Hans-Peter Gittel

Analysis 2

(für Physiker)

Sommersemester 2016 und Wintersemester 2016/2017

Universität Leipzig, Institut für Mathematik

Version vom 25. November 2016

Es wird keinerlei Gewähr für die Richtigkeit und Vollständigkeit der hier vorliegenden Mitschrift gegeben.

Diese Mitschrift wurde ursprünglich von Felix Dietzsch mit Hilfe von LaTeX 2_ε im Sommersemester 2008 gesetzt. Die jeweils aktuellste Version der Mitschrift ist erhältlich unter:

`http://www.math.uni-leipzig.de/~gittel/ma2phys16.html`

Inhaltsübersicht

1 Metrische Räume **1**

2 Differentialrechnung im $\mathbb{R}^N$ **41**

3 Integralrechnung im $\mathbb{R}^N$ **71**

Inhaltsverzeichnis

1 Metrische Räume **1**

1.1 Topologische Grundbegriffe 1

1.1.1 Der allgemeine Abstandsbegriff 1

1.1.2 Konvergenz 2

1.1.3 Klassifikation von Punkten 3

1.1.4 Klassifikation von Mengen 4

1.1.4.1 Eigenschaften offener Mengen 5

1.1.4.2 Eigenschaften abgeschlossener Mengen 5

1.2 Vollständigkeit und Kompaktheit 6

1.2.1 Vollständige metrische Räume 6

1.2.2 Kompakte Mengen 8

1.3 Funktionenfolgen und Funktionenreihen 9

1.3.1 Problemstellung 9

1.3.2 Gleichmäßige Konvergenz von Funktionenfolgen 10

1.3.3 Gleichmäßige Konvergenz von Funktionenreihen 12

1.3.4 Anwendung des neuen Konvergenzbegriffes 13

1.3.4.1 Potenzreihen 13

1.3.4.2 Fourierreihen 14

1.4 Stetige Abbildungen metrischer Räume 17

1.4.1 Definitionen und Beispiele 17

1.4.2 Charakterisierung stetiger Abbildungen 19

1.4.3 Weitere Eigenschaften stetiger Abbildungen 19

1.4.4 Kontraktive Abbildungen 21

1.4.4.1 Kontraktionen 21

1.4.4.2 Fixpunkte 22

1.4.4.3 Anwendungen 25

1.5 Systeme gewöhnlicher Differentialgleichungen erster Ordnung 25

1.5.1 Allgemeine Existenz- und Eindeutigkeitssätze 25

1.5.1.1 Eine Differentialgleichung 25
1.5.1.2 Differentialgleichungssysteme 27
1.5.1.3 Bedeutung von Differentialgleichungssystemen 1. Ordnung . . 28
1.5.2 Lineare Differentialgleichungssysteme 29
1.5.2.1 Struktur der Lösungen 30
1.5.2.2 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten 33
1.5.2.3 Anwendung auf lineare Differentialgleichungen n-ter Ordnung mit konstanten Ko

2 Differentialrechnung im $\mathbb{R}^N$ 41
2.1 Differenzierbarkeit und Ableitung 41
2.1.1 Partielle Ableitung 41
2.1.1.1 Definition 41
2.1.1.2 Geometrische Deutung 42
2.1.2 Partielle Ableitungen höherer Ordnung 43
2.1.2.1 Partielle Differenzierbarkeit und Stetigkeit 44
2.1.3 Totale Ableitung 45
2.1.3.1 Definition und Beispiel 45
2.1.3.2 Folgerungen aus der totalen Differenzierbarkeit 45
2.1.4 Differentiation von Vektorfunktionen 47
2.2 Die Kettenregel 48
2.2.1 Herleitung 48
2.2.2 Anwendung auf skalare Funktionen 49
2.2.2.1 Summenregel, Produktregel, Quotientenregel 49
2.2.2.2 Spezialfälle der Kettenregel 50
2.2.2.3 Richtungsableitung 50
2.2.2.4 Geometrische Deutung von $\nabla f = \operatorname{grad} f$ 51
2.3 Vektorfelder 51
2.3.1 Potentialfelder 51
2.3.1.1 Integrabilitätsbedingungen 52
2.3.1.2 Exakte Differentialgleichungen 53
2.3.2 (Differential-)Operationen mit Vektorfeldern 54
2.3.2.1 Das Vektorprodukt 54
2.3.2.2 Rotation und Divergenz 55
2.3.2.3 Der Nablaoperator 57
2.3.2.4 Rechenregeln 57
2.4 Taylorscher Lehrsatz und Anwendungen 58
2.4.1 Der Taylorsche Lehrsatz im $\mathbb{R}^N$ 58
2.4.1.1 Herleitung der Taylorformel 58

2.4.1.2 Spezialfälle . . . 59
2.4.2 Berechnung lokaler Extremstellen skalarer Funktionen . . . 59
2.4.2.1 Notwendige Bedingung . . . 59
2.4.2.2 Hinreichende Bedingung . . . 60
2.4.2.3 Ergänzende Betrachtungen . . . 60
2.4.2.4 Beispiele . . . 61
2.5 Implizite Funktionen . . . 62
2.5.1 Der Hauptsatz über implizite Funktionen . . . 62
2.5.1.1 Problemstellung . . . 62
2.5.1.2 Hauptsatz . . . 64
2.5.1.3 Spezialfall . . . 65
2.5.1.4 Korollar zum Hauptsatz 2.53 . . . 65
2.5.2 Extrema von Funktionen unter Nebenbedingungen . . . 66
2.5.2.1 Problemstellung . . . 66
2.5.2.2 Lagrangesche Multiplikatoren . . . 67
2.5.2.3 Lagrangefunktion . . . 68

3 Integralrechnung im $\mathbb{R}^N$ 71
3.1 Ergänzungen zum Riemann-Integral im $\mathbb{R}^1$ . . . 71
3.1.1 Uneigentliche Integrale . . . 71
3.1.1.1 Integrale über unbeschränkte Intervalle . . . 71
3.1.1.2 Integrale über unbeschränkte Funktionen . . . 72
3.1.1.3 Absolute Integrierbarkeit . . . 72
3.1.2 Parameterintegrale . . . 74
3.2 Mehrfache Riemann-Integrale . . . 76
3.2.1 Einführung . . . 76
3.2.1.1 Geometrische Motivation . . . 76
3.2.1.2 Bezeichnungen und Definitionen . . . 76
3.2.1.3 Eigenschaften . . . 77
3.2.2 Berechnung von mehrdimensionalen Riemann-Integralen . . . 77
3.2.2.1 Der Satz von Fubini . . . 77
3.2.3 Integration über beschränkte Mengen . . . 79
3.2.3.1 Definition und Eigenschaften . . . 79
3.2.3.2 Nullmengen . . . 79
3.2.3.3 Der Satz von Cavalieri . . . 80
3.2.3.4 Volumen von Rotationskörpern $\mathcal{K} \subset \mathbb{R}^3$ . . . 81
3.2.4 Integration über Normalbereiche . . . 81
3.3 Substitutionsregel . . . 83

3.3.1 Der Transformationssatz 83
3.3.2 Beispiele 84

Literaturverzeichnis

[1] Fischer, H. und H. Kaul: Mathematik für Physiker 1. Teubner 1990

[2] Forster, O.: Analysis 2. Teubner 1989

[3] Heuser, H.: Lehrbuch der Analysis II. Teubner 1990

[4] Jänich, K.: Analysis für Physiker und Ingenieure. Springer 1990

[5] Kerner, H. und W. von Wahl: Mathematik für Physiker. Springer 2006

[6] Königsberger, K.: Analysis 2. Springer 1993

[7] Zeidler, E.(Hrsg.): Teubner-Taschenbuch der Mathematik. Teubner 1996

Kapitel 1

Metrische Räume

1.1 Topologische Grundbegriffe

1.1.1 Der allgemeine Abstandsbegriff

Definition 1.1. Sei $X \neq \emptyset$. Eine Abbildung $d : X \times X \to \mathbb{R}$ heißt *Metrik* oder *Abstand*, falls für alle $x, y, z \in X$ gilt:

1. $d(x, y) \geq 0, d(x, y) = 0 \Leftrightarrow x = y$ (positive Definitheit)
2. $d(x, y) = d(y, x)$ (Symmetrie)
3. $d(x, y) \leq d(x, z) + d(z, y)$ (Dreiecksungleichung)

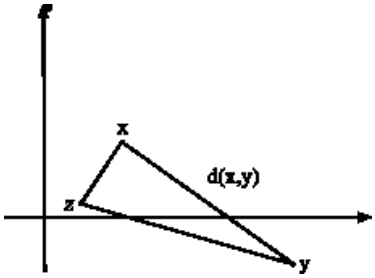
(X, d) heißt dann *metrischer Raum*.

Beispiel 1.2. a. $\mathbb{R}^1 : d(x, y) = |x - y|$

b. $\mathbb{R}^N : d(x, y) = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2} = \sqrt{\langle x - y, x - y\rangle} = \sqrt{(x - y)^T(x - y)}$

(*euklidische Norm* bzw. *Standardmetrik*)

$\mathbb{R}^2$:



c. Allgemein: Jeder normierte Vektorraum ist ein metrischer Raum mit $d(x, y) = \|x - y\|$

d. In jeder beliebigen Menge $X \neq \emptyset$ wird durch $d(x,y) = \begin{cases} 1 & \text{für } x \neq y \\ 0 & \text{für } x = y. \end{cases}$ eine Metrik definiert. Diese heißt *diskrete Metrik* auf X.

Bemerkung 1.3. Mit (X,d) ist auch (M,d) für jedes $M \subset X, M \neq \emptyset$ wieder ein metrischer Raum.

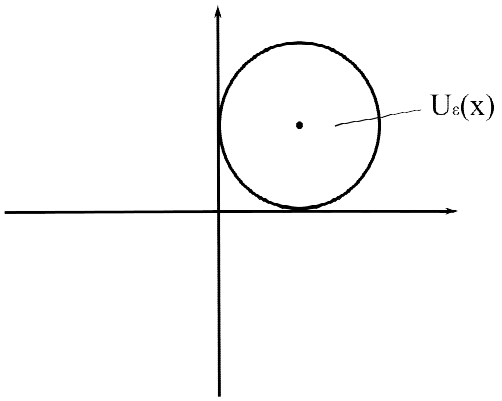
1.1.2 Konvergenz

Für den Konvergenzbegriff ist die Definition der *ε-Umgebung* $U_\varepsilon(x)$ für $x \in X$, X metrischer Raum, entscheidend:

$$U_\varepsilon(x) := \{y \in X \,|d(y,x) < \varepsilon\}\,.$$

Bemerkung 1.4. Ist X ein normierter Raum, so ist $U_\varepsilon(x) = \{y \in X \,|\|x-y\| \leq \varepsilon\}\,.$

$\mathbb{R}^2$ mit Standardmetrik:



Definition 1.5. Eine Folge $(x_n)_{n=1}^\infty \subset X$ heißt *konvergent* gegen $x \in X$ ($\lim\limits_{n\to\infty} = x$ bzw. $x_n \to x$ für $n \to \infty$), wenn für alle $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert.mit:

$$\underbrace{x_n \in U_\varepsilon(x)}_{\Leftrightarrow\ d(x_n,x)<\varepsilon} \text{ für alle } n \geq n_0.$$

Folgerung 1.6. $\lim\limits_{n\to\infty} x_n = x$ in $X \Leftrightarrow \lim\limits_{n\to\infty} d(x_n,x) = 0$ in $\mathbb{R}$.

Beispiel 1.7.

$$\mathbb{R}^N: \qquad d(x_n,x) = \sqrt{\sum_{i=1}^N (x_{ni}-x_i)^2} \to 0 \text{ für } n \to \infty$$

$$\Updownarrow$$

$$x_{ni} \to x_i \text{ für } n \to \infty,\ i = 1,\ldots,n;$$

$$x_n = (x_{n1}, x_{n2}, \ldots, x_{nN}),\ x = (x_1, x_2, \ldots, x_N).$$

Ergebnis. Konvergenz im $\mathbb{R}^N$ ist äquivalent zur koordinatenweisen Konvergenz.

Beispiel. $\lim\limits_{n\to\infty} \underbrace{\left(\sqrt[n]{n}, \frac{10^n}{n!}, \left(1+\frac{1}{n}\right)^n\right)}_{\in\mathbb{R}^3} = (1,0,\mathrm{e})$

Allgemein. Sei X ein *endlichdimensionaler* normierter Raum. Dann ist die Konvergenz in X bezüglich jeder Norm äquivalent zur koordinatenweise Konvergenz bezogen auf eine festgewählte Basis von X.
Man sagt: Alle Normen in X sind äquivalent (siehe auch Heuser II, Seite 18).

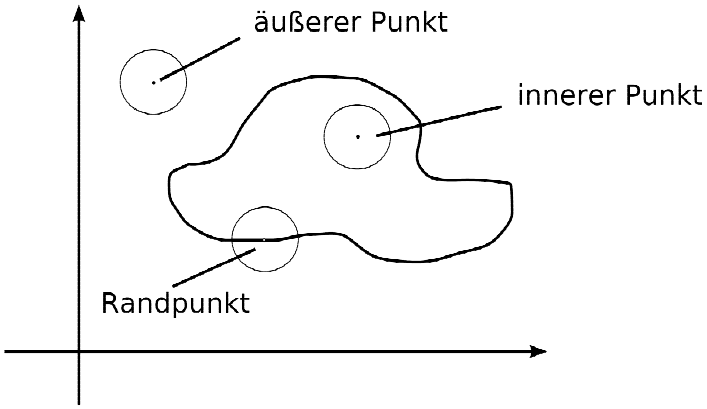
1.1.3 Klassifikation von Punkten

Definition 1.8. Sei X ein metrischer Raum, $M \subseteq X$. Ein Punkt $x \in X$ heißt

a. *innerer Punkt* von M, wenn $U_\varepsilon(x) \subset M$ für ein $\varepsilon > 0$.

