^t)^{-1} = (d\Phi^t d\Phi)^{-1} = g^{-1} \end{aligned}$$

mit der von Φ induzierte Metrik

$$g = d\Phi^t d\Phi = (g_{ij}) = \left(\sum_{k=1}^n \frac{\partial \Phi^k}{\partial x^i} \frac{\partial \Phi^k}{\partial x^j} \right)_{1 \leq i,j \leq n}$$

nach Bemerkung 9.6.1.iii). Wir schreiben

$$h = g^{-1} =: (g^{ij})_{1 \leq i,j \leq n}.$$

Mit Satz 9.5.2 und der Identität $|\det(d\Phi)| = \sqrt{\det(g)}$ folgt

$$\begin{aligned} 2E(v) &= \int_{\Omega'} |\nabla v|^2 d\mu = \sum_{i,j=1}^n \int_{\Phi(\Omega)} \left((g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j}) \circ \Psi \right) d\mu \\ &= \sum_{i,j=1}^n \int_{\Omega} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j} \sqrt{\det(g)} d\mu =: 2E_g(u). \end{aligned}$$

Analog erhalten wir für $\varphi \in C^1(\bar{\Omega})$, $\psi = \varphi \circ \Psi \in C^1(\bar{\Omega}')$ mit $|g| = \det(g)$ die Identität

$$\begin{aligned} E(v + \psi) &= E(v) + E(\psi) + \int_{\Omega'} \nabla v \nabla \psi d\mu = E_g(u + \varphi) \\ &= E_g(u) + E_g(\varphi) + \sum_{i,j} \int_{\Omega} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial \varphi}{\partial x^j} \sqrt{|g|} d\mu; \end{aligned} \tag{9.8.5}$$

also

$$\int_{\Omega'} \nabla v \nabla \psi d\mu = \sum_{i,j=1}^n \int_{\Omega} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial \varphi}{\partial x^j} \sqrt{|g|} d\mu.$$

Sei $g(\cdot, \cdot) = (\cdot, \cdot)_g$ das von g induzierte Skalarprodukt mit

$$g(\xi, \eta) = \sum_{i,j} g_{ij} \xi^i \eta^j, \quad \xi = (\xi^i)_{1 \leq i \leq n}, \quad \eta = (\eta^j)_{1 \leq j \leq n} \in \mathbb{R}^n.$$

Definition 9.8.4. Der Gradient von u bezüglich g an einer Stelle $x \in \Omega$ ist der Vektor $\nabla_g u(x)$ mit

$$(\nabla_g u)^j = \sum_{i=1}^n g^{ij} \frac{\partial u}{\partial x^i}, \quad 1 \leq j \leq n.$$

Bemerkung 9.8.4. i) Aus Definition 9.8.4 folgt die zu Definition 8.3.2 analoge Identität

$$\forall w \in T_x \mathbb{R}^n \cong \mathbb{R}^n : (\nabla_g u(x), w)_g = du(x)w.$$

ii) Damit können wir deuten

$$\sum_{i,j} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial \varphi}{\partial x^j} = d\varphi(\nabla_g u) = (\nabla_g u, \nabla_g \varphi)_g;$$

insbesondere

$$E_g(u) = \frac{1}{2} \int_{\Omega} \sum_{i,j=1}^n g_{ij} \nabla_g u^i \nabla_g \varphi^j \sqrt{|g|} \, d\mu = \frac{1}{2} \int_{\Omega} (\nabla_g u, \nabla_g u)_g \sqrt{|g|} \, d\mu.$$

Satz 9.8.5. Sei $u \in C^2(\Omega) \cap C^1(\overline{\Omega})$, $u = v \circ \Phi$ mit $v \in C^2(\Omega') \cap C^1(\overline{\Omega}')$. Es sind äquivalent:

i) Die Funktion v ist harmonisch, $\Delta v = 0$;

ii) es gilt

$$\forall \varphi \in C_0^1(\Omega) : \int_{\Omega} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial \varphi}{\partial x^j} \sqrt{|g|} \, d\mu = 0,$$

wobei wir die Einsteinsche Summenkonvention verwenden;

iii) die Funktion u erfüllt die Gleichung

$$\Delta_g u = \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^i} \left(g^{ij} \sqrt{|g|} \frac{\partial u}{\partial x^j} \right) = 0.$$

Definition 9.8.5. Der Operator Δ_g heisst **Laplace-Beltrami-Operator** in der Metrik g .

Beweis von Satz 9.8.5. i) \Leftrightarrow ii): Mit Satz 9.8.3 und (9.8.5) ist die Identität $\Delta v = 0$ äquivalent zur Minimalität von $E_g(u)$. Analog zu (9.8.4) erhalten wir daraus die Identität ii), und umgekehrt.

ii) \Leftrightarrow iii): Analog zu (9.8.4) folgt aus der Identität ii) für $\varphi \in C^1(\Omega)$ mit $\varphi = 0$

auf $\partial\Omega$ mit Satz 9.8.2 die Gleichung

$$\begin{aligned} \int_{\Omega} \underbrace{\sqrt{|g|} g^{ij}}_{=K} \underbrace{\frac{\partial u}{\partial x^i} \frac{\partial \varphi}{\partial x^j}}_{=\nabla f} d\mu &= - \int \frac{1}{\sqrt{|g|}} \frac{\partial}{\partial x^j} \left(\sqrt{|g|} g^{ij} \frac{\partial u}{\partial x^i} \right) \varphi \sqrt{|g|} d\mu \\ &= - \int_{\Omega} \Delta_g u \varphi \sqrt{|g|} d\mu. \end{aligned}$$

Wie im Beweis von Satz 9.8.3 ist dies äquivalent zur Gleichung $\Delta_g u = 0$. \square

Beispiel 9.8.4. i) *Polarkoordinaten in \mathbb{R}^2 .* Wähle

$$U = \{(r, \theta); r > 0, -\pi < \theta < \pi\}, \quad V = \mathbb{R}^2 \setminus \{(x^1, 0); x^1 \leq 0\},$$

und sei $\Phi \in C^\infty(U; \mathbb{R}^2)$ die Abbildung mit

$$\Phi : \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} x^1 \\ x^2 \end{pmatrix} = \begin{pmatrix} r \cos(\theta) \\ r \sin(\theta) \end{pmatrix}.$$

Dann gilt

$$d\Phi(r, \theta) = \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix},$$

und wir erhalten

$$g = d\Phi^t d\Phi = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix}$$

mit $\sqrt{|g|} = \sqrt{\det(g)} = r$ und

$$g^{-1} = (g^{ij}) = \begin{pmatrix} 1 & 0 \\ 0 & r^{-2} \end{pmatrix}.$$

Für $u = v \circ \Phi \in C^2(\Omega) \cap C^1(\overline{\Omega})$, $\Omega \subset U$ folgt

$$E_g(u) = \frac{1}{2} \int_{\Omega} g^{ij} \frac{\partial u}{\partial x^i} \frac{\partial u}{\partial x^j} \sqrt{|g|} \underbrace{d\mu(r, \theta)}_{drd\theta} = \frac{1}{2} \int_{\Omega} \left(\left| \frac{\partial u}{\partial r} \right|^2 + \frac{1}{r^2} \left| \frac{\partial u}{\partial \theta} \right|^2 \right) r dr d\theta$$

und

$$\Delta_g u = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}.$$

ii) Die Funktion $u = u(r) = \log(r)$, $r > 0$ erfüllt $\Delta_g u = 0$; insbesondere ist $v(x) = \log(|x|)$ harmonisch auf $\mathbb{R}^2 \setminus \{0\}$.