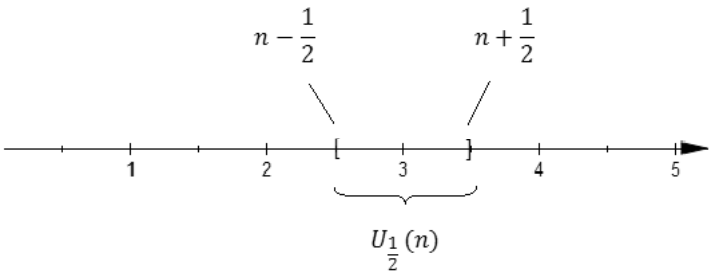
b. *äußerer Punkt* von M, wenn $U_\varepsilon(x) \subset X \setminus M$ für ein $\varepsilon > 0$.

c. *Randpunkt* von M, wenn $U_\varepsilon(x) \cap M \neq \emptyset$ und $U_\varepsilon(x) \cap (X \setminus M) \neq \emptyset$ für alle $\varepsilon > 0$.



d. *Häufungspunkt* von M, wenn $(U_\varepsilon(x) \cap M) \setminus \{x\} \neq \emptyset$ für alle $\varepsilon > 0$.

e. *isolierter Punkt* von M, wenn $U_\varepsilon(x) \cap M = \{x\}$ für ein $\varepsilon > 0$.
(Damit ist x ist kein Häufungspunkt von M.)
Beispiel. Für $\mathbb{N} \subset \mathbb{R}$ ist jedes $n \in \mathbb{N}$ isolierter Punkt von $\mathbb{N}$.



Bemerkung 1.9.

$$x \text{ ist Häufungspunkt von } M \;\Leftrightarrow\; \text{Es existiert } (x_n) \subset M \text{ mit } x = \lim_{n\to\infty} x_n,\ x_n \neq x.$$

Beweis. Siehe Übungsaufgabe.

Definition 1.10.

Das *Innere* int$M$ ist die Menge aller inneren Punkte von M.

Das *Äußere* ext$M$ ist die Menge aller äußeren Punkte von M.

Der *Rand* ∂M ist die Menge aller Randpunkte von M.

$$\Rightarrow \quad X = \text{int}M \cup \text{ext}M \cup \partial M \quad \text{(disjunkte Zerlegung von } X\text{)}$$

$\overline{M} := M \cup \partial M$ heißt *Abschluss* oder *abgeschlossene Hülle* von M.

Beispiel 1.11. i. $I = [a, b)\,, \text{int}I = (a, b)\,, \partial I = \{a, b\}\,, \overline{I} = [a, b]$

ii. $I = (a, \infty)\,, \text{int}I = I, \partial I = \{a\}\,, \overline{I} = [a, \infty)$

iii. int$U_\varepsilon(x) = U_\varepsilon(x)$, denn jeder Punkt $z \in U_\varepsilon(x)$ ist ein innerer Punkt, denn sei $d(y, z) < \underbrace{\varepsilon - d(x, z)}_{:=\varepsilon_1}$, so

$$d(y, x) \leq d(y, z) + d(z, x) < \varepsilon - d(x, z) + d(x, z) = \varepsilon \;\Rightarrow\; U_{\varepsilon_1}(z) \subseteq U_\varepsilon(x).$$

Bemerkung 1.12. Charakterisierung von $\overline{M}$:

$$x \in \overline{M} \;\Leftrightarrow\; \text{Es existiert } (x_n) \subset M \text{ mit } x = \lim_{n\to\infty} x_n.$$

Beweis. Siehe Übungsaufgabe.

1.1.4 Klassifikation von Mengen

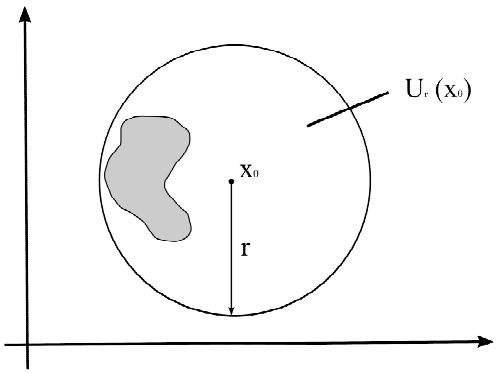
Definition 1.13. Die Menge M heißt

a. *offen*, wenn $M = \text{int}M$.

b. *abgeschlossen*, wenn $M = \overline{M}$.

c. *dicht* in X, wenn $\overline{M} = X$.

d. *beschränkt*, wenn $M \subseteq U_r(x_0)$ für ein $r > 0, x_0 \in X$.



Beispiel 1.14. i. Im $\mathbb{R}^1$ ist $\mathbb{Q}$ eine dichte Menge, denn $\overline{\mathbb{Q}} = \mathbb{R}$ (Jede reelle Zahl lässt sich als Grenzwert einer Folge rationaler Zahlen darstellen). Man sagt auch: $\mathbb{Q}$ liegt dicht in $\mathbb{R}$.

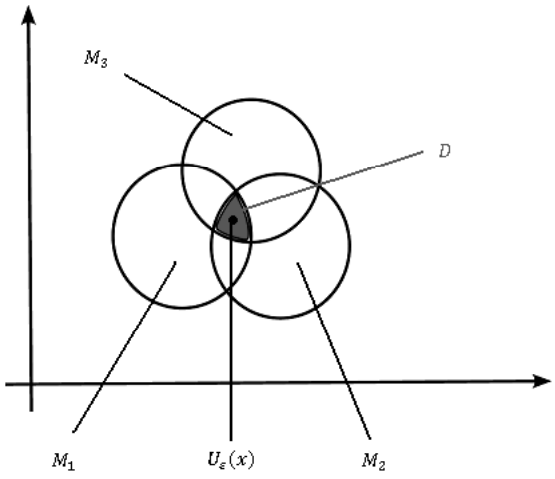
ii. $U_\varepsilon(x)$ ist stets offen in X, speziell (a, b) offen in $\mathbb{R}$.

1.1.4.1 Eigenschaften offener Mengen

1. $M \subseteq X$ ist offen $\Leftrightarrow$ Mit jedem $x \in M$ liegt auch $U_\varepsilon(x)$ in M für ein $\varepsilon > 0$.
2. $\emptyset, X$ sind offene Mengen.
3. Die Vereinigung beliebig vieler offener Mengen ist wieder offen (Beweis siehe Übungsaufgabe).
4. Der Durchschnitt endlich vieler offener Mengen ist wieder offen.

   Beweis. Sei $M_1, \ldots, M_k \subseteq X$ offen und $D := \bigcap_{j=1}^{k} M_j$. Für $x \in D$ muss $x \in M_j$ für jedes $j = 1, \ldots, k$ sein. Da M_j offen ist, existiert $\varepsilon_j > 0$ mit $U_{\varepsilon_j}(x) \subseteq M_j$. Setze

$$\varepsilon := \min_{j=1,\ldots,k} \varepsilon_j \;\Rightarrow\; U_\varepsilon(x) \subseteq U_{\varepsilon_j}(x) \subseteq M_j \text{ für jedes } j \;\Rightarrow\; U_\varepsilon(x) \subseteq \bigcap_{j=1}^{k} M_j = D.$$



Gegenbeispiel. Der Durchschnitt unendlich vieler offener Mengen muss nicht offen sein. $\mathbb{R}^1 : U_{\frac{1}{n}}(0) = \left(-\frac{1}{n}, \frac{1}{n}\right)$ ist offen für jedes $n \in \mathbb{N}$. $\bigcap_{n=1}^{\infty} U_{\frac{1}{n}}(0) = \bigcap_{n=1}^{\infty} \left(-\frac{1}{n}, \frac{1}{n}\right) = \{0\}$. Jedoch ist $\{0\}$ eine abgeschlossene Menge.

1.1.4.2 Eigenschaften abgeschlossener Mengen

1. $M \subseteq X$ ist abgeschlossen $\Leftrightarrow$ Jeder Häufungspunkt von M gehört zu M.

   Beweis. (i) ($\Rightarrow$) Sei $M = \overline{M}$ und x ein Häufungspunkt von M, d.h. es existiert $(x_n) \subset M$ mit $\lim_{n\to\infty} x_n = x$ und $x_n \neq x$ für alle $n \in \mathbb{N}$. Nach Charakterisierung von $\overline{M}$ muss $x \in \overline{M}$ gelten. Mit Vorraussetzung ist $x \in M$.

(ii) ($\Leftarrow$) Wir zeigen $\overline{M} \subseteq M$. Sei $x \in \overline{M}$. Dann gibt es $(x_n) \subset M$ mit $\lim\limits_{n\to\infty} x_n = x$. Wäre $x \in X \setminus M$, so folgt daraus $x_n \neq x$ für alle $n \in \mathbb{N}$, also ist x Häufungspunkt von M. Nach Vorraussetzung muss $x \in M$ sein $\Rightarrow$ Widerspruch. Demzufolge ist $x \in M$.

2. M ist abgeschlossen $\Leftrightarrow$ $X \setminus M$ offen ist, denn
„$X \setminus M$ offen“ bedeutet: Mit $y \in X \setminus M$ gibt es $U_\varepsilon(y) \subset X \setminus M$ für ein $\varepsilon > 0$. D.h., kein $y \in X \setminus M$ kann Häufungspunkt von M sein. Somit müssen alle Häufungspunkte von M in M liegen. Mit 1. folgt, dass M abgeschlossen ist.

3. $\emptyset, X$ sind abgeschlossen.

4. Der Durchschnitt beliebig vieler abgeschlossener Mengen ist wieder abgeschlossen. Die Vereinigung endlich vieler abgeschlossener Mengen ist wieder abgeschlossen.

1.2 Vollständigkeit und Kompaktheit

1.2.1 Vollständige metrische Räume

Definition 1.15. Sei (X, d) ein metrischer Raum. Dann heißt $(x_n) \subset X$ *Cauchyfolge* oder *Fundamentalfolge*, falls für jedes $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert mit $d(x_n, x_m) < \varepsilon$ für alle $n, m \geq n_0$.

Folgerung 1.16. Wie im $\mathbb{R}^1$ erhält man

1. Jede konvergente Folge aus X ist Cauchyfolge.

2. Jede Cauchyfolge ist beschränkt.

Definition 1.17. a. X heißt *vollständig*, wenn jede Cauchy-Folge aus X gegen einen Grenzwert in X konvergiert.

b. Ein vollständiger normierter Raum heißt auch *Banachraum* (Banach 1892 - 1945).

Beispiel 1.18. i. $\mathbb{R}^1$ ist vollständig.

ii. $\mathbb{R}^N$ ist vollständig, denn sei (x_n) eine Cauchyfolge in $\mathbb{R}^N$, so gilt

$$\|x_n - x_m\| = \underbrace{\sqrt{\sum_{i=1}^{N} (x_{ni} - x_{mi})^2}}_{\geq |x_{ni} - x_{mi}| \text{ für jedes } i=1,\ldots,N} < \varepsilon \text{ für } n, m \geq n_0(\varepsilon)$$

$\Rightarrow$ $(x_{ni})_{n=1}^{\infty}$ ist Cauchyfolge in $\mathbb{R}$. Nach Cauchyschem Konvergenzkriterium folgt $x_{ni} \to x_i \in \mathbb{R}$ für $n \to \infty$. Damit ergibt sich die koordinatenweise Konvergenz von (x_n) gegen $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$, d.h. $\lim\limits_{n\to\infty} x_n = x$ in $\mathbb{R}^N$ bezüglich der euklidischen Norm.

Folgerung 1.19. $\mathbb{R}^N$ und $\mathbb{C}^N$ sind vollständig, d.h. sie sind Banachräume bezüglich jeder Norm.

Satz 1.20. Der unendlichdimensionale normierte Raum $C([a,b])$ ist vollständig bezüglich

$$\|f\| = \max_{x\in[a,b]} |f(x)| \quad (\textit{Maximum}\text{- bzw. }\textit{Supremumnorm} \text{ auf } C([a,b])).$$

Beweis. (i). Sei $(f_n) \subset C([a,b])$ eine Cauchyfolge bezüglich $\|.\|$, so gilt

$$\|f_n - f_m\| = \max_{x\in[a,b]} \underbrace{|f_n(x) - f_m(x)|}_{\geq |f_n(x)-f_m(x)| \text{ für jedes } x\in[a,b]} < \varepsilon \text{ für alle } n,m \geq n_0(\varepsilon) \tag{1.2.1.1}$$

Für festes x ist $(f_n(x))_{n=1}^{\infty}$ eine Cauchyfolge in $\mathbb{R}$. Nach Cauchykriterium konvergiert $(f_n(x))$. Wir schreiben $\lim\limits_{n\to\infty} f_n(x) = f(x)$. Somit ist $f : [a,b] \to \mathbb{R}$.

(ii). Wir müssen zeigen, dass f stetig auf $[a,b]$ ist. Aus (1.2.1.1) folgt mit $n \to \infty$:

$$|f(x) - f_m(x)| \leq \varepsilon \text{ für alle } m \geq n_0 \text{ und alle } x \in [a,b]. \tag{1.2.1.2}$$

Nachweis der Stetigkeit von f mittels ε-δ-Kriterium unter Benutzung des "3ε-Tricks": Sei $\varepsilon > 0, x, y \in [a,b]$, dann

$$\begin{aligned} |f(x) - f(y)| &\leq \big|(f(x) - f_{n_0}(x)) + (f_{n_0}(x) - f_{n_0}(y)) + (f_{n_0}(y) - f(y))\big| \\ &\leq \underbrace{\big|f(x) - f_{n_0}(x)\big|}_{\leq\varepsilon} + \big|f_{n_0}(x) - f_{n_0}(y)\big| + \underbrace{\big|f_{n_0}(y) - f(y)\big|}_{\leq\varepsilon}. \end{aligned}$$

Wegen Stetigkeit von f_{n_0} gilt $\big|f_{n_0}(x) - f_{n_0}(y)\big| < \varepsilon$ für alle y mit $|x-y| < \delta$,
$\delta = \delta(\varepsilon, n_0(\varepsilon)) =: \tilde{\delta}(\varepsilon))$

$$\Rightarrow \quad |f(x) - f(y)| < 3\varepsilon \quad \text{für dieselben } x, y \quad \Rightarrow \quad f \in C([a,b]).$$

(iii). Aus (1.2.1.2) folgt

$$\|f - f_m\| \leq \varepsilon \text{ für alle } m \geq n_0 \quad \Leftrightarrow \quad \lim_{n\to\infty} f_m = f \text{ in } C([a,b]).$$

Bemerkung 1.21.

i. Bezüglich der durch das Skalarprodukt $\langle f, g\rangle := \int\limits_a^b f(x)g(x)\mathrm{d}x$ bestimmten Norm ist $C([a,b])$ nicht vollständig (s. Vorlesung Analysis III, IV).

ii. Konvergenz bezüglich der Maximumnorm heißt auch *gleichmäßige Konvergenz* der Funktionenfolge (f_n) im Gegensatz zur *punktweisen Konvergenz* der Funktionenfolge (f_n), d.h. von $(f_n(x)) \subset \mathbb{R}$ für jedes $x \in [a,b]$.