9.9 k -Formen im \mathbb{R}^n

Seien $k, n \in \mathbb{N}$.

Definition 9.9.1. Eine k -lineare Abbildung $\lambda: \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{k\text{-mal}} \rightarrow \mathbb{R}$ heisst **alternierend** oder **schiefsymmetrisch**, falls für $v_1, \dots, v_k \in \mathbb{R}^n$ gilt:

$$\lambda(v_1, \dots, v_i, \dots, v_j, \dots, v_k) = -\lambda(v_1, \dots, v_j, \dots, v_i, \dots, v_k).$$

Setze

$$\Lambda^k \mathbb{R}^n = \{ \lambda: \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{k\text{-mal}} \rightarrow \mathbb{R}; \lambda \text{ } k\text{-linear und alternierend} \}.$$

Für $\mu \in \Lambda^k \mathbb{R}^n$, $\nu \in \Lambda^l \mathbb{R}^n$ sei $\mu \wedge \nu \in \Lambda^{k+l} \mathbb{R}^n$ die alternierende $(k+l)$ -lineare Abbildung mit

$$\begin{aligned} \mu \wedge \nu (x_1, \dots, x_{k+l}) \\ = C(k, l)^{-1} \sum_p \operatorname{sgn}(p) \mu(x_{p(1)}, \dots, x_{p(k)}) \nu(x_{p(k+1)}, \dots, x_{p(k+l)}), \end{aligned}$$

wobei wir über alle Permutationen p von $\{1, \dots, k+l\}$ summieren mit

$$\operatorname{sgn}(p) = (-1)^{\#\{\text{Permutationen in } p\}}$$

und mit $C(k, l) = C(k)C(l)$, wobei $C(k) = k!$ die Anzahl aller möglichen Permutationen von $\{1, \dots, k\}$ bezeichnet.

Definition 9.9.2.

$$\wedge: \Lambda^k \mathbb{R}^n \times \Lambda^l \mathbb{R}^n \rightarrow \Lambda^{k+l} \mathbb{R}^n$$

heisst **schiefes Produkt**.

Beispiel 9.9.1. i) Falls $k = 1$, so ist nach Beispiel 8.1.3.ii) die Familie der Differentiale dx^1, \dots, dx^n eine Basis für $\Lambda^1 \mathbb{R}^n \cong L(\mathbb{R}^n; \mathbb{R})$.

ii) Für $k = 2$ erhalten wir für $\Lambda^2 \mathbb{R}^n$ die Basis

$$dx^i \wedge dx^j, \quad 1 \leq i < j \leq n.$$

Beachte, dass gilt

$$dx^i \wedge dx^j(e_k, e_l) = \delta_k^i \delta_l^j - \delta_l^i \delta_k^j, \quad 1 \leq i, j \leq n, 1 \leq k, l \leq n;$$

insbesondere

$$dx^i \wedge dx^j = -dx^j \wedge dx^i, \quad 1 \leq i < j \leq n.$$

iii) Analog gilt für beliebige $k \leq n$:

$$\Lambda^k \mathbb{R}^n = \operatorname{span}\{dx^{i_1} \wedge \dots \wedge dx^{i_k}; 1 \leq i_1 < i_2 < \dots < i_k \leq n\}.$$

iv) Offenbar gilt:

$$\forall k > n: \Lambda^k \mathbb{R}^n = \{0\}.$$

Definition 9.9.3. Sei $\Omega \subset \mathbb{R}^n$ offen. Eine **k-Form** der Klasse C^m auf Ω ist eine Abbildung

$$x \mapsto \lambda(x) = \sum_{1 \leq i_1 < \dots < i_k \leq n} \lambda_{i_1, \dots, i_k}(x) dx^{i_1} \wedge \dots \wedge dx^{i_k} \in \Lambda^k T_x \mathbb{R}^n \cong \Lambda^k \mathbb{R}^n$$

mit Koeffizienten $\lambda_{i_1, \dots, i_k} \in C^m(\Omega)$. Wir schreiben $\lambda \in C^m(\Omega; \Lambda^k \mathbb{R}^n)$.

Beispiel 9.9.2. i) Die **Volumenform** im \mathbb{R}^n hat die Darstellung

$$\omega = dx^1 \wedge \dots \wedge dx^n \in C^\infty(\Omega; \Lambda^n \mathbb{R}^n).$$

ii) Für $n = 2$ gilt:

$$\omega = dx^1 \wedge dx^2 = -dx^2 \wedge dx^1.$$

iii) Im Falle $n = 3$ gilt

$$\omega = dx^1 \wedge dx^2 \wedge dx^3 = dx^2 \wedge dx^3 \wedge dx^1 = dx^3 \wedge dx^1 \wedge dx^2.$$

Sei

$$\lambda = \sum_{i_1 < \dots < i_k} \lambda_{i_1, \dots, i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in C^1(\Omega; \Lambda^k \mathbb{R}^n),$$

wobei $\Omega \subset \mathbb{R}^n$ offen.