1.2.2 Kompakte Mengen

Definition 1.22. Eine Menge $M \subseteq X$ heißt *(folgen-)kompakt*, wenn jede Folge aus M eine konvergente Teilfolge mit Grenzwert in M besitzt.

Beispiel 1.23. $[a, b]$ mit $a < b$ ist eine kompakte Menge im $\mathbb{R}^1$, da nach Satz von Bolzano-Weierstraß jede unendliche Teilmenge von $[a, b]$ einen Häufungspunkt hat.

Satz 1.24. Jede kompakte Menge $M \subseteq X$ ist beschränkt und abgeschlossen bezüglich der Metrik d von X.

Beweis. (i). Wir zeigen $M = \overline{M}$. 1.Fall: M endlich $\Rightarrow$ M hat keinen Häufungspunkt.
2. Fall: M unendlich. Sei x Häufungspunkt von M, d.h. $x = \lim\limits_{n\to\infty} x_n, x_n \in M, x_n \neq x \forall n \in \mathbb{N}$. Wegen Kompaktheit von M besitzt (x_n) eine Teilfolge $(x_{n_k}) \subseteq (x_n)$ mit $\lim\limits_{k\to\infty} x_{n_k} = \tilde{x}, \tilde{x} \in M \Rightarrow x = \tilde{x}$ wegen Eindeutigkeit des Grenzwertes und Teilfolgenprinzip. Somit folgt $x \in M$.

(ii). Annahme: Sei M nicht beschränkt, d.h. zu jedem $r > 0$ und $y \in X$ gibt es $x \in M$ mit $x \notin U_r(y)$, d.h. $d(x, y) \geq r$. Hier ist $x = x(y, r)$. Hält man $y \in X$ fest und wählt $r = n$, so ergibt sich $x_n := x(y, n) \in M$ und $d(x_n, y) \geq n$. Jedoch existiert $(x_{n_k}) \subseteq (x_n)$ mit $\lim\limits_{k\to\infty} x_{n_k} = \tilde{x} \in M$. Daraus folgt $d(x_{n_k}, \tilde{x}) \leq 1$ für $k \geq k_0(1)$.

$$n_k \leq d(x_{n_k}, y) \leq d(x_{n_k}, \tilde{x}) + d(\tilde{x}, y) \leq 1 + d(\tilde{x}, y) \text{ für alle } k \geq k_0(1).$$

Da (n_k) unbeschränkt ist, ergibt sich der Widerspruch. M muss somit beschränkt sein.

Satz 1.25. (*Satz von Bolzano-Weierstraß* im $\mathbb{R}^N$)
Jede beschränkte Menge M des $\mathbb{R}^N$ besitzt einen Häufungspunkt.

Beweis. $M \subset \mathbb{R}^N$ ist beschränkt $\Leftrightarrow$ $\|x\| = \sqrt{\sum\limits_{i=1}^{N} x_i^2} \leq R$ für alle $x \in M$ mit einem $R > 0$. Sei $(x_n) \subset M$, so ist $|x_{n,1}| \leq R$, d.h. $(x_{n,1})$ ist beschränkte Menge im $\mathbb{R}^1$. $\Rightarrow$ Es existiert $(x_{n_1,1})_{n_1=1}^{\infty} \subseteq (x_{n,1})$ mit $\lim\limits_{n\to\infty} x_{n_1,1} = \tilde{x}_1$.
$(x_{n_1,2}) \subseteq (x_{n,2})$ ist beschränkt im $\mathbb{R}^1$. $\Rightarrow$ Es existiert $(x_{n_2,2})_{n_2=1}^{\infty} \subseteq (x_{n_1,2})$ mit $\lim\limits_{n_2\to\infty} x_{n_2,2} = \tilde{x}_2$ und $\lim\limits_{n_2\to\infty} x_{n_2,1} = \tilde{x}_1$. Nach N Schritten erhält man

$$(x_{n_N})_{N=1}^{\infty} \subseteq (x_n) \text{ mit } \lim_{n_N\to\infty} x_{n_N,i} = \tilde{x}_i \text{ für alle } i = 1, \ldots, n \Leftrightarrow \left\|x_{n_N} - \tilde{x}\right\| \to 0 \text{ für } n_N \to \infty$$

d.h. $\tilde{x} := (\tilde{x}_1, \ldots, \tilde{x}_N)$ ist Häufungspunkt von M.

Ergebnis. Im $\mathbb{R}^N$ und $\mathbb{C}^N$ ist die Kompaktheit von M äquivalent zur Beschränktheit und Abgeschlossenheit von M.

Bemerkung 1.26. In allgemeinen metrischen Räumen folgt aus Beschränktheit und Abgeschlossenheit *nicht* die Kompaktheit der Menge.

Gegenbeispiel. In einem *unendlichdimensionalen* euklidischen Vektorraum ist

$$\overline{U_1(0)} = \left\{ x \in V \,\middle|\, \|x\| = \sqrt{\langle x, x\rangle} \leq 1 \right\}.$$

Diese Menge ist zwar beschränkt und abgeschlossen, aber nicht kompakt, denn V besitzt ein Orthonormalsystem $(x_n)_{n=1}^{\infty} \subset V$, d.h. $\langle x_n, x_m\rangle = \delta_{nm} = \begin{cases} 1 & \text{für } m = n \\ 0 & \text{für } m \neq n. \end{cases}$ *(Kroneckersymbol).*(Kronecker 1823 - 1891)

$$\|x_n - x_m\|^2 = \langle x_n - x_m, x_n - x_m\rangle = \underbrace{\langle x_n, x_n\rangle}_{=1} - \underbrace{\langle x_n, x_m\rangle}_{=0} - \underbrace{\langle x_m, x_n\rangle}_{=0} + \underbrace{\langle x_m, x_m\rangle}_{=1} = 2$$

für alle $n, m \in \mathbb{N}, n \neq m$. Das heißt, keine Teilfolge von (x_n) kann konvergieren.

1.3 Funktionenfolgen und Funktionenreihen

1.3.1 Problemstellung

Wir betrachten die Folge $(f_n)_{n=1}^{\infty}$ von reellen Funktionen $f_n : I \to \mathbb{R}, I \subseteq \mathbb{R}$ (*Funktionenfolge*). Dann ist für jedes $x \in I$: $(f_n(x))_{n=1}^{\infty}$ eine reelle Zahlenfolge. Sei $f(x) := \lim\limits_{n\to\infty} f_n(x)$. $f(x)$ heißt *Grenzfunktion* von (f_n).
Fragestellung nach der Vertauschbarkeit von Grenzprozessen, d.h. wann gilt

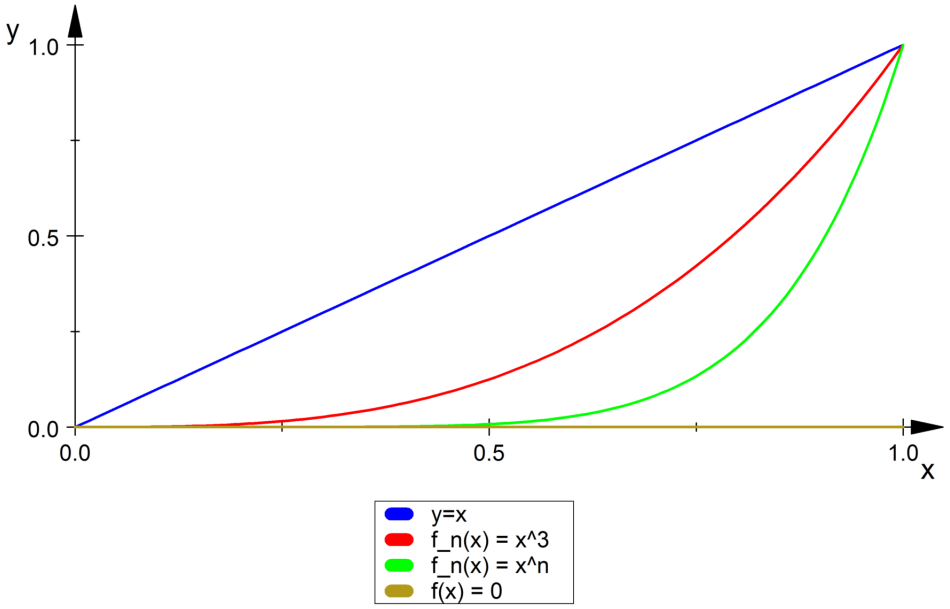
$$\frac{\mathrm{d}}{\mathrm{d}x} \lim_{n\to\infty} f_n(x) = \frac{\mathrm{d}}{\mathrm{d}x} f(x) = \lim_{n\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} f_n(x)\,,$$

$$\int_a^b \lim_{n\to\infty} f_n(x)\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\mathrm{d}x\ ?$$

Beispiel 1.27. i.

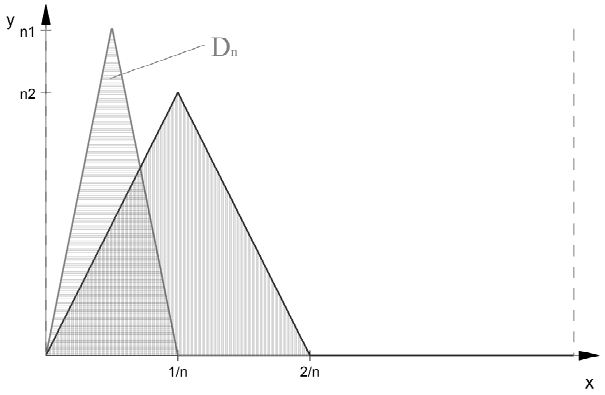
$$f_n(x) = x^n, x \in [0,1]\,, f_n \in C^{\infty}([0,1])$$

$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x = 1\,. \end{cases}$$



Beobachtung. Die Grenzfunktion einer Folge stetiger Funktionen muss nicht stetig sein.

ii. Folgende Grafik einer Funktionenfolge sei gegeben:



$$f_n(x) = \begin{cases} n^2 x & \text{für } 0 \le x < \frac{1}{n} \\ 2n - n^2 x & \text{für } \frac{1}{n} \le x < \frac{2}{n} \\ 0 & \text{für } \frac{2}{n} \le x \le 1 \end{cases}.$$

Dann gilt $\lim\limits_{n\to\infty} f_n(x) = 0$ für alle $x \in [0,1]$. Klar für $x = 0$.

$$\text{Sei } x > 0: \text{ Dann ist } n > \frac{2}{x} \Leftrightarrow x > \frac{2}{n} \Rightarrow f_n(x) = 0.$$

$$\int_0^1 f_n(x)\mathrm{d}x = |D_n| = \tfrac{1}{2} \cdot \tfrac{2}{n} \cdot n = 1 \to 1 \text{ für } n \to \infty, \textit{ aber } \int_0^1 \lim_{n\to\infty} f_n(x) = \int_0^1 0 \,\mathrm{d}x = 0.$$

1.3.2 Gleichmäßige Konvergenz von Funktionenfolgen

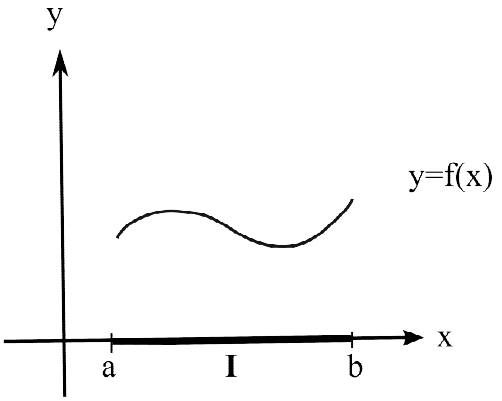
Definition 1.28. Sei $f_n : I \to \mathbb{R}, I \subseteq \mathbb{R}, n \in \mathbb{N}$.

a. (f_n) heißt *punktweise konvergent* gegen $f(x)$, falls $\lim\limits_{n\to\infty} f_n(x) = f(x)$ für jedes $x \in I$, d.h. $\forall x \in I, \forall \varepsilon > 0$ existiert $n_0 = n_0(\varepsilon, x) : |f_n(x) - f(x)| < \varepsilon$ für $n \geq n_0$.

b. (f_n) heißt *gleichmäßig konvergent* gegen f, falls $\|f_n - f\| = \sup\limits_{x\in I} |f_n(x) - f(x)| \to 0$ für $n \to \infty$.

Bemerkung 1.29. i. Die gleichmäßige Konvergenz ist eine „stärkere“ Eigenschaft als die punktweise Konvergenz, denn die gleichmäßige Konvergenz von (f_n) impliziert die punktweise Konvergenz von (f_n).
Aber: Die Umkehrung gilt nicht (siehe Beispiele 1 und 2).

ii. Die gleichmäßige Konvergenz besagt, dass das n_0 in der ε-n_0-Abschätzung für $\lim\limits_{n\to\infty} f_n(x) = f(x)$ *unabhängig von* $x \in I$ gewählt werden kann.