Definition 9.9.4. Die **äußere Ableitung** von λ ist die $(k+1)$ -Form $d\lambda$ mit

$$\begin{aligned} d\lambda &= \sum_{i_1 < \dots < i_k} d\lambda_{i_1, \dots, i_k} \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} \\ &= \sum_{i_1 < \dots < i_k} \sum_{j=1}^n \frac{\partial \lambda_{i_1, \dots, i_k}}{\partial x^j} dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k}. \end{aligned}$$

Beispiel 9.9.3. i) Sei $n = 2$, $\lambda = g dx + h dy \in C^1(\Omega; \Lambda^1 \mathbb{R}^2)$. Dann gilt

$$\begin{aligned} d\lambda &= dg \wedge dx + dh \wedge dy = \frac{\partial g}{\partial y} dy \wedge dx + \frac{\partial h}{\partial x} dx \wedge dy \\ &= \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) dx \wedge dy = \left(\frac{\partial h}{\partial x} - \frac{\partial g}{\partial y} \right) \omega, \end{aligned}$$

wobei $\omega = dx \wedge dy = d\mu$ die Volumenform von \mathbb{R}^2 ist.

ii) Sei $n = 3$, $\varphi = \varphi^1 dy \wedge dz + \varphi^2 dz \wedge dx + \varphi^3 dx \wedge dy \in C^1(\mathbb{R}^3; \Lambda^1 \mathbb{R}^3)$. Dann gilt

$$\begin{aligned} d\varphi &= \frac{\partial \varphi^1}{\partial x^1} dx^1 \wedge dx^2 \wedge dx^3 - \frac{\partial \varphi^2}{\partial x^2} dx^2 \wedge dx^1 \wedge dx^3 + \frac{\partial \varphi^3}{\partial x^3} dx^3 \wedge dx^1 \wedge dx^2 \\ &= \sum_{i=1}^3 \frac{\partial \varphi^i}{\partial x^i} dx^1 \wedge dx^2 \wedge dx^3 = \operatorname{div}(\varphi) \omega, \end{aligned}$$

mit der Volumenform $\omega = dx^1 \wedge dx^2 \wedge dx^3$.

Satz 9.9.1. Es gilt $d^2 = 0$.

Beweis. Für $\lambda = \sum_{i_1 < \dots < i_k} \lambda_{i_1, \dots, i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in C^2(\Omega; \Lambda^k \mathbb{R}^n)$ gilt

$$d^2 \lambda = \sum_{i_1 < \dots < i_k} d^2 \lambda_{i_1, \dots, i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

Es genügt daher, den Fall $k = 0$ zu betrachten. Für $f \in C^2(\Omega)$ erhalten wir jedoch wegen Satz 8.5.1 und der Antisymmetrie von $dy^l \wedge dy^m$

$$d^2 f = d\left(\sum_{j=1}^n \frac{\partial f}{\partial x^j} dx^j\right) = \sum_{i,j=1}^n \underbrace{\frac{\partial^2 f}{\partial x^i \partial x^j}}_{\text{symm.}} \underbrace{dx^i \wedge dx^j}_{\text{antisymm.}} = 0,$$

analog zu Beispiel 9.4.5.ii). □

Definition 9.9.5. i) Eine k -Form λ heißt **geschlossen**, falls $d\lambda = 0$;

ii) λ heißt **exakt**, falls eine $(k-1)$ -Form ν existiert mit $\lambda = d\nu$.

Bemerkung 9.9.1. Nach Satz 9.9.1 ist eine exakte Form stets geschlossen.

9.9.1 Integration von k -Formen

Sei $U \subset \mathbb{R}^k$ offen, $\Phi = (\Phi^i)_{1 \leq i \leq n} \in C^1(U; \mathbb{R}^n)$ injektive Immersion, Ω beschränkt und Jordan-messbar mit $\Omega \subset \overline{\Omega} \subset U$, und sei $S = \Phi(\Omega) \subset M = \Phi(U)$.

Definition 9.9.6. Für $\lambda = \sum_{i_1 < \dots < i_k} \lambda_{i_1, \dots, i_k} dx^{i_1} \wedge \dots \wedge dx^{i_k} \in C^0(\overline{S}; \Lambda^k \mathbb{R}^n)$ heißt $\Phi^* \lambda \in C^0(\overline{\Omega}; \Lambda^k \mathbb{R}^k)$ mit

$$\Phi^* \lambda(v_1, \dots, v_k) \Big|_{w=w_0} = \lambda(\Phi(w_0))(d\Phi(w_0)v_1(w_0), \dots, d\Phi(w_0)v_k(w_0)),$$

oder

$$\Phi^* \lambda = \sum_{i_1 < \dots < i_k} (\lambda_{i_1, \dots, i_k} \circ \Phi) d\Phi^{i_1} \wedge \dots \wedge d\Phi^{i_k}$$

die mit Φ zurückgezogene (**pull-back**) k -Form auf Ω .

Beispiel 9.9.4. i) Sei $k = 1$, $\gamma \in C^1([0, 1]; \mathbb{R}^n)$, $\lambda \in C^0(\mathbb{R}^n; \Lambda^1 \mathbb{R}^n)$. Dann gilt

$$(\gamma^* \lambda)(t) = \lambda(\gamma(t)) \dot{\gamma}(t) dt$$

mit $\dot{\gamma}(t) = \frac{d\gamma(t)}{dt}$, und

$$\int_0^1 \gamma^* \lambda = \int_0^1 \lambda(\gamma(t)) \dot{\gamma}(t) dt = \int_\gamma \lambda.$$

ii) Sei $k = n$, $\Phi \in C^1(U; \mathbb{R}^n)$ ein Diffeomorphismus auf $V = \Phi(U)$, und sei $\omega = dx^1 \wedge \dots \wedge dx^n$ die Volumenform im \mathbb{R}^n . Dann gilt

$$\Phi^* \omega(e_1, \dots, e_n) = \omega\left(\frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^n}\right) = \det(d\Phi).$$

Somit können wir den Transformationssatz 9.5.1 für orientierungserhaltende Φ mit $\det(d\Phi) > 0$ in folgender Form schreiben:

$$\mu_n(\Phi(\Omega)) = \int_{\Omega} \det(d\Phi) d\mu = \int_{\Omega} \Phi^* \omega.$$

Insbesondere für $\Phi = id$ erhalten wir

$$\mu_n(\Omega) = \int_{\Omega} d\mu = \int_{\Omega} \omega;$$

das heisst, das "Volumenelement" $d\mu$ können wir auch in der Form $d\mu = \omega = dx^1 \wedge \dots \wedge dx^n$ darstellen.