Gleichmäßige Konvergenz von f_n gegen f heißt:
$f_n(x)$ liegt im ε-*Schlauch* um f für alle $x \in I$ und $n \geq \tilde{n}_0(\varepsilon)$
(bei beliebigem ε).

iii. Für $f_n \in C(I)$ ist die gleichmäßige Konvergenz dasselbe wie die Konvergenz im Banachraum $C(I)$.

Folgerung 1.30. 1. Wegen Vollständigkeit von $C(I)$ gilt das *Cauchy-Kriterium für gleichmäßige Konvergenz*: $(f_n) \subset C(I)$ konvergiert gleichmäßig auf I genau dann, wenn zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ existiert mit

$$|f_n(x) - f_m(x)| < \varepsilon \quad \text{für alle } n, m \geq n_0 \text{ und alle } x \in I.$$

2. **Satz.** *(Satz über die Stetigkeit der Grenzfunktion)*
Sei $(f_n) \subset C(I)$ gleichmäßig konvergent gegen f, so ist $f \in C(I)$.
(Siehe “3ε-Trick“ im Beweis der Vollständigkeit von $C(I)$.)

3. **Satz.** *(Satz über die Vertauschbarkeit von Integration und Grenzwert)*
Sei (f_n) gleichmäßig konvergent gegen f und $F_n(x) := \int\limits_a^x f_n(t)\mathrm{d}t$, $F(x) := \int\limits_a^x f(t)\mathrm{d}t$ für $x \in [a, b]$. Dann gilt: (F_n) konvergiert gleichmäßig gegen F und

$$\int_a^b \lim_{n\to\infty} f_n(t)\mathrm{d}t = \lim_{n\to\infty} \int_a^b f_n(t)\mathrm{d}t \quad (\Leftrightarrow\ F_n(b) \to F(b)).$$

Beweis. Nach 2. ist $f \in C(I)$ und

$$|f_n(t) - f(t)| \leq \sup_{t\in[a,x]} |f_n(t) - f(t)| = \|f_n - f\|_{C([a,x])} < \varepsilon$$

für alle $n \geq \tilde{n}_0(\varepsilon)$ und alle $t \in [a, x]$.

$$\Rightarrow \underbrace{\left| \int_a^x (f_n(t) - f(t))\,\mathrm{d}t \right|}_{=|F_n(x)-F(x)|} \leq \int_a^x |f_n(t) - f(t)|\,\mathrm{d}t \leq \int_a^x \varepsilon\,\mathrm{d}t = \varepsilon(x-a) \leq \varepsilon(b-a)$$

für *jedes* $x \in (a, b]$, d.h. $|F_n(x) - F(x)| \leq \varepsilon(b-a)$ für alle $n \geq \tilde{n}_0(\varepsilon)$ und alle $x \in [a, b]$, also gleichmäßige Konvergenz von (F_n).

4. **Satz.** *(Satz über die Vertauschbarkeit von Differentiation und Grenzwertbildung)*
Sei $(f_n)) \subset C^1(I)$ und (f_n') gleichmäßig konvergent gegen g und $\lim_{n\to\infty} f_n(x_0) = A$ für ein $x_0 \in [a, b]$. Dann konvergiert (f_n) gleichmäßig gegen f, wobei $f \in C^1(I)$, $f' = g$, $A = f(x_0)$.

   Beweis. Beweis mittels Hauptsatz der Differential- und Integralrechnung.

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f_n'(t)\,\mathrm{d}t\,, \quad \text{Setze} \quad f(x) := A + \int_{x_0}^x g(t)\,\mathrm{d}t.$$

   Dann gilt: $f(x_0) = A$, $f' = g \in C(I)$ wegen 2. Außerdem folgt aus 3.: $\lim_{n\to\infty} f_n = f$ gleichmäßig auf I, weil (f_n') gleichmäßig gegen g konvergiert.

Bemerkung 1.31. In den Beispielen 1.27.i. und ii. konvergieren die Funktionenfolgen (f_n) zwar punktweise auf dem Intervall $[0, 1]$, aber *nicht gleichmäßig*. In Beispiel i. ist die Eigenschaft aus 2. und in Beispiel ii. die Eigenschaft aus 3. verletzt.
Auf $[0, r]$ mit $r < 1$ ist die Konvergenz von $(x^n)_{n=1}^\infty$ gegen 0 gleichmäßig, denn

$$|x^n - 0| = x^n \leq r^n < \varepsilon \text{ für jedes } x \in [0, r] \text{ und } \ln(r^n) = n \ln r < \ln \varepsilon \Leftrightarrow n > \frac{\ln \varepsilon}{\ln r}.$$

Wähle $n_0(\varepsilon) := \left[\frac{\ln \varepsilon}{\ln r}\right] + 1$, dann gilt: $\|x^n - 0\|_{C([0,r])} < \varepsilon$ für alle $n \geq n_0(\varepsilon)$.

1.3.3 Gleichmäßige Konvergenz von Funktionenreihen

Definition 1.32. Die *Funktionenreihe* $\sum_{n=0}^\infty f_n$ heißt *gleichmäßig konvergent* gegen f, wenn die Funktionenfolge $(s_k)_{k=0}^\infty$, $s_k(x) := \sum_{n=0}^k f_n(x)$ gleichmäßig gegen f konvergiert.
$(f_n : I \to \mathbb{R}, I \subseteq \mathbb{R})$

Bemerkung 1.33. i. Mit dieser Definition übertragen sich die Folgerungen 1.29.1. - 4. zur gleichmäßigen Konvergenz auf Funktionenreihen bei denselben Vorraussetzungen an f_n.

ii. Sätze gelten entsprechend auch für komplexwertige Funktionen $f_n : I \to \mathbb{C}$.

Satz 1.34. *(Majorantenkriterium für gleichmäßige Konvergenz von Funktionenreihen)* Wenn $|f_n(x)| \leq c_n$ für alle $n \in \mathbb{N}_0$ und alle $x \in I$ gilt sowie $\sum_{n=0}^{\infty} c_n$ konvergiert, dann ist $\sum_{n=0}^{\infty} f_n$ gleichmäßig konvergent auf I.

Beweis. Beweis über Cauchy-Kriterium für gleichmäßige Konvergenz von (s_k).

$$|s_m(x) - s_k(x)| = \left| \sum_{n=k+1}^{m} f_n(x) \right| \leq \sum_{n=k+1}^{m} \underbrace{|f_n(x)|}_{\leq c_n} \leq \sum_{n=k+1}^{m} c_n \leq \varepsilon$$

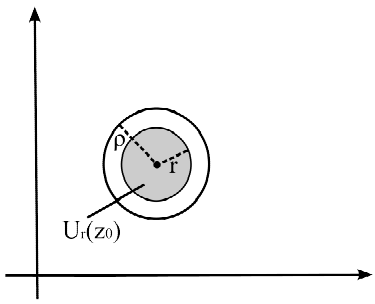
für $m > k \geq n_0(\varepsilon)$, weil $\sum_{n=0}^{\infty} c_n$ konvergiert. Diese Ungleichung gilt für alle $x \in I$, daraus folgt die gleichmäßige Konvergenz von (s_k).

Beispiel 1.35. Sei die Reihe $\sum_{n=1}^{\infty} \underbrace{\frac{1}{n^2} \sin(nx)}_{f_n(x)}$ gegeben. Wegen $|f_n(x)| = \left|\frac{1}{n^2} \sin(nx)\right| \leq \frac{1}{n^2}$ für alle $n \in \mathbb{N}$, und weil $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergiert, liefert das Majorantenkriterium die gleichmäßige Konvergenz dieser Reihe für $x \in \mathbb{R}$. Damit folgt: $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^2} \sin(nx)$ ist stetig auf $\mathbb{R}$.

1.3.4 Anwendung des neuen Konvergenzbegriffes

1.3.4.1 Potenzreihen

Hat die Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ den Konvergenzradius $\varrho > 0$, dann konvergiert $\sum_{n=0}^{\infty} a_n(z - z_0)$ für $|z - z_0| < \varrho$ absolut. Aus dem Majorantenkriterium folgt die gleichmäßige Konvergenz von $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ für $z \in \mathbb{C}$ mit $|z - z_0| \leq r$ wobei $0 < r < \varrho$ ist, denn wählt man ein $\tilde{z} \in \mathbb{C}$ mit $|\tilde{z} - z_0| = r$, so konvergiert $\sum_{n=0}^{\infty} \underbrace{|a_n(\tilde{z} - z_0)^n|}_{=|a_n| r^n}$ und es gilt



$$|a_n(z - z_0)^n| = |a_n| \, |z - z_0|^n \leq |a_n| \, r^n$$

für alle $n \in \mathbb{N}_0, |z - z_0| \leq r$.

1.3.4.2 Fourierreihen

Wir betrachten die *trigonometrische Reihe*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \qquad a_n, b_n \in \mathbb{R}, x \in [-\pi, \pi] \tag{1.3.4.3}$$

Die Reihe habe als Summe die Funktion $f(x)$.

Satz 1.36. Konvergiert die Reihe (1.3.4.3) gleichmäßig auf $[-\pi, \pi]$ gegen f, so gilt

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\, \mathrm{d}x\,, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx)\, \mathrm{d}x \qquad \text{für } n \in \mathbb{N}_0. \tag{1.3.4.4}$$

Beweis. Sei $m \in \mathbb{N}_0$ fest. Dann konvergiert

$$\frac{a_0}{2} \cos(mx) + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \cos(mx)$$

gleichmäßig auf $[-\pi, \pi]$ gegen $f(x)\cos(mx)$.

$$m = 0: \int_{-\pi}^{\pi} f(x)\cos(0x)\, \mathrm{d}x = \frac{a_0}{2} \underbrace{\int_{-\pi}^{\pi} \mathrm{d}x}_{=2\pi} + \sum_{n=1}^{\infty} \Big(a_n \underbrace{\int_{-\pi}^{\pi} \cos(nx)\, \mathrm{d}x}_{=0} + b_n \underbrace{\int_{-\pi}^{\pi} \sin(nx)\, \mathrm{d}x}_{=0} \Big)$$

$$\Rightarrow \quad a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, \mathrm{d}x$$

$$m \in \mathbb{N}: \int_{-\pi}^{\pi} f(x)\cos(mx)\, \mathrm{d}x = \frac{a_0}{2} \underbrace{\int_{-\pi}^{\pi} \cos(mx)\, \mathrm{d}x}_{=\langle \cos(mx), 1\rangle = 0}$$

$$+ \sum_{n=1}^{\infty} \Big(a_n \underbrace{\int_{-\pi}^{\pi} \cos(nx)\cos(mx)\, \mathrm{d}x}_{=\langle \cos(nx), \cos(mx)\rangle = \pi\delta_{mn}} + b_n \underbrace{\int_{-\pi}^{\pi} \sin(nx)\cos(mx)\, \mathrm{d}x}_{\langle \sin(nx), \cos(mx)\rangle = 0} \Big).$$

Da $\{\cos(nx), \sin(nx) \,|n \in \mathbb{N}\} \cup \{1\}$ in $C([-\pi, \pi])$ ein Orthogonalsystem bezüglich des Skalarproduktes $\langle f, g\rangle := \int_{-\pi}^{\pi} fg\, \mathrm{d}x$ bildet, gilt:

$$\langle \cos(mx), \sin(nx)\rangle = \langle \cos(mx), \cos(nx)\rangle = \langle \sin(mx), \sin(nx)\rangle = 0 \quad \text{für } m \neq n$$
$$\langle \cos(mx), \sin(mx)\rangle = 0, \langle \cos(mx), \cos(mx)\rangle = \langle \sin(mx), \sin(mx)\rangle = \pi \quad \text{für } m \in \mathbb{N}.$$

Damit ergibt sich: $\int_{-\pi}^{\pi} f(x)\cos(mx)\, \mathrm{d}x = \pi a_m$.

Durch Multiplikation mit $\sin(mx)$ folgt analog die Formel für b_n.

Bemerkung 1.37. i. Die a_n, b_n heißen *Fourierkoeffizienten* von f. Die Reihe (1.3.4.3) heißt *Fourierreihe* zu f. (Fourier 1768 - 1830)

ii. Konvergiert die Fourierreihe (1.3.4.3) für alle $x \in [-\pi, \pi]$, so auch für alle $x \in \mathbb{R}$, da $f(x+2k\pi) = f(x)$ für $k \in \mathbb{Z}$ wegen der Periodizität von $\cos(nx), \sin(nx)$ (Die Grenzfunktion f ist wieder 2π-periodisch).

iii. Sei $g \in R([-a, a])$.

Ist g ungerade, d.h. $g(x) = -g(-x)$ für alle $x \in [-a, a]$, so ist $\int_{-a}^{a} g(x)\,\mathrm{d}x = 0$.

Ist g gerade, d.h. $g(-x) = g(x)$ für alle $x \in [-a, a]$, so ist $\int_{-a}^{a} g(x)\,\mathrm{d}x = 2\int_{0}^{a} g(x)\,\mathrm{d}x$.

Anwendung auf Fourierkoeffizienten (1.3.4.4):

(i) f ist gerade auf $[-\pi, \pi] \Rightarrow f(x)\sin(nx)$ ist ungerade $\Rightarrow b_n = 0$. $f(x)\cos(nx)$ ist gerade $\Rightarrow a_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\cos(nx)\,\mathrm{d}x$. Die Fourierreihe zu f ist somit eine reine Kosinusreihe.

(ii) f ist ungerade auf $[-\pi, \pi] \Rightarrow b_n = \frac{2}{\pi}\int_{0}^{\pi} f(x)\sin(nx)\,\mathrm{d}x$, $a_n = 0$. Die Fourierreihe zu f ist somit eine reine Sinusreihe.

Sei jetzt umgekehrt $f \in R([-\pi, \pi])$[I] gegeben. Dann kann man die Fourierkoeffizienten von f gemäß der Formeln (1.3.4.4) ausrechnen.
Fragestellung: Wann konvergiert die Fourierreihe (1.3.4.3) und wann ist ihre Summe gleich $f(x)$?
Ergebnisse.