Die vorangehenden Beispiele motivieren die folgende Definition

Definition 9.9.7. Für $S = \Phi(\Omega)$, $\lambda \in C^0(\overline{S}; \Lambda^k \mathbb{R}^n)$ wie oben setze

$$\int_S \lambda := \int_{\Omega} \Phi^* \lambda.$$

Bemerkung 9.9.2. i) Für Darstellungen Φ_1, Φ_2 derselben k -dimensionalen Untermannigfaltigkeit $S = \Phi_1(\Omega_1) = \Phi_2(\Omega_2) \subset \mathbb{R}^n$ definiert $\Phi_1^{-1} \circ \Phi_2 = f \in C^1(\Omega_2; \mathbb{R}^k)$ analog zu Bemerkung 9.6.1.iv) einen Diffeomorphismus. Nach Definition 9.9.5 erhalten wir für $\lambda \in C^0(\overline{S}; \Lambda^k \mathbb{R}^n)$ die Darstellungen $\Phi_{1,2}^* \lambda$ mit

$$\Phi_i^* \lambda(w)(e_1, \dots, e_k) = \lambda(\Phi_i(w))(d\Phi_i(w)e_1, \dots, d\Phi_i(w)e_k), \quad i = 1, 2.$$

Mit $\Phi_2 = \Phi_1 \circ f$ folgt

$$\Phi_2^* \lambda = f^*(\Phi_1^* \lambda) = ((\Phi_1^* \lambda) \circ f) \det(df).$$

Falls die Parametrisierungen Φ_1 und Φ_2 gleich orientiert sind im Sinne, dass $\det(df) > 0$, so gilt nach Satz 9.5.2 also

$$\int_{\Omega_1} \Phi_1^* \lambda = \int_{\Omega_2} \Phi_2^* \lambda;$$

das heisst, $\int_S \lambda$ ist unabhängig von der Darstellung $S = \Phi(\Omega)$ bei fester Orientierung von S .

ii) Sei $\Omega \subset \mathbb{R}^3$ ein Normalbereich

$$\Omega = \{x = (x^i); (x^1, x^2) \in Q, 0 \leq x^3 \leq \psi(x^1, x^2)\}$$

wie im Beweis vom Satz von Gauss, Satz 9.8.1, mit einem Quader $Q \subset \mathbb{R}^2$ und einer Funktion $0 \leq \psi \in C^1(\overline{Q})$. Für $K = (K^i)_{1 \leq i \leq 3} \in C^1(\overline{\Omega}; \mathbb{R}^3)$ sei

$$\lambda = K^1 dy \wedge dz + K^2 dz \wedge dx + K^3 dx \wedge dy \in C^1(\overline{\Omega}; \Lambda^1 \mathbb{R}^3)$$

wie in Beispiel 9.9.3.ii). Falls $K^1 = K^2 = 0$, so erhalten wir mit der Parametrisierung

$$\Phi(x^1, x^2) = (x^1, x^2, \psi(x^1, x^2)), \quad (x^1, x^2) \in Q,$$

des “oberen” Teils $S = \mathcal{G}(\psi) = \Phi(Q)$ von $\partial\Omega$ wegen

$$\Phi^*\lambda = (K^3 \circ \Phi)d\Phi^1 \wedge d\Phi^2 = (K^3 \circ \Phi)dx^1 \wedge dx^2$$

die Gleichheit

$$\int_S \lambda = \int_Q \Phi^*\lambda = \int_Q K^3(x^1, x^2, \psi(x^1, x^2)) d\mu_2(x^1, x^2).$$

Analog ergibt die Parametrisierung $\Phi = id$ des unteren Teils von $\partial\Omega$ die Identität

$$\int_{Q \times \{0\}} \lambda = \int_Q \lambda = \int_Q K^3(x^1, x^2, 0) d\mu_2(x^1, x^2).$$

Verlangen wir, dass mit der äusseren Normale n für die Parametrisierung Φ eines jeden Teils von $\partial\Omega$ gelten soll

$$\det\left(n, \frac{\partial\Phi}{\partial x^1}, \frac{\partial\Phi}{\partial x^2}\right) > 0,$$

so folgt mit Beispiel 9.9.3.ii) und der Rechnung (9.8.1) im Beweis von Satz 9.8.1

$$\begin{aligned} \int_{\Omega} d\lambda &= \int_{\Omega} \frac{\partial K^3}{\partial x^3} d\mu_3 \\ &= \int_Q (K^3(x^1, x^2, \psi(x^1, x^2)) - K^3(x^1, x^2, 0)) d\mu_2(x^1, x^2) = \int_{\partial\Omega} \lambda, \end{aligned}$$

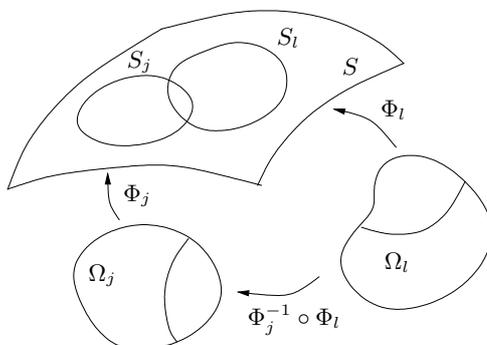
da die übrigen Randkomponenten nichts beitragen. Zusammen mit dem entsprechenden Resultat für die Komponenten K^1, K^2 erhalten wir für allgemeines λ die Identität

$$\int_{\Omega} d\lambda = \int_{\partial\Omega} \lambda.$$

Für eine allgemeine k -dimensionale Untermannigfaltigkeit S kann man Orientierbarkeit analog zu Bemerkung 9.9.2.i) wie folgt erklären.

Definition 9.9.8. S heisst orientierbar, falls $S \subset \bigcup_{j=1}^J M_j$, mit $M_j = \Phi_j(\Omega_j)$ mit injektiven Immersionen $\Phi_j \in C^1(U_j; \mathbb{R}^n)$ auf offenen Mengen $\Omega_j \subset \bar{\Omega}_j \subset U_j \subset \mathbb{R}^k$, $1 \leq j \leq J$, wobei für alle $1 \leq j, l \leq J$ auf dem Definitionsbereich des “Kartenwechsels” $\Phi_j^{-1} \circ \Phi_l$ gilt

$$\det(d(\Phi_j^{-1} \circ \Phi_l)) > 0.$$



Bemerkung 9.9.3. i) Für eine orientierbare, kompakte k -dimensionale Untermannigfaltigkeit S ist dann $\int_S \lambda$ mittels Definition 9.9.7 für jede k -Form $\lambda \in C^0(\overline{S}; \Lambda^k \mathbb{R}^n)$ bis auf das Vorzeichen wohldefiniert.

ii) Im Falle $k = 2$ induziert die Wahl orientierter “Karten” $(\Phi_j)_{1 \leq j \leq J}$ gemäss Definition 9.9.8 für derartiges S analog zu Beispiel 9.7.1 eine damit kompatible Orientierung von ∂S .

9.9.2 Der Satz von Stokes im \mathbb{R}^n

Sei $U \subset \mathbb{R}^k$ offen, $\Phi \in C^2(U; \mathbb{R}^n)$ injektive Immersion, Ω beschränkt mit $\partial\Omega \in C^1_{pw}$ und $\Omega \subset \overline{\Omega} \subset U$, $S = \Phi(\Omega) \subset M = \Phi(U)$ orientierbar, und sei $\lambda \in C^1(W; \Lambda^l \mathbb{R}^n)$ in einer Umgebung W von \overline{S} mit $1 \leq k, l \leq n$.

Lemma 9.9.1. *Es gilt $d(\Phi^* \lambda) = \Phi^*(d\lambda) \in C^0(\overline{\Omega}; \Lambda^{l+1} \mathbb{R}^2)$.*

Beweis. Der Einfachheit halber beschränken wir uns auf den Fall $l = 1$. Sei

$$\lambda = \sum_{j=1}^n \lambda_j dx^j$$

mit

$$d\lambda = \sum_{j=1}^n d\lambda_j \wedge dx^j = \sum_{i,j=1}^n \frac{\partial \lambda_j}{\partial x^i} dx^i \wedge dx^j = \sum_{i < j} \left(\frac{\partial \lambda_j}{\partial x^i} - \frac{\partial \lambda_i}{\partial x^j} \right) dx^i \wedge dx^j.$$

Es folgt

$$\Phi^*(d\lambda) = \sum_{i < j} \left(\left(\frac{\partial \lambda_j}{\partial x^i} - \frac{\partial \lambda_i}{\partial x^j} \right) \circ \Phi \right) d\Phi^i \wedge d\Phi^j.$$

Weiter gilt

$$\Phi^* \lambda = \sum_{j=1}^n (\lambda_j \circ \Phi) d\Phi^j$$

mit $d\Phi^j = \sum_{l=1}^k \frac{\partial \Phi^j}{\partial y^l} dy^l$. Da andererseits $d^2 \Phi^j = 0$ gemäss Satz 9.9.1, folgt

$$\begin{aligned} d(\Phi^* \lambda) &= \sum_{j=1}^n d(\lambda_j \circ \Phi) d\Phi^j = \sum_{i,j=1}^n \left(\frac{\partial \lambda_j}{\partial x^i} \circ \Phi \right) d\Phi^i \wedge d\Phi^j \\ &= \sum_{i < j} \left(\left(\frac{\partial \lambda_j}{\partial x^i} - \frac{\partial \lambda_i}{\partial x^j} \right) \circ \Phi \right) d\Phi^i \wedge d\Phi^j = \Phi^*(d\lambda), \end{aligned}$$

wie gewünscht; analog für $l \geq 2$. □

Satz 9.9.2. (Stokes) *Sei S eine kompakte, orientierbare 2-dimensionale Untermannigfaltigkeit von \mathbb{R}^n der Klasse C^2 mit relativem Rand ∂S von der Klasse C^1_{pw} , und sei $\lambda \in C^1(W; \Lambda^1 \mathbb{R}^n)$ in einer Umgebung W von S . Dann gilt*

$$\int_{\partial S} \lambda = \int_S d\lambda,$$

wobei ∂S verträglich mit S orientiert ist gemäss Bemerkung 9.9.3.ii).

Beweis. i) Sei zunächst $S = \Phi(\Omega)$, wobei $\Phi \in C^2(U; \mathbb{R}^n)$ injektive Immersion auf einer offenen Menge $U \subset \mathbb{R}^2$, $\Omega \subset \bar{\Omega} \subset U$ beschränkt mit $\partial\Omega \in C_{pw}^1$ und $\Phi: \partial\Omega \rightarrow \partial S$ mit der Orientierung von ∂S verträglich. Dann gilt gemäss Definition 9.9.6, Lemma 9.9.1 und Satz 9.4.2 die Identität

$$\int_{\partial S} \lambda = \int_{\partial\Omega} \Phi^* \lambda = \int_{\Omega} d(\Phi^* \lambda) = \int_{\Omega} \Phi^*(d\lambda) = \int_S d\lambda,$$

wie behauptet.

ii) Für den allgemeinen Fall überdecke S mit paarweise disjunkten Flächenstücken $S_j = \Phi_j(\Omega_j) \subset M_j = \Phi_j(U_j)$, wobei $\Phi_j \in C^2(U_j; \mathbb{R}^n)$ injektive Immersion einer offenen Menge $U_j \subset \mathbb{R}^2$, $\Omega_j \subset \bar{\Omega}_j \subset U_j$ beschränkt mit C_{pw}^1 -Rand, $1 \leq j \leq J$. Weiter seien alle M_j sowohl untereinander als auch mit der von ∂S induzierten Orientierung verträglich orientiert. Dann gilt

$$\int_S d\lambda = \sum_j \int_{S_j} d\lambda = \sum_j \int_{\partial S_j} \lambda = \int_{\partial S} \lambda,$$

da sich die Wegintegrale längs gemeinsamer (“innerer”) Kanten von S_j und S_l für $j \neq l$ aufheben. \square

Es folgt ein zu Satz 9.7.2 analoges Kriterium für exakte 1-Formen.

Satz 9.9.3. (Poincaré) Sei $W \subset \mathbb{R}^n$ offen und einfach zusammenhängend, $\lambda \in C^1(W; \Lambda^1 \mathbb{R}^n)$. Dann sind äquivalent:

- i) λ ist exakt;
- ii) $d\lambda = 0$.

Beweis. $i) \Rightarrow ii)$: Nach Bemerkung 9.9.1 ist jede exakte 1-Form geschlossen.

$ii) \Rightarrow i)$: Sei $\lambda \in C^1(W; \Lambda^1 \mathbb{R}^n)$ geschlossen, $\Gamma \subset W$ einfach geschlossen, $S \subset W$ orientiertes Flächenstück mit $\partial S = \Gamma$ gemäss Definition 9.7.5. Dann gilt nach Satz 9.9.2

$$\int_{\Gamma} \lambda = \pm \int_{\partial S} \lambda = \pm \int_S d\lambda = 0;$$

und nach Satz 8.4.2 existiert $f \in C^2(W)$ mit $\lambda = df$; das heisst, λ ist exakt. \square

Sei S eine orientierbare, kompakte k -dimensionale Untermannigfaltigkeit im \mathbb{R}^n . Analog zu Definition 9.7.2 können wir eine Orientierung von S wie folgt erklären.

Definition 9.9.9. Sei $l = n - k$. Eine **Orientierung** von S ist eine Familie

$$\nu = (\nu_1, \dots, \nu_l) \in C^0(\bar{S}; \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{l\text{-mal}})$$

von punktweise orthonormalen Vektorfeldern $\nu_i(p) \in T_p^\perp S$, $p \in S$, $1 \leq i \leq l$.