(1). Konvergiert $\sum(|a_n| + |b_n|)$, so konvergiert die Fourierreihe (1.3.4.3) gleichmäßig auf $[-\pi, \pi]$ wegen des Majorantenkriteriums, da

$$|a_n\cos(nx) + b_n\sin(nx)| \le |a_n|\underbrace{|\cos(nx)|}_{\le 1} + |b_n|\underbrace{|\sin(nx)|}_{\le 1} \le |a_n| + |b_n|$$

(2). Ist $f \in C^2([-\pi, \pi])$ mit $f(-\pi) = f(\pi)$, $f'(-\pi) = f'(\pi)$, so konvergiert die Fourierreihe zu f

[I] $R(I)$ ist der Raum der Riemann-integrierbaren Funktionen auf dem Intervall I

gleichmäßig auf $[-\pi, \pi]$, denn

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\ \underbrace{\cos(nx)}_{=\left(\frac{1}{n}\sin(nx)\right)'}\ \mathrm{d}x = \frac{1}{\pi}\left(\underbrace{f(x)\frac{1}{n}\sin(nx)\Big|_{-\pi}^{\pi}}_{=0} - \frac{1}{n}\int_{-\pi}^{\pi} f'(x)\ \underbrace{\sin(nx)}_{=\left(\frac{1}{n}\cos(nx)\right)'}\ \mathrm{d}x\right)$$

$$= -\frac{1}{n\pi}\left(f'(x)\left(-\frac{1}{n}\cos(nx)\right)\Big|_{-\pi}^{\pi} + \frac{1}{n}\int_{-\pi}^{\pi} f''(x)\cos(nx)\ \mathrm{d}x\right)$$

$$= \frac{1}{n^2\pi}\left(f'(\pi)\cos(n\pi) - \underbrace{f'(-\pi)}_{=f'(\pi)}\underbrace{\cos(-n\pi)}_{=\cos(n\pi)} - \int_{-\pi}^{\pi} f''(x)\cos(nx)\ \mathrm{d}x\right)$$

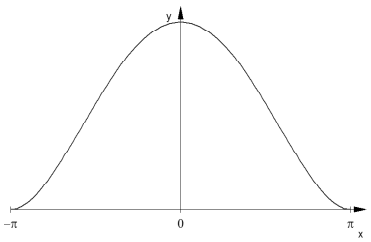
$$|a_n| = \frac{1}{n^2\pi}\left|\int_{-\pi}^{\pi} f''(x)\cos(nx)\ \mathrm{d}x\right| \le \frac{1}{n^2\pi}\int_{-\pi}^{\pi}\underbrace{|f''(x)|}_{\le\|f''\|}\underbrace{|\cos(nx)|}_{\le 1}\ \mathrm{d}x$$

$$\le \frac{2\pi}{n^2\pi}\|f''\|_{C([-\pi,\pi])}$$

$\Rightarrow \sum |a_n|$ konvergiert nach Majorantenkriterium. Analog konvergiert $\sum |b_n|$. Somit ist nach (1) die Fourierreihe zu f gleichmäßig konvergent.

Beispiel 1.38. Ein solches f ist zum Beispiel

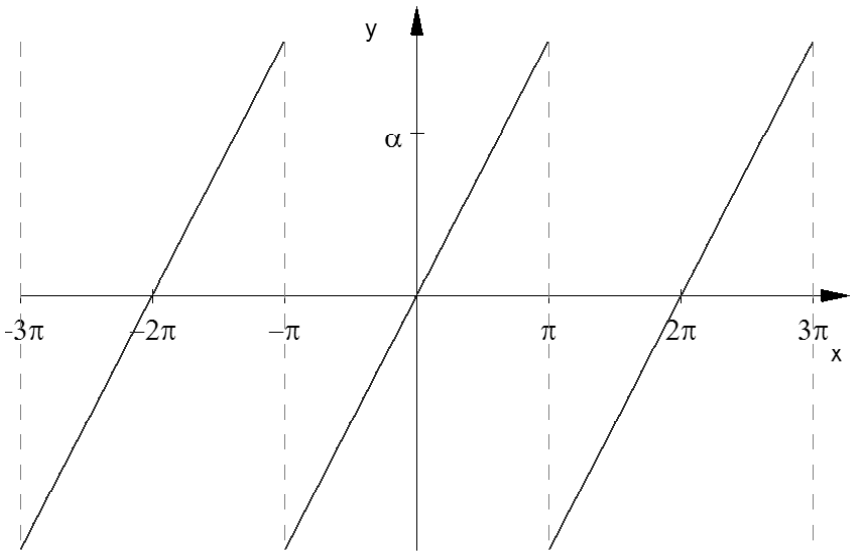
$$f(x) = \left(x^2 - \pi^2\right)^2, f(-\pi) = f(\pi) = 0$$
$$f'(x) = 4x\left(x^2 - \pi^2\right), f'(-\pi) = f'(\pi) = 0$$



(3). Ist $f \in C^1([-\pi, \pi])$ mit $f(-\pi) = f(\pi)$, so konvergiert die die Fourierreihe (1.3.4.3) gleichmäßig gegen f auf $[-\pi, \pi]$.
Noch allgemeiner gilt: Haben f und f' nur endlich viele Sprungstellen $s_0, \ldots, s_k \in [-\pi, \pi]$ (sonst sind f, f' stetig), so konvergiert die Fourierreihe (1.3.4.3) gleichmäßig gegen f in jedem kompakten Teilintervall, welches kein s_j enthält. In s_j konvergiert die Fourierreihe gegen $\frac{1}{2}\left(\lim_{x\to s_j-} f(x) + \lim_{x\to s_j+} f(x)\right)$.
(Beweis s. Heuser II, Seiten 133-148.)

Beispiel 1.39.

$$f(x) = \begin{cases} x & \text{für } x \in (-\pi, \pi) \\ \alpha & \text{für } x = \{-\pi, \pi\} \end{cases},\ f(x + 2k\pi) = f(x)$$



f ist ungerade auf $(-\pi,\pi) \Rightarrow f(x) = \sum_{n=1}^{\infty} b_n \sin(nx)$ mit

$$b_n = \frac{2}{\pi}\int_0^{\pi} f(x)\sin(nx)\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\sin(nx)\,\mathrm{d}x = \frac{2}{n}(-1)^{n+1} \quad \text{für } n \in \mathbb{N}.$$

$$\sin x - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} - \ldots + \ldots = \begin{cases} \frac{x}{2} & \text{für } x \in (-\pi,\pi) \\ \frac{1}{4}\Big(\underbrace{\lim_{x\to\pi-} f(x)}_{=\pi} + \underbrace{\lim_{x\to\pi+} f(x)}_{=-\pi}\Big) = 0 & \text{für } x = \{-\pi,\pi\}\,. \end{cases}$$

Mit $x = \frac{\pi}{2}$ folgt

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - + \ldots = \sum_{l=0}^{\infty} \frac{(-1)^l}{2l+1}$$

Bemerkung. Später wird mit funktionalanalytischen Methoden bewiesen, dass für eine große Klasse von Funktionen f deren Fourierreihe im quadratischen Mittel auf $[-\pi,\pi]$ gegen f konvergiert, das heißt im Sinne der Metrik $d(f,g) = \sqrt{\int_{-\pi}^{\pi} |f-g|^2 \,\mathrm{d}x}$.

1.4 Stetige Abbildungen metrischer Räume

1.4.1 Definitionen und Beispiele

Seien $(X,d),(Y,\rho)$ metrische Räume und $f : X \to Y$.

Definition 1.40. a. f heißt *stetig* in $x_0 \in X$, falls für jede Folge $x_n \subset X$ mit $\lim_{n\to\infty} x_n = x_0$ gilt: $\lim_{n\to\infty} f(x_n) = f(x_0)$

b. f heißt *stetig auf* X, falls f stetig in allen $x \in X$ ist. Bezeichnung: $f \in C(X, Y)$.

c. Eine bijektive Abbildung $f : X \to Y$ heißt *Homöomorphismus*, falls f und f^{-1} stetig sind.

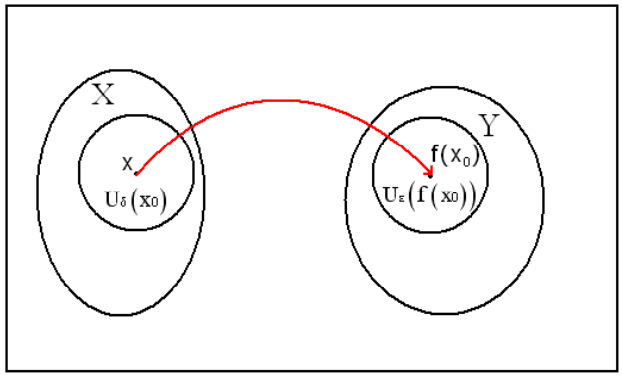
Bemerkung 1.41. 1. Wie bei reellen Funktionen ist a) äquivalent zum ε - δ - Kriterium, d.h.:
f ist stetig in $x_0 \in X$

$\Updownarrow$

für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert mit $\rho(f(x), f(x_0)) < \varepsilon$ für alle $x \in X$ mit $d(x, x_0) < \delta$

$\Updownarrow$

für alle $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ existiert mit $f(U_\delta(x_0)) \subseteq U_\varepsilon(f(x_0))$.



2. Aus 1. ergibt sich, dass die Hintereinanderausführung stetiger Funktionen f, g mit $f : X \to Y, g : Y \to Z$ wieder stetig ist ($g \circ f : X \to Z, g \circ f(x) := g(f(x))$.

Beispiel 1.42. i. $\mathrm{id} : X \to X, \mathrm{id}(x) = x$ ist stetig.

ii. Für Abbildungen $f : D \subseteq \mathbb{R}^N \to \mathbb{R}^M$ ist die Stetigkeit von f in $x_0 \in D$ gleichbedeutend zur Stetigkeit jeder Koordinatenfunktion $f_1, \ldots, f_M$ in $x_0 \in D$, wobei

$$f(x) = (f_1(x_1, \ldots, x_M), \ldots, f_M(x_1, \ldots, x_M))$$

gilt mit $f_i : D \to \mathbb{R}, i = 1, \ldots, M$, denn sei

$$x_n \in D, \lim_{n\to\infty} x_n = x_0 \;\Leftrightarrow\; \lim_{n\to\infty} x_{n,j} = x_{0,j} \text{ für alle } j = 1, \ldots, N.$$

Dann muss gelten

$$f(x_n) \to f(x_0) \;\Leftrightarrow\; f_i(x_n) \to f_i(x_0) \text{ für alle } i = 1, \ldots, M, \text{ wenn } n \to \infty.$$

iii. Jede lineare Abbildung $T : \mathbb{R}^N \to \mathbb{R}^M$ ist stetig, denn mit einer Darstellungsmatrix

$$A = (a_{ij})_{i=1,\ldots,M;j=1,\ldots,N} \text{ ist } T(x) = Ax = \Big(\sum_{j=1}^{N} a_{ij}x_j\Big)_{i=1,\ldots,M}, a_{ij} \in \mathbb{R}$$

iv. Eine reguläre Matrix $A \in \mathbb{R}^{N\times N}$ (d.h. $\det A \neq 0$) vermittelt einen Homöomorphismus $T : \mathbb{R}^N \to \mathbb{R}^N$, denn $T(x) = Ax$ ist stetig und $T^{-1}(x) = A^{-1}x$ ist ebenfalls stetig.

v. In einem normierten Vektorraum $(V, \|\cdot\|)$ ist $\|\cdot\| : V \to \mathbb{R}$ eine stetige Abbildung, denn

$$|\|x_n\| - \|x_0\|| \leq \|x_n - x_0\|$$

(aus Dreiecksungleichung und Homogenität von $\|\cdot\|$)

1.4.2 Charakterisierung stetiger Abbildungen

1. $f : X \to Y$ ist stetig genau dann, wenn für jede offene Menge $M \subseteq Y$

$$\text{das Urbild} \quad f^{-1}(M) = \{x \in X \,|\, f(x) \in M\}$$

eine offene Menge in X ist.

Beweis. (a) ($\Rightarrow$) Sei f stetig auf X und $x_0 \in f^{-1}(M)$.
Zu zeigen: Es existiert $U_r(x_0) \subseteq f^{-1}(M)$.

$$x_0 \in f^{-1}(M) \Leftrightarrow f(x_0) \in M \,.$$

Wegen Offenheit von M gibt es $U_\varepsilon(f(x_0)) \subset M$ mit $\varepsilon > 0$. Wegen Stetigkeit von f existiert $\delta = \delta(\varepsilon)$ mit

$$f(U_\delta(x_0)) \subset U_\varepsilon(f(x_0)) \subset M \;\Rightarrow\; U_\delta(x_0) \subset f^{-1}(M).$$

(b) ($\Leftarrow$) Wir zeigen Stetigkeit von f in $x_0 \in X$ mittels ε - δ - Kriterium.
Sei $\varepsilon > 0$. Damit ist $M := U_\varepsilon(f(x_0))$ eine offene Menge von Y. Nach Vorraussetzung muss auch $f^{-1}(M)$ offen sein. Das bedeutet, dass $x_0 \in f^{-1}(M)$ ein innerer Punkt ist, also existiert ein $\delta > 0$, so dass

$$U_\delta(x_0) \subset f^{-1}(M) \;\Leftrightarrow\; f(U_\delta(x_0)) \subset U_\varepsilon(f(x_0))$$

2. $f : X \to Y$ ist stetig genau dann, wenn für jede abgeschlossene Menge $M \subseteq Y$ ist $f^{-1}(M)$ abgeschlossen in X.
(Beweis über Komplementbildung aus 1, siehe Übungsaufgabe)

Bemerkung 1.43. Beachte: $f(A)$ muss nicht offen sein, auch wenn $A \subseteq X$ offen ist. Z.B.:

$$f(x) = 1,\; f : \mathbb{R} \to \mathbb{R}, f(\underbrace{\mathbb{R}}_{\text{offen}}) = \underbrace{\{1\}}_{\text{abgeschlossen}}$$

1.4.3 Weitere Eigenschaften stetiger Abbildungen

Satz 1.44. Sei $K \subseteq X$ kompakt, $f : X \to Y$ stetig. Dann ist auch $f(K)$ kompakt.