Sei nun $S = \Phi(\Omega)$, wobei $\Phi \in C^2(U; \mathbb{R}^n)$ injektive Immersion auf einer offenen Menge $U \subset \mathbb{R}^k$, $\Omega \subset \bar{\Omega} \subset U$ beschränkt mit $\partial\Omega \in C_{pw}^1$, und sei $\nu = (\nu_1, \dots, \nu_l) \in C^0(\bar{S}; \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{l\text{-mal}})$ eine Orientierung für S .

Definition 9.9.10. Die Parametrisierung Φ von S ist mit der Orientierung $\nu = (\nu_1, \dots, \nu_l)$ **verträglich**, falls gilt

$$\det(\nu_1 \circ \Phi, \dots, \nu_l \circ \Phi, \frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^k}) > 0 \text{ in } \bar{\Omega}.$$

Bemerkung 9.9.4. i) Sei $S = \Phi(\Omega) \subset M = \Phi(U)$ wie oben, und sei die Parametrisierung Φ mit der Orientierung $\nu = (\nu_1, \dots, \nu_l)$ verträglich. Mit der Volumenform ω im \mathbb{R}^n erhalten wir

$$\begin{aligned} & \det(\nu_1 \circ \Phi, \dots, \nu_l \circ \Phi, \frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^k}) \\ &= \omega(\nu_1 \circ \Phi, \dots, \nu_l \circ \Phi, \frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^k}) = \Phi^*(i(\nu)\omega), \end{aligned}$$

wobei $i(\nu)\omega$ die **mit ν kontrahierte Volumenform** bezeichnet, mit

$$(i(\nu)\omega)(X_1, \dots, X_k) = \omega(\nu_1, \dots, \nu_l, X_1, \dots, X_k)$$

für beliebige Vektorfelder X_1, \dots, X_k in W .

ii) Seien $S = \Phi(\Omega) \subset M = \Phi(U)$, $\nu = (\nu_1, \dots, \nu_l)$ wie in i). Für $p = \Phi(x) \in \partial S$ mit $x \in \partial\Omega$ sei weiter

$$\nu_{l+1}(p) \in T_p M \cap T_p^\perp \partial S$$

die äussere Normale an S in M . OBdA gelte

$$\nu_{l+1}(p) = \frac{\partial \Phi(x)}{\partial x^1}.$$

Dann ist $\bar{\nu} = (\nu_1, \dots, \nu_{l+1})$ eine Orientierung für ∂S , und wegen

$$\begin{aligned} & \det(\nu_1 \circ \Phi, \dots, \nu_l \circ \Phi, \nu_{l+1} \circ \Phi, \frac{\partial \Phi}{\partial x^2}, \dots, \frac{\partial \Phi}{\partial x^k}) \\ &= \det(\nu_1 \circ \Phi, \dots, \nu_l \circ \Phi, \frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^k}) > 0 \end{aligned}$$

ist $\Phi|_{\partial\Omega}$ mit $\bar{\nu}$ verträglich.

iii) Wie in Bemerkung 9.6.1.iii) kann man zeigen, dass für mit $\nu = (\nu_1, \dots, \nu_l)$ verträgliches Φ gilt

$$\Phi^*(i(\nu)\omega)(e_1, \dots, e_k) = \omega(\nu_1 \circ \Phi, \nu_2 \circ \Phi, \dots, \nu_l \circ \Phi, \frac{\partial \Phi}{\partial x^1}, \dots, \frac{\partial \Phi}{\partial x^k}) = \sqrt{\det(g)},$$

wobei $g = d\Phi^t d\Phi$.

Aufgrund der letzten Bemerkung können wir analog zu Definition 9.6.1 im Falle $k = 2$ nun für beliebiges $k \in \mathbb{N}$ das k -dimensionale Volumen von S definieren.

Definition 9.9.11. Sei $S = \Phi(\Omega)$ wie oben mit Orientierung $\nu = (\nu_1, \dots, \nu_l)$, und sei Φ mit ν verträglich. Dann bezeichnet

$$\mu_k(S) := \int_S i(\nu)\omega = \int_\Omega \Phi^*(i(\nu)\omega) = \int_\Omega \sqrt{\det(g)} \, d\mu_k$$

das k -dimensionale Volumen von S .

Analog zu Bemerkung 9.6.1.iv) folgt, dass das k -dimensionale Volumen von S unabhängig von der Parametrisierung von S erklärt ist.

9.9.3 Der allgemeine Satz von Stokes

Nach diesen Vorbereitungen können wir nun Satz 9.8.1 auf beliebige Dimensionen $n \in \mathbb{N}$ erweitern. Sei $\Omega \subset \mathbb{R}^n$ offen und beschränkt mit $\partial\Omega \in C_{pw}^1$ und sei N die äussere Normale entlang $\partial\Omega$, $K = (K^i)_{1 \leq i \leq n} \in C^1(\overline{\Omega}; \mathbb{R}^n)$.

Wie im Beweis von Satz 9.8.1 sei zunächst Ω ein Normalbereich bzgl. aller Achsen x^1, \dots, x^n . Vereinfachend gelte wieder mit einem Quader $Q \subset \mathbb{R}^{n-1}$ und einer Funktion $0 \leq \psi \in C^1(Q)$ die Darstellung

$$\Omega = \{x = (x^i); x' = (x^1, \dots, x^{n-1}) \in Q, 0 \leq x^n \leq \psi(x^1, \dots, x^{n-1})\}.$$

Analog zu (9.8.1) folgt

$$\int_{\Omega} \frac{\partial K^n}{\partial x^n} d\mu_n = \int_Q (K^n(x', \psi(x')) - K^n(x', 0)) d\mu_{n-1}(x').$$

Mit der Parametrisierung $\Phi(x') = (x', \psi(x'))^t \in C^1(Q; \mathbb{R}^n)$, $x' \in Q$, des oberen Teils $S = \mathcal{G}(\psi) = \Phi(Q)$ von $\partial\Omega$ erhalten wir

$$N = \frac{1}{\sqrt{1 + |\nabla\psi|^2}} \begin{pmatrix} -\partial\psi/\partial x^1 \\ \vdots \\ -\partial\psi/\partial x^{n-1} \\ 1 \end{pmatrix}.$$

Weiter ist $v_1 = d\psi^t$ Eigenvektor von $g = d\Phi^t d\Phi$ zum Eigenwert $1 + |\nabla\psi|^2$, und jeder Vektor $v \perp v_1$ ist offenbar Eigenvektor von g zum Eigenwert 1; also

$$\det(g) = 1 + |\nabla\psi|^2.$$

Falls $K^1 = \dots = K^{n-1} = 0$, so folgt

$$\int_S K \cdot N do_{n-1} = \int_Q K^n(x', \psi(x')) d\mu_{n-1}(x'),$$

wobei do_{n-1} das $n-1$ -dimensionale Oberflächenelement auf S ist mit der Darstellung $do_{n-1} = \sqrt{\det(g)} d\mu_{n-1}$ bzgl. Φ gemäss Bemerkung 9.9.4.iii), und

$$\int_{\Omega} \frac{\partial K^n}{\partial x^n} d\mu_n = \int_{\partial\Omega} K \cdot N do_{n-1}$$

wie im Beweis von Satz 9.8.1. Argumentieren wir weiter analog zu Satz 9.8.1, so erhalten wir das folgende Resultat.

Satz 9.9.4. (Gauss) Sei $\Omega \subset \mathbb{R}^n$ beschränkt und von der Klasse C_{pw}^1 mit äusserem Normalenvektor N , und sei $K \in C^1(\overline{\Omega}; \mathbb{R}^n)$. Dann gilt

$$\int_{\Omega} \operatorname{div}(K) d\mu_n = \int_{\partial\Omega} K \cdot N do_{n-1},$$

wobei N die äussere Normale ist.

Bemerkung 9.9.5. Für ein Gebiet $\Omega \subset \mathbb{R}^n$ und $K = (K^i)_{1 \leq i \leq n} \in C^1(\overline{\Omega}; \mathbb{R}^n)$ wie in Satz 9.9.4 sei

$$\lambda = \sum_{i=1}^n (-1)^{i-1} K^i dx^1 \wedge \dots \wedge \widehat{dx^i} \wedge \dots \wedge dx^n,$$

wobei der Term $\widehat{dx^i}$ im jeweiligen Produkt entfernt wird. Es gilt

$$\begin{aligned} d\lambda &= \sum_{i=1}^n (-1)^{i-1} \frac{\partial K^i}{\partial x^i} dx^i \wedge dx^1 \wedge \dots \wedge \widehat{dx^i} \wedge \dots \wedge dx^n \\ &= \left(\sum_{i=1}^n \frac{\partial K^i}{\partial x^i} \right) dx^1 \wedge \dots \wedge dx^n = \operatorname{div}(K) d\mu_n. \end{aligned}$$

wie in Beispiel 9.9.3.ii). Falls $K^1 = \dots = K^{n-1} = 0$, so erhalten wir mit der Parametrisierung

$$\Phi(x') = (x', \psi(x')), \quad x' \in Q,$$

des “oberen” Teils $S = \mathcal{G}(\psi) = \Phi(Q)$ von $\partial\Omega$ wegen

$$\Phi^* \lambda = (K^n \circ \Phi) d\Phi^1 \wedge \dots \wedge d\Phi^{n-1} = (K^n \circ \Phi) dx^1 \wedge \dots \wedge dx^{n-1}$$

andererseits die Darstellung

$$\int_S \lambda = \int_Q \Phi^* \lambda = \int_Q K^n(x', \psi(x')) d\mu_{n-1}(x') = \int_S K \cdot N d\sigma_{n-1};$$

analog für den unteren Teils von $\partial\Omega$. Da die übrigen Randkomponenten nichts beitragen, folgt die Gleichheit

$$\int_{\Omega} d\lambda = \int_{\partial\Omega} \lambda.$$

Zusammen mit dem entsprechenden Resultat für die übrigen Komponenten erkennen wir, dass wir Satz 9.9.4 auch in der Form

$$\int_{\Omega} d\lambda = \int_{\partial\Omega} \lambda.$$

für allgemeines $\lambda \in C^1(\overline{\Omega}; \Lambda^{n-1}\mathbb{R}^n)$ schreiben können.

Aus Bemerkung 9.9.5 erhalten wir den allgemeinen Stokesschen Satz.

Satz 9.9.5. (Stokes) Sei S eine kompakte, orientierbare k -dimensionale Untermannigfaltigkeit von \mathbb{R}^n der Klasse C^2 mit Orientierung $\nu = (\nu_1, \dots, \nu_k)$ und relativem Rand ∂S von der Klasse C^1_{pw} , und sei $\lambda \in C^1(W; \Lambda^{k-1}\mathbb{R}^n)$ in einer Umgebung W von S . Dann gilt

$$\int_{\partial S} \lambda = \int_S d\lambda,$$

wobei ∂S verträglich mit S orientiert ist gemäss Bemerkung 9.9.4.ii).

Beweis. Der Beweis ist vollkommen analog zum Beweis von Satz 9.9.2. Sei zunächst $S = \Phi(\Omega)$, wobei $\Phi \in C^2(U; \mathbb{R}^n)$ injektive Immersion auf einer offenen Menge $U \subset \mathbb{R}^k$, $\Omega \subset \overline{\Omega} \subset U$ beschränkt mit $\partial\Omega \in C_{pw}^1$ und $\Phi: \Omega \rightarrow S$ mit der Orientierung von S verträglich. Dann gilt gemäss Definition 9.9.6, Lemma 9.9.1 und Bemerkung 9.9.5 die Gleichheit

$$\int_{\partial S} \lambda = \int_{\partial\Omega} \Phi^* \lambda = \int_{\Omega} d(\Phi^* \lambda) = \int_{\Omega} \Phi^*(d\lambda) = \int_S d\lambda,$$

wie behauptet. Den allgemeinen Fall erhält man nun genauso wie im Beweis von Satz 9.9.2. \square

Index

- k -Form, 303
- Äquivalenzklasse, 13
- Äquivalenzrelation, 12
- äquivalente Normen, 92
- äussere Ableitung, 271, 303
- überdeckungskompakt, 72

- Abbildung, 8
- abelsche Gruppe, 19
- abgeschlossen, 67
- Ableitung, 109
- Abschluss, 68
- Absolutbetrag, 36
- absolute Konvergenz, 59
- abzählbar, 29
- Additionstheorem für Log, 100
- algebraisch vollständig, 37
- alternierende Abbildung, 302
- alternierende harmonische Reihe, 51
- analytische Fortsetzung, 139
- analytische Funktion, 138
- Anfangswertproblem, 142
- Argument, 120
- Aussagen, 3
- Auswahlaxiom, 15
- Axiom, 3

- Banachraum, 106
- Bernoullische Ungleichung, 28
- beschränkt, 53
- beschränkt, nach oben, 23
- beschränkt, nach unten, 23
- beschränkte Folge, 44
- bijektiv, 10

- Cauchy-Folge, 50, 52
- Cauchy-Schwarzsche Ungleichung, 136
- Cavalierisches Prinzip, 262
- charakteristisches Polynom, 146

- Dedekindscher Schnitt, 18
- Descartesches Blatt, 236

- dichte Menge, 95
- Diffeomorphismus, 243, 276
- Differential, 199
- Differentialform, 210
- Differentialgleichungen, 140
- Differentialgleichungssystem, 141
- differenzierbar, 109, 199, 226
- Dirichlet-Energie, 297
- divergente Folge, 40
- Divergenz, 292