Beweis. Siehe Übungsaufgabe.

Folgerung 1.45. *(Verallgemeinerung des Extremalsatzes von Weierstraß)*
Ist $f : X \to \mathbb{R}$ stetig (symbolisch: $f \in C(X, \mathbb{R})$) und $K \subseteq X$ kompakt, so hat f auf K ein Minimum und ein Maximum.

Beweis. Nach obigem Satz ist $f(K)$ eine kompakte Teilmenge von $\mathbb{R}$, das heißt $f(K)$ ist beschränkt und abgeschlossen. Gemäß Supremum- und Infimumprinzip in $\mathbb{R}$ existiert

$$G := \sup f(K) = \sup \{f(x) \,|x \in K\} = \sup_K f \,,$$
$$g := \inf f(K) = \inf \{f(x) \,|x \in K\} = \inf_K f \,.$$

Nach Definition von G und g gibt es Folgen$(G_n), (g_n) \subset f(K)$ mit $\lim\limits_{n\to\infty} G_n = G$, $\lim\limits_{n\to\infty} g_n = g$. Wegen Abgeschlossenheit von $f(K)$ ist $G, g \in f(K)$, das heißt es existieren $x_1, x_2 \in K$, so dass gilt

$$f(x_1) = G = \max_K f, \; f(x_2) = g = \min_K f \,.$$

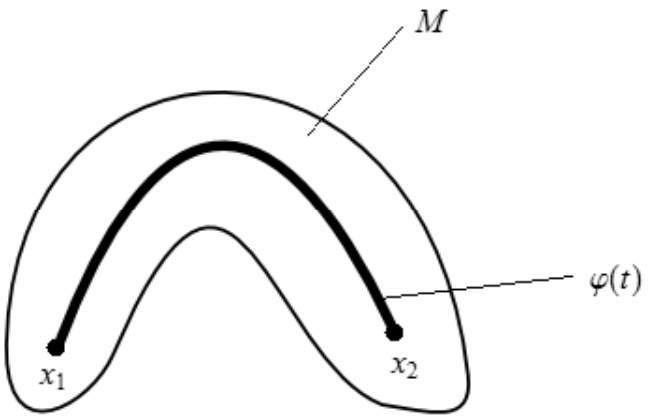
Beispiel 1.46.

$$f(x) = \left|x_4^{16} x_2 \cos x_1\right| \mathrm{e}\left(x_1 x_3 + x_4^2\right), \quad f : \mathbb{R}^4 \to \mathbb{R}^1$$

f hat ein Maximum und ein Minimum auf $[-1, 0] \times [0, 2] \times [14, 17] \times [-\pi, \pi]$.

Zur Übertragung des Zwischenwertsatzes für reelle stetige Funktionen benötigen wir den Begriff der zusammenhängenden Menge.

Definition 1.47. a. $M \subseteq X$ heißt *(weg-)zusammenhängend*, wenn für alle $x_1, x_2 \in M$ ein $\varphi \in C([a, b], X)$ existiert mit $\varphi(a) = x_1$, $\varphi(b) = x_2$ und $\varphi(t) \in M$ für alle $t \in [a, b]$.



b. φ heißt *Weg* oder *Kurve* in M vom Anfangspunkt x_1 zum Endpunkt x_2.

c. Insbesondere nennt man φ *Polygonzug*, wenn $\varphi(t)$ nur aus endlich vielen Geradenstücken besteht unter der Vorraussetzung, dass X ein normierter Vektorraum ist.

d. Speziell heißt eine nichtleere offene und zusammenhängende Menge ein *Gebiet.*

Beispiel 1.48. i. $[a, b)$ ist zusammenhängende Menge des $\mathbb{R}^1$, jedoch kein Gebiet. (a, b) ist eine zusammenhängende Menge des $\mathbb{R}^1$ und auch Gebiet.

ii. $M = [-1, 0) \cup (0, 1]$ ist nicht zusammenhängend.

iii. $\mathbb{N} \subset \mathbb{R}$ ist nicht zusammenhängend.

iv. $U_\varepsilon(x_0) \subset V$ ist zusammenhängend mit V als normiertem Vektorraum, denn die Verbindungsgerade G zwischen 2 Punkten $x_1, x_2 \in U_\varepsilon(x_0)$ liegt vollständig in $U_\varepsilon(x_0)$.

$$G = \{x \in V |\, x = \underbrace{x_1 + t(x_2 - x_1)}_{=:\varphi(t)},\ t \in [0, 1]\}$$

Dann ist $\varphi : [0, 1] \to V$ stetig und $\varphi(0) = x_1, \varphi(1) = x_2$. Sei $x \in G$, so untersuche

$$\begin{aligned} \|x - x_0\| &= \|x_1 + t(x_2 - x_1) - \underbrace{1}_{=1-t+t} \cdot x_0\| \\ &= \|t(x_2 - x_0) + (1 - t)(x_1 - x_0)\| \\ &\leq \|t(x_2 - x_0\| + \|(1 - t)(x_1 - x_0)\| \\ &= t\underbrace{\|x_2 - x_0\|}_{<\varepsilon} + (1 - t)\underbrace{\|x_1 - x_0\|}_{<\varepsilon} \quad (t \in [0, 1]) \\ &< \varepsilon\ . \end{aligned}$$

v. $M = \left\{(x_1, x_2) \in \mathbb{R}^2 \,\middle|\, (x_1 - 1)^2 + x_2^2 < 1 \text{ oder } (x_1 + 1)^2 + x_2^2 < 1\right\}$ ist nicht zusammenhängend, weil $M = U_1((1, 0)) \cup U_1((-1, 0))$.

Satz 1.49. Sei $f \in C(X, Y)$ und $M \subseteq X$ eine zusammenhängende Menge. Dann ist auch $f(M) \subseteq Y$ zusammenhängend.

Beweis. Sei $y_1, y_2 \in f(M)$, dann existieren $x_1, x_2 \in M$ mit $f(x_i) = y_i$ für $i = 1, 2$. Wegen Zusammenhang von M gibt es ein $\varphi \in C([a, b], X)$ mit $\varphi(t) \in M$ für alle $t \in [a, b]$, $\varphi(a) = x_1$, $\varphi(b) = x_2$. Dann liefert $f \circ \varphi : [a, b] \to Y$ eine stetige Abbildung mit

$$(f \circ \varphi)(a) = f(x_1) = y_1,\ (f \circ \varphi)(b) = f(x_2) = y_2 \text{ und } (f \circ \varphi)(t) = f(\underbrace{\varphi(t)}_{\in M}) \in f(M)\ .$$

Spezialfall: Ist $f \in C(X, \mathbb{R})$ und $M \subseteq X$ zusammenhängend, dann nimmt f jeden Wert zwischen $f(x_1)$ und $f(x_2)$ an für $x_1, x_2 \in M$, denn $f(M)$ ist zusammenhängend in $\mathbb{R}$, d.h. ein Intervall, welches $f(x_1)$ und $f(x_2)$ enthält.

1.4.4 Kontraktive Abbildungen

1.4.4.1 Kontraktionen

Definition 1.50. Sei X ein metrischer Raum und $f : X \to X$. Gilt mit einer Konstanten $q \in [0, 1)$:

$$d(f(x), f(y)) \leq q \cdot d(x, y) \text{ für alle } x, y \in X, \tag{1.4.4.5}$$

so heißt f eine *Kontraktion* oder *kontraktive Selbstabbildung* auf X und q *Kontraktionskonstante* von f.

Folgerung 1.51. Jede Kontraktion f ist stetig, denn aus der Definition folgt:

$$0 \le d(f(x), f(x_n)) \le q \cdot \underbrace{d(x, x_n)}_{\to 0 \text{ für } n\to\infty}$$

für irgendeine Folge $(x_n) \subset X$ mit $\lim\limits_{n\to\infty} x_n = x \;\Rightarrow\; d(f(x), f(x_n)) \to 0$.

Beispiel 1.52. i. $f(x) = \frac{1}{10}\mathrm{e}^x$ ist eine Kontraktion auf $[0, 1]$, denn

$$0 \le x \le 1 \;\Rightarrow\; \frac{1}{10} \le \frac{1}{10}\mathrm{e}^x \le \frac{1}{10}\mathrm{e} \;\Rightarrow\; f(x) \in [0, 1]\,.$$

Überprüfung auf Kontraktion:

$$|f(x) - f(y)| = \frac{1}{10}\,|\mathrm{e}^x - \mathrm{e}^y| = \frac{1}{10}\left|(\mathrm{e}^x)'\right|_{x=\xi}(x - y);\; \xi = x + \vartheta(y - x), \vartheta \in (0, 1)\,,$$
$$|f(x) - f(y)| = \frac{1}{10}\mathrm{e}^\xi\,|x - y| \le \frac{1}{10}\mathrm{e}\,|x - y| \quad \text{für alle } x, y \in [0, 1] \;\Rightarrow\; q = \frac{\mathrm{e}}{10} \in (0, 1)\,.$$

ii. Sei eine lineare Abbildung $\mathbb{C}^N \to \mathbb{C}^N$ gegeben durch die Darstellungsmatrix $A = \left(a_{ij}\right)_{i,j=1,\dots,N}$. Wann ist A eine Kontraktion auf $\mathbb{C}^N$?

$$\|Ax - Ay\|^2 = \|A(x - y)\|^2 = \sum_{i=1}^{N}\left|\sum_{j=1}^{N} a_{ij}\left(x_j - y_j\right)\right|^2$$

Aus Cauchy-Schwarzscher Ungleichung folgt $\left|\sum\limits_{j=1}^{N} a_{ij}\left(x_j - y_j\right)\right|^2 \le \sum\limits_{j=1}^{N}\left|a_{ij}\right|^2 \sum\limits_{j=1}^{N}\left|x_j - y_j\right|^2$.

$$\Rightarrow\; \|Ax - Ay\|^2 \le \underbrace{\sum_{j=1}^{N}\left|x_j - y_j\right|^2}_{=\|x-y\|^2}\underbrace{\sum_{i,j=1}^{N}\left|a_{ij}\right|^2}_{=:\|A\|_2^2}$$

Ergebnis: A ist eine Kontraktion auf $\mathbb{C}^N$, falls $q = \|A\|_2 < 1$ ist.

1.4.4.2 Fixpunkte

Kontraktionen spielen die entscheidende Rolle bei der Betrachtung von Fixpunktgleichungen für $f : X \to X$, das heißt für die Bestimmung solcher $x \in X$ mit $f(x) = x$. x nennt man *Fixpunkt* von f.
In normierten Vektorräumen X ist die Fixpunktgleichung äquivalent zur Lösung der nichtlinearen Gleichung $f(x) - x = 0$. Umgekehrt ist

$$F(x) = 0 \;\Leftrightarrow\; x = x - G(F(x))$$

mit $G(0) = 0$, $G(y) \neq 0$ für $y \neq 0$ (Fixpunktgleichung für $\mathrm{id} - G \circ F$).

Satz 1.53. *(Banachscher Fixpunktsatz)*
Sei (X, d) ein vollständiger metrischer Raum und f eine Kontraktion auf X. Dann gilt:

a. Die Abbildung f besitzt genau einen Fixpunkt $x \in X$.

b. Der Fixpunkt x ergibt sich durch *sukzessive Approximation*:

$$x_{n+1} := f(x_n),\ n \in \mathbb{N}_0 \quad \textit{(Iterationsfolge)} \tag{1.4.4.6}$$

mit beliebigem Startpunkt $x_0 \in X$ und $x = \lim\limits_{n\to\infty} x_n$.

c. Fehlerabschätzung

$$d(x_n, x) \le \frac{q^n}{1-q}\, d(x_1, x_0)$$

Bemerkung 1.54. Der Fixpunktsatz gilt auch, wenn X ersetzt wird durch ein abgeschlossenes $M \subset X$.

Beispiel 1.55. $f(x) = \frac{1}{10}e^x$, $x \in [0,1]$. f ist eine Kontraktion auf $[0,1]$. Somit existiert eine Lösung $x \in [0,1]$ der Gleichung $\frac{1}{10}e^x = x$.

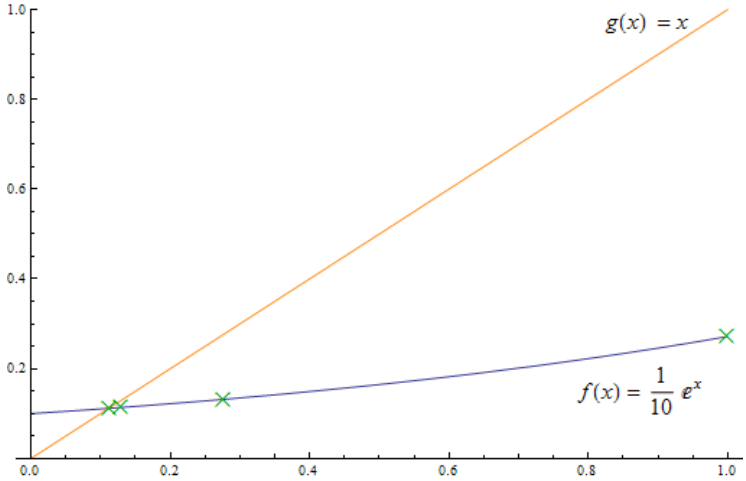
$x_0 = 1$

$x_1 = f(x_0) = \dfrac{1}{10}e \approx 0.278$

$x_2 = f(x_1) = \dfrac{1}{10}e\left(\dfrac{e}{10}\right) \approx 0.131$

$x_3 = f(x_2) \approx 0.114$

$x_4 = f(x_3) \approx 0.112$



Berechnung von x auf $\frac{1}{100}$ genau: Benutze 1.53.c

$$\underbrace{d(x_n, x)}_{=|x_n - x|} \le \frac{\left(\frac{e}{10}\right)^n}{1-\frac{e}{10}}\ \underbrace{d(x_1, x_0)}_{=|x_1-x_0|=\left|\frac{e}{10}-1\right|} \quad \Rightarrow\ |x_n - x| \le \frac{e^n}{10^n} < \frac{1}{100} \ \Leftrightarrow\ e^n < 10^{n-2}$$

Dies gilt für $n \ge 4$, also Ergebnis: $x \in \left[x_4 - \frac{1}{100}, x_4 + \frac{1}{100}\right] = [0.102, 0.122]$.

Beweis. (Beweis des Banachschen Fixpunktsatzes)

(i). Eindeutigkeit des Fixpunktes. Seien $x, \tilde{x} \in X$ Fixpunkte von f, dann ist

$$d(x, \tilde{x}) = d(f(x), f(\tilde{x})) \underset{(1.4.4.5)}{\le} q d(x, \tilde{x}) \ \Rightarrow\ 0 \le \underbrace{(q-1)}_{<0} d(x, \tilde{x})$$

$$\Rightarrow\ d(x, \tilde{x}) = 0 \ \Leftrightarrow\ x = \tilde{x}\,.$$

(ii). Die Iterationsfolge (x_n) ist eine Cauchyfolge.

$$\begin{aligned} d(x_{k+1}, x_k) &\underset{(1.4.4.6)}{=} d(f(x_k), f(x_{k-1})) \underset{(1.4.4.5)}{\leq} qd(x_k, x_{k-1}) \underset{(1.4.4.6)}{=} qd(f(x_{k-1}), f(x_{k-2})) \\ &\underset{(1.4.4.5)}{\leq} q^2 d(x_{k-1}, x_{k-2}) \underset{(1.4.4.5),(1.4.4.6)}{\leq} \dots \underset{(1.4.4.5),(1.4.4.6)}{\leq} q^k d(x_1, x_0) \\ \Rightarrow\ d(x_{n+p}, x_n) &\leq d(x_{n+p}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\leq d(x_{n+p}, x_{n+2}) + d(x_{n+2}, x_{n+1}) + d(x_{n+1}, x_n) \\ &\vdots \\ &\leq \sum_{l=0}^{p-1} \underbrace{d(x_{n+l+1}, x_{n+l})}_{\leq q^{n+l} d(x_1, x_0)} \quad \text{(für } p \in \mathbb{N}\text{, Übungsaufgabe)} \\ \Rightarrow d(x_{n+p}, x_n) &\leq d(x_1, x_0) \sum_{l=0}^{p-1} q^{n+l} \end{aligned}$$

Wegen $q \in [0, 1)$ lässt sich weiter abschätzen

$$d(x_{n+p}, x_n) \leq d(x_1, x_0) q^n \sum_{l=0}^{\infty} q^l = \frac{q^n}{1-q} d(x_1, x_0). \tag{1.4.4.7}$$

Weil $\lim\limits_{n\to\infty} q^n = 0$, folgt $d(x_{n+p,x_n}) < \varepsilon$ für $n \geq n_0(\varepsilon)$ und alle $p \in \mathbb{N}$. Somit ist (x_n) eine Cauchyfolge.

(iii). Existenz des Fixpunktes. Wegen Vollständigkeit von X konvergiert (x_n) gegen ein $x \in X$. Wir zeigen, dass x Fixpunkt von f ist.

$$d(f(x_n), f(x)) \underset{(1.4.4.5)}{\leq} qd(x_n, x) \to 0 \ \text{ für } \ n \to \infty \ \Rightarrow \ f(x_n) \to f(x).$$

Der Grenzübergang $n \to \infty$ in der Iterationsgleichung (1.4.4.6) liefert

$$\underbrace{\lim_{n\to\infty} x_{n+1}}_{=x} = \underbrace{\lim_{n\to\infty} f(x_n)}_{=f(x)}.$$

(iv). Beweis der Fehlerabschätzung 1.53.c. Wählt man ein festes n in (1.4.4.7) und betrachtet den Grenzübergang $p \to \infty$, so ergibt sich

$$\underbrace{\lim_{p\to\infty} d(x_{n+p}, x_n)}_{=d(x,x_n)} \leq \frac{q^n}{1-q} d(x_1, x_0)$$

wegen Stetigkeit von d (siehe Übungsaufgabe).

1.4.4.3 Anwendungen

1. Das *Newton-Verfahren* dient der angenäherten Berechnung einer Lösung von $f(x) = 0$ für eine reelle Funktion f.

$$f(x) = 0 \;\Leftrightarrow\; x = x - \frac{f(x)}{f'(x)} := g(x) \;\text{ für } f' \neq 0$$

$$x_{n+1} = g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)} \qquad \textit{(Newton-Folge).}$$

   (s. Vorlesung Numerik)

2. Existenz- und Eindeutigkeitssatz zum Anfangswertproblem für gewöhnliche Differentialgleichungen (s. Abschnitt 1.5).

1.5 Systeme gewöhnlicher Differentialgleichungen erster Ordnung

1.5.1 Allgemeine Existenz- und Eindeutigkeitssätze

1.5.1.1 Eine Differentialgleichung

In diese Abschnitt wird ein allgemeiner Existenz- und Eindeutigkeitssatz für das Anfangswertproblem zur Differentialgleichung

$$y' = f(x, y)\,, \quad y(x_0) = y_0 \tag{1.5.1.8}$$

(und danach für Differentialgleichungssysteme) hergeleitet. Der Begriff der Lipschitz-Stetigkeit spielt dabei eine wichtige Rolle.

Definition 1.56. Eine Funktion g heißt *Lipschitz-stetig* im Intervall $[c, d]$, wenn es eine positive Konstante L *(Lipschitz-Konstante)* gibt, so dass

$$|f(z_1) - f(z_2)| \leq L\,|z_1 - z_2| \qquad \text{für alle } z_1, z_2 \in [c, d]. \tag{1.5.1.9}$$

Bemerkung 1.57. 1. Ist die Funktion $g \in C^1([c, d])$, dann genügt g in diesem Intervall einer Lipschitz-Bedingung. Das folgt unmittelbar aus dem Mittelwertsatz der Differentialrechnung, denn für $z_1, z_2 \in [c, d]$ ist

$$\left|\frac{g(z_1) - g(z_2)}{z_1 - z_2}\right| = |g'(\xi)| \leq \max_{[c,d]} |g'| = \|g'\|_{C([c,d])}$$

mit einem Zwischenwert $\xi = z_1 + \vartheta(z_2 - z_1)$, $\vartheta \in (0, 1)$.
Die beschränkte Differenzierbarkeit einer Funktion stellt also eine hinreichende (aber nicht notwendige) Bedingung für ihre Lipschitz-Stetigkeit dar.

2. Die Funktion $f : y \mapsto \sqrt{|y|}$ stetig auf $\mathbb{R}$. Das Anfangswertproblem

$$y' = \sqrt{|y|}\,,\ y(0) = 0$$

besitzt jedoch mehrere Lösungen, etwa $y_1 \equiv 0$ oder $y_2 = \frac{1}{4}x|x|$.
Somit kann die Stetigkeit der rechten Seite in (1.5.1.8) nicht ausreichend sein für die Eindeutigkeit der Lösung dieses Anfangswertproblem.

Satz 1.58. *(Satz von Picard-Lindelöf)*
Die rechte Seite f in (1.5.1.8) sei stetig im Rechteck

$$Q = \{(x, y) \in \mathbb{R}^2 : |x - x_0| \le a, |y - y_0| \le b\} = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b].$$

Für $(x, y) \in Q$ gelte $|f(x, y)| \le M$. Ferner genüge f einer Lipschitz-Bedingung

$$|f(x, y) - f(x, \bar{y})| \le L\,|y - \bar{y}| \qquad \text{für alle } (x, y), (x, \bar{y}) \in Q \tag{1.5.1.10}$$

bezüglich der zweiten Variablen. Wählt man $\alpha = \min\left\{a, \frac{b}{M}, \frac{1}{L+1}\right\}$, dann gibt es im Intervall $I := [x_0 - \alpha, x_0 + \alpha]$ genau eine stetig differenzierbare Lösung y des Anfangswertproblems (1.5.1.8).

Beweis. (i). Mittels Hauptsatz der Differential- und Integralrechnung sieht man: Eine stetig differenzierbare Funktion y ist genau dann Lösung unseres Anfangswertproblems (1.5.1.8), wenn sie (lediglich) stetig ist und der *Integralgleichung*

$$y(x) = y_0 + \int_{x_0}^{x} f(z, y(z))dz =: T[y](x) \tag{1.5.1.11}$$

für alle $x \in I$ genügt. Die Lösung des Anfangswertproblems (1.5.1.8) und die Lösung der Integralgleichung (1.5.1.11) innerhalb der Klasse stetiger Funktionen sind also äquivalente Problemstellungen.

(ii). Die Integralgleichung (1.5.1.11) ist eine Fixpunktgleichung für die Abbildung T. Diese wollen wir mittels des Banachschen Fixpunktsatzes 1.53 lösen. Dazu überprüfen wir dessen Voraussetzungen. Sei

$$\mathcal{M} := \left\{ y \in C(I) \,\middle|\, \max_{z \in I} |y(z) - y_0| = \|y - y_0\|_{C(I)} \le b \right\}.$$

Dann ist $\mathcal{M}$ ein abgeschlossene Teilmenge von $C(I)$, und T hat folgende Eigenschaften (für $x_0 \le x \le x_0 + \alpha$, analog für $x_0 - \alpha \le x \le x_0$):

a) $$|T[y](x) - y_0| \le \int_{x_0}^{x} \underbrace{|f(z, y(z))|}_{\le M} dz \le M \underbrace{(x - x_0)}_{\le \alpha} \le M\frac{b}{M} = b$$

für $y \in \mathcal{M}$, also $T : \mathcal{M} \to \mathcal{M}$;

b) $$|T[y](x) - T[\bar{y}](x)| \le \int_{x_0}^{x} \underbrace{|f(z, y(z)) - f(z, \bar{y}(z))|}_{\le L|y(z) - \bar{y}(z)|} dz \le L \max_{z \in I} |y(z) - \bar{y}(z)| \underbrace{(x - x_0)}_{\le \alpha}$$

$$\le \frac{L}{L+1}\|y - \bar{y}\|_{C(I)}$$

für $y, \bar{y} \in \mathcal{M}$, also T ist Kontraktion auf $\mathcal{M}$.

Weitere wichtige Sätze:

(a). Existenzsatz von Peano (mittels Euler-Cauchyschen Polygonzügen, wobei nur Stetigkeit der rechten Seite f vorausgesetzt wird)

(b). Existenz- und Eindeutigkeitssatz von Cauchy (mittels Potenzreihenansatz)

1.5.1.2 Differentialgleichungssysteme

Alle angeführten Sätze gelten sinngemäß auch für Systeme

$$y_i' = f_i(x, y_1, \ldots, y_n); \quad i = 1, \ldots, n \tag{1.5.1.12}$$

von n Differentialgleichungen 1. Ordnung für n gesuchte Funktionen $y_1, \ldots, y_n$ der unabhängigen Variablen x. Das Anfangswertproblem für (1.5.1.12) besteht in der Vorgabe der n Funktionswerte $y_i(x_0) = y_{i0}$ in x_0. Fasst man sowohl die Unbekannten y_i als auch die rechten Seiten f_i zu Spaltenvektoren

$$\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{bzw.} \quad \vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

zusammen, dann kann das Anfangswertproblem zu (1.5.1.12) kurz geschrieben werden (bei komponentenweiser Differentiation)

$$\vec{y}\,' = \vec{f}(x, \vec{y}), \qquad \vec{y}(x_0) = \vec{y}_0. \tag{1.5.1.13}$$

Insbesondere gilt der Picard-Lindelöfsche Existenz- und Eindeutigkeitssatz für Differentialgleichungssystem unter folgenden Voraussetzungen:

$$\vec{f} \in C(Q, \mathbb{R}^n), \quad Q = [x_0 - a, x_0 + a] \times [y_{01} - b, y_{01} + b] \times \cdots \times [y_{0n} - b, y_{0n} + b] \subset \mathbb{R}^{1+n},$$
$$|f_i(x, \vec{y})| \le M, \quad \left|f_i(x, \vec{y}) - f_i(x, \vec{\bar{y}})\right| \le L \max_{j=1,\ldots,n} \left|y_j - \bar{y}_j\right| \quad \text{für alle } (x, \vec{y}), (x, \vec{\bar{y}}) \in Q,\ i = 1, \ldots, n.$$

Beispiel 1.59. Radioaktive Zerfallsreihe: Eine radioaktive Ausgangssubstanz S_1 zerfällt in S_2, S_2 zerfällt in $S_3, \ldots$, z.B. U^{238} zerfällt über 13 radioaktive Zwischenstufen in stabiles Blei. Sei $\lambda_i > 0$ die Zerfallskonstante von S_i: $S_1 \overset{\lambda_1}{\to} S_2 \overset{\lambda_2}{\to} S_3 \to \cdots \overset{\lambda_{n-1}}{\to} S_n$. Das zugehöriges DGL-

System für $m_i = m_i(t)$ (Masse der Substanz S_i zur Zeit $t \geq 0$)

$$\begin{aligned}
\frac{dm_1}{dt} &= -\lambda_1 m_1 \\
\frac{dm_2}{dt} &= \lambda_1 m_1 - \lambda_2 m_2 \qquad \text{(Herleitung: } m_2(t+\Delta t) = m_2(t) + \lambda_1 m_1(t)\Delta t - \lambda_2 m_2(t)\Delta t\text{)} \\
&\vdots \\
\frac{dm_{n-1}}{dt} &= \lambda_{n-2} m_{n-2} - \lambda_{n-1} m_{n-1} \\
\frac{dm_n}{dt} &= \lambda_{n-1} m_{n-1}
\end{aligned}$$

$$\text{bzw.} \quad \frac{d}{dt}\begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} = \begin{pmatrix} -\lambda_1 & & & 0 \\ \lambda_1 & -\lambda_2 & & \\ & & \ddots & \\ 0 & & \lambda_{n-1} & 0 \end{pmatrix}\begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix}.$$

Anfangswerte: $m_1(0) = M,\ m_2(0) = \cdots = m_n(0) = 0$.