- Einbettung, 250
- einfach zusammenhängend, 273, 291
- Einheitswurzel, 121
- Einsteinsche Summenkonvention, 200
- Elementarfigur, 261
- Elementarinhalt, 253
- Epigraph, 132
- Euklidische Norm, 33
- Eulersche Zahl, 45
- exakte 1-Form, 214
- exakte Form, 304
- Exponentialreihe, 56

- Feinheit, 253
- Fibonacci Zahlen, 39
- Flächeninhalt, 283
- Fluss, 275, 286
- folgenkompakt, 72
- Folgenkriterium, 86
- Fundamentallösung, 144, 194
- Funktion, 8

- geometrische Reihe, 28, 39, 54
- gerichtetes Längenelement, 217
- geschlossene Form, 304
- glatte Funktion, 136
- Gleichgewichtslösung, 153
- gleichmächtig, 29
- gleichmässige Konvergenz, 102
- gleichmassig stetig, 94

- Gradient, 300
 Gradientenfeld, 210
 Grenzwert, 81
 Gruppe, 19

 Häufungspunkt, 46
 harmonische Funktion, 297
 harmonische Reihe, 51, 54
 Hausdorffraum, 80
 Hausdorffsches Trennungsaxiom, 80
 Heine-Borel-Eigenschaft, 72
 Hesse-Matrix, 222
 Hessische Form, 223
 Homöomorphismus, 91

 Identivität, 14
 Imaginärteil, 36
 Immersion, 249
 Immersionsatz, 249
 induktive Menge, 26
 Infimum, 23
 injektiv, 10
 innerer Punkt, 65
 Integral, 155
 inverse Abbildung, 11

 Jordan-messbar, 260
 Jordan-Normalform, 145
 Jordan-Nullmenge, 263
 Jordansches Mass, 260

 kanonische Basis, 32
 kompakt, 72
 komplexe Multiplikation, 35
 Konjugation, 36
 konjugiert konvexe Funktion, 134
 konjugierte Exponenten, 135
 konservatives Vektorfeld, 217
 kontrahierende Abbildung, 186
 kontrahierte Form, 309
 Kontraktionsprinzip, 186
 konvergente Folge, 40
 Konvergenzradius, 58
 konvex, 132
 konvexe Menge, 76
 kritischer Punkt, 223, 241
 kritischer Wert, 223

 Lagrange-Multiplikator, 241
 Lagrangefunktion, 241
 Landau-Symbole, 202

 Laplace-Beltrami-Operator, 300
 Laplace-Operator, 296
 Leibniz-Reihe, 52
 Limes inferior, 47
 Limes superior, 47
 linear geordnet, 15
 Linearkombination, 32
 Lipschitz stetig, 84
 Lipschitz-Konstante, 84
 lokal Lipschitz stetig, 86
 lokale Lösung, 185
 lokale Minimalstelle, 130

 maximales Element, 15
 Maximum, 24
 Metrik, 34, 52
 metrisch vollständig, 52
 metrischer Raum, 52
 Minimum, 24
 Minkowski-Ungleichung, 136
 monoton fallend, 44
 monoton wachsend, 44, 98

 natürlicher Logarithmus, 100
 negativ definit, 223
 negativ orientiert, 272, 288
 Norm, 90
 Normalbereich, 269
 Normalraum, 246
 Normkonvergenz, 103
 null-homotop, 291

 obere Schranke, 15
 oberes Riemann-Integral, 169
 offen, 65
 offener Ball, 65
 offener Kern, 68
 Ordnungsgraph, 14
 Ordnungsrelation, 14
 Ordnungsvollständigkeit, 20
 orientierbare Untermannigfaltigkeit, 306
 Orientierung, 287, 308
 orientierungserhaltend, 212
 orientierungsumkehrend, 212
 orthogonal, 33
 orthonormal, 33

 Partialbruchzerlegung, 164
 Partialsummen, 54
 partikuläre Lösung, 150

- Polarform, 120
Polarkoordinaten, 120, 279
positiv definit, 222
positiv orientiert, 272, 288
Potential, 217
Potentialfeld, 217
Potenzreihe, 57
pull-back, 304
punktweise Konvergenz, 102
- Quader, 253
Quellstärke, 275
- Rand, 68
Rang, 235
Realteil, 36
Rechtssystem, 272, 288
reell analytisch, 225
Reflexivität, 12, 14
regulärer Punkt, 235
regulärer Wert, 243
relativ abgeschlossen, 77
relativ kompakt, 78
relativ offen, 77
relativer Rand, 287
Relativtopologie, 77
Richtungsableitung, 206
Riemann-integrabel, 170, 255
Riemann-Integral, 170, 255
Riemannsche Fläche, 121
Rotation, 289
- schiefes Produkt, 302
Schranke, obere, 23
Schranke, untere, 23
schwach harmonisch, 299
senkrecht, 33
separabel, 107
Separation der Variablen, 161
singulärer Wert, 243
skalares Flächenelement, 283
Skalarprodukt, 33
Stammfunktion, 155
Standardbasis, 32
stetig, 82, 88
stetig ergänzbar, 82
stetige Fortsetzung, 94
Supremum, 23
Supremumsnorm, 90
surjektiv, 10
- Symmetrie, 12
- Tangentialraum, 210, 244
Taylor-Polynom, 129, 222
Taylor-Reihe, 136, 225
Teilfolge, 46
Topologie, 67
topologischer Raum, 79
topologisches Kriterium, 86
total geordnet, 15
Transitivität, 12, 14
Treppenfunktion, 167, 253
- Umgebung, 70, 79, 85
Umkehrabbildung, 11
Umkehrfunktion, 12
unbestimmtes Integral, 156
uneigentlich R-integrabel, 180
unteres Riemann-Integral, 169
Untermannigfaltigkeit, 243
Untermannigfaltigkeit mit Rand, 287
Urbildfunktion, 11
- Variation-der-Konstanten Formel, 162
Vektorraum, 32
Verfeinerung, 254
Verzinsung, 39
Vollständigkeitsaxiom, 20
Volumenform, 271, 303
- Wahrheitstafel, 4
Wegintegral, 212
wegzusammenhängend, 101
Weierstrass-Kriterium, 86
Wirbelstärke, 272
- Zahlen, ganze, 17
Zahlen, natürliche, 17
Zahlen, rationale, 17
Zahlenstrahl, 17
Zahlkörper, 17
Zentralfeld, 217
Zerlegung, 253
Zirkulation, 272
zusammenhängend, 75