1.5.1.3 Bedeutung von Differentialgleichungssystemen 1. Ordnung

Diese Bedeutung wird ersichtlich, wenn man die *allgemeine explizite Differentialgleichung n-ter Ordnung*

$$y^{(n)}(x) = f(x, y'(x), y''(x), \ldots, y^{(n-1)}(x)), \quad x \in I \subseteq \mathbb{R} \tag{1.5.1.14}$$

betrachtet. Diese ist äquivalent zum Differentialgleichungssystem der folgenden Bauart: Setze $y_i(x) := y^{(i-1)}(x)$ für $i = 1, \ldots, n$, dann gilt

$$\begin{aligned}
y_1' &= y_2 \\
y_2' &= y_3 \\
&\vdots \\
y_{n-1}' &= y_n \\
y_n' &= f(x, y_1, \ldots, y_n) \qquad \text{(wegen (1.5.1.14)).}
\end{aligned}$$

Ähnlich lassen sich Systeme von expliziten Differentialgleichungen höherer Ordnung durch Einführung zusätzlicher Unbekannter fur die Ableitungen der gesuchten Funktionen in ein Differentialgleichungssystem 1. Ordnung überführen. Damit gelten die allgemeinen Existenz- und Eindeutigkeitssätze auch für solche Differentialgleichungen bzw. Systeme.

1.5.2 Lineare Differentialgleichungssysteme

Das System:

$$\begin{aligned} y_1'(x) &= a_{11}(x)y_1(x) + a_{12}(x)y_2(x) + \cdots + a_{1n}(x)y_n(x) + s_1(x) \\ y_2'(x) &= a_{21}(x)y_1(x) + a_{22}(x)y_2(x) + \cdots + a_{2n}(x)y_n(x) + s_2(x) \\ &\vdots \\ y_n'(x) &= a_{n1}(x)y_1(x) + a_{n2}(x)y_2(x) + \cdots + a_{nn}(x)y_n(x) + s_n(x) \end{aligned} \tag{1.5.2.15}$$

mit gegebenen Funktionen $a_{jk}, s_j \in C(I)$ $I \subseteq \mathbb{R}$ Intervall, heißt *lineares Differentialgleichungssystem 1. Ordnung*. Mit der *Koeffizientenmatrix* $A(x) = (a_{jk}(x))_{j,k=1,\ldots,n}$ und dem

$$\textit{Absolutgliedvektor} \text{ bzw. } \textit{Störung(svektor)} \quad \vec{s}(x) = \begin{pmatrix} s_1(x) \\ \vdots \\ s_n(x) \end{pmatrix}$$

schreibt sich (1.5.2.15) für den gesuchten Funktionenvektor $\vec{y} = y(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix}$ in der Form

$$\vec{y}'(x) = A(x)\vec{y}(x) + \vec{s}(x)$$

Das zu (1.5.2.15) gehörige *homogene* lineare Differentialgleichungssystem lautet

$$\vec{y}'(x) = A(x)\vec{y}(x) \tag{1.5.2.16}$$

Ist $\vec{s} \neq 0$ in (1.5.2.15), so heißt das System *inhomogen*.
Wortwörtlich lässt sich das Superpositionsprinzip für eine lineare Differentialgleichung 1. Ordnung auf System (1.5.2.15) übertragen. Existenz und Eindeutigkeit der Lösung des Anfangswertproblems für (1.5.2.15) ist gesichert durch den Satz von Picard-Lindelöff, denn die rechte Seite dieses Differentialgleichungssystems ist $f(x, \vec{y}) = A(x)\vec{y} + \vec{s}(x)$. Dafür gilt:

$$\begin{aligned} |f_j(x,\vec{y}) - f_j(x,\vec{\tilde{y}})| = \left|\sum_{k=1}^{n} a_{jk}(x)(y_k - \tilde{y}_k)\right| &\leq \sum_{k=1}^{n} |a_{jk}(x)| \underbrace{|y_k - \tilde{y}_k|}_{\leq \max\limits_{k=1,\ldots,n} |y_k - \tilde{y}_k|} \\ &\leq \underbrace{n \max_{j,k=1,\ldots,n} \|a_{jk}\|_{C(I)}}_{=:L} \max_{k=1,\ldots,n} |y_k - \tilde{y}_k| \quad \text{für alle } x \in I, \text{ und alle } \vec{y}, \vec{\tilde{y}} \in \mathbb{R}^n. \end{aligned}$$

D.h. rechte Seite von (1.5.2.15) genügt einer Lipschitzbedingung bez. $\vec{y}$ unabhängig von x und unabhängig von von den Anfangsdaten $x_0 \in I$, $\vec{y}_0 \in \mathbb{R}^n$. Damit kann gezeigt werden, dass die Lösung eines linearen Differentialgleichungssystems im *gesamten* Stetigkeitsintervall I der Daten a_{ij}, b_i existiert.

1.5.2.1 Struktur der Lösungen

Definition 1.60. i. p stetige Funktionenvektoren $\vec{z}_1, \ldots, \vec{z}_p$ heißen *linear unabhängig* auf I, wenn sie als Elemente des Vektorraums $C(I, \mathbb{R}^n)$ linear unabhängig sind, d.h., wenn aus

$$c_1\vec{z}_1(x) + \cdots + c_p\vec{z}_p(x) = 0 \quad \text{für alle } x \in I$$

folgt $c_1 = \cdots = c_p = 0$. Andernfalls heißen sie *linear abhängig*.

ii. n linear unabhängige Lösungsvektoren $\vec{y}_1, \ldots, \vec{y}_n$ des homogenen Differentialgleichungssystems (1.5.2.16) nennt man *Fundamentalsystem (Integralsystem, Hauptsystem)* zu (1.5.2.16) und $Y(x) := \left(\vec{y}_1(x), \ldots, \vec{y}_n(x)\right)$ *Fundamentalmatrix*.

Bemerkung 1.61. Versteht man die Differentiation einer Matrix komponentenweise, so genügt die Fundamentalmatrix Y der Differentialgleichung

$$Y'(x) = A(x)Y(x) \tag{1.5.2.17}$$

Satz 1.62. *(Kriterium für Fundamentalsystem)*
Sei $a_{jk} \in C(I)$, $I \subseteq \mathbb{R}$ ein Intervall. Dann bilden n Lösungen $\vec{y}_1, \ldots, \vec{y}_n$ des linearen homogenen Differentialgleichungssystems (1.5.2.16) genau dann ein Fundamentalsystem auf I, wenn ihre *Wronski-Determinante*

$$W[\vec{y}_1, \ldots, \vec{y}_n](x) := \det Y(x)$$

für alle Punkte $x \in I$ verschieden von 0 ist.

Beweis. Übungsaufgabe

Satz 1.63. *(Struktursatz)*
Seien a_{jk}, s_j stetig auf I, dann gilt:

a. Das homogene Differentialgleichungssystem (1.5.2.16) besitzt stets ein Fundamentalsystem.

b. Jede Lösung von (1.5.2.16) kann als Linearkombination aus den Funktionenvektoren $\vec{y}_1, \ldots, \vec{y}_n$ eines Fundamentalsystems dargestellt werden, d.h.:

$$\vec{y}_H(x) = c_1\vec{y}_1(x) + c_2\vec{y}_2(x) + \cdots + c_n\vec{y}_n(x) = Y(x)\vec{c}, \quad \vec{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n. \tag{1.5.2.18}$$

c. Die allgemeine Lösung $\vec{y} = \vec{y}(x)$ des inhomogenen Differentialgleichungssystems (1.5.2.15) lautet

$$\vec{y}(x) = \vec{y}_S(x) + \vec{y}_H(x),$$

wobei $\vec{y}_S$ eine spezielle Lösung von (1.5.2.15) und y_H allgemeine Lösung aus b. ist.

d. Eine spezielle Lösung $\vec{y}_S$ von (1.5.2.15) erhält man durch den Ansatz

$$\vec{y}_S(x) = Y(x)c(x) = c_1(x)\vec{y}_1(x) + c_2(x)\vec{y}_2(x) + \cdots + c_n(x)\vec{y}_n(x) \tag{1.5.2.19}$$

mit einem zu bestimmenden stetig differenzierbaren Funktionenvektor $\vec{c} : I \to \mathbb{R}^n$.

Beweis.

zu a. Konstruktion eines Fundamentalsystems $\vec{y}_1, \ldots, \vec{y}_n \in C^1(I, \mathbb{R}^n)$:
Sei $\vec{e}_j$ der j-te Einheitsvektor im $\mathbb{R}^n$, $E_n = (\vec{e}_1, \ldots, \vec{e}_n)$.Nach dem Existenz- und Eindeutigkeitssatz zu (1.5.2.16) existiert eine eindeutige Lösung mit den Anfangswerten

$$\vec{y}(x_0) = \vec{e}_j \quad \text{für } x_0 \in I.$$

Diese Lösung bezeichnen wir mit $\vec{y}_j$, $j = 1, \ldots, n$. Betrachten wir $w(x) := \det(\vec{y}_1, \ldots, \vec{y}_n)(x) = \det(Y)(x)$, so gilt $w(x_0) = \det(E_n) = 1 \neq 0$. Aus dem Kriterium für Fundamentalsysteme folgt, dass $\vec{y}_1, \ldots, \vec{y}_n$ linear unabhängig sind.

zu b. Sei $\vec{y}_H$ eine beliebige Lösung des homogenen Differentialgleichungssystems (1.5.2.16) auf I. Für einen Wert $x_0 \in I$ setzen wir $\vec{y}_H(x_0) =: \vec{y}_0$ Mit einem Fundamentalsystem $\vec{y}_1, \ldots, \vec{y}_n$ hat das lineare algebraische Gleichungssystem

$$Y(x_0)\vec{c} = \begin{pmatrix} y_{11}(x_0) & \ldots & y_{n1}(x_0) \\ \vdots & \ddots & \vdots \\ y_{1n}(x_0 & \ldots & y_{nn}(x_0) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \vec{y}_0, \quad \vec{c} \in \mathbb{R}^n$$

für $\vec{c}$ hat eine eindeutig bestimmte Lösung $\tilde{\vec{c}}$, da $\det(Y(x_0) \neq 0$ ist. Mit diesem $\tilde{\vec{c}}$ bildet man

$$\tilde{\vec{y}}(x) := Y(x)\tilde{\vec{c}} = \tilde{c}_1\vec{y}_1(x) + \cdots + \tilde{c}_n\vec{y}_n(x).$$

Dann ist $\tilde{\vec{y}} \in C^1(I, \mathbb{R}^n)$ und eine Lösung von (1.5.2.16) nach dem Superpositionsprinzip. Damit lösen $\tilde{\vec{y}}$ und $\vec{y}_H$ das gleiche Anfangswertproblem mit den Anfangswerten $\vec{y}(x_0) = \vec{y}_0$. Wegen eindeutiger Lösbarkeit dieses Anfangswertproblems muss $\tilde{\vec{y}}(x) = \vec{y}_H(x)$ für alle $x \in I$ sein. $\Rightarrow$ Behauptung (1.5.2.18).

zu c. Aussage folgt wieder aus dem Superpositionsprinzip.

zu d. Differentiation des Ansatzes (1.5.2.19) für $\vec{y}_S$ liefert

$$\vec{y}_S'(x) = Y'(x)\vec{c}(x) + Y(x)\vec{c}\,'(x), \quad \text{Einsetzen in (1.5.2.15): } \underbrace{Y'}_{=AY}\vec{c} + Y\vec{c}\,' = AY\vec{c} + \vec{s}.$$

Damit ist der Funktionenvektor $\vec{y}_S$ genau dann eine Lösung von (1.5.2.15), wenn gilt

$$Y\vec{c}\,' = \vec{s} \;\Leftrightarrow\; \vec{c}\,' = Y^{-1}\vec{s}$$

D.h. $\vec{c} = \vec{c}(x)$ lässt sich durch einfache Integration berechnen $\vec{c}(x) = \int_{x_0}^{x} Y^{-1}(\xi)\vec{s}(\xi)d\xi$, wobei Integration stets komponentenweise zu verstehen ist.

Bemerkung 1.64. 1. Aus Beweis zu d.: $\vec{y}_S(x) = \vec{y}(x)\int_{x_0}^{x} Y^{-1}(\xi)\vec{s}(\xi)d\xi = \int_{x_0}^{x} \vec{y}(x)Y^{-1}(\xi)\vec{s}(\xi)d\xi$.

2. Die Lösungsmenge des homogenen Differentialgleichungssystems (1.5.2.16) bildet ein n-dimensionalen linearen Unterraum $\text{lin}\{\vec{y}_1, \ldots, \vec{y}_n\}$ des Vektorraums $C(I, \mathbb{R}^n)$, die von (1.5.2.15) einen entsprechenden affinen Unterraum.

3. $(n + 1)$ Lösungsvektoren zu (1.5.2.15) sind stets linear abhängig.

Beispiel 1.65. Gesucht: Lösung (y_1, y_2) des Anfangswertproblems

$$\begin{aligned} y_1' &= y_2 + 1 \\ y_2' &= -y_1 \end{aligned} \qquad \text{mit} \quad y_1(0) = 1,\ y_2(0) = 2\ .$$

1. Schritt Finden eines Fundamentalsystems zum homogenen Differentialgleichungssystem

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= -y_1 \end{aligned} : \tag{1.5.2.20}$$

$$y_1'' = y_2' = -y_1 \ \Rightarrow\ 2y_1'y_1'' = -2y_1y_1' \ \Rightarrow\ \frac{d}{dx}(y_1')^2 = -\frac{d}{dx}y_1^2,\ (y_1')^2 = A - y_1^2$$

$$\Rightarrow\ y_1' = \pm\sqrt{A - y_1^2},\ A > 0,\ A = B^2, B \in \mathbb{R}$$

$$\Rightarrow\ \arcsin\left(\frac{y_1}{B}\right) = \int \frac{dy_1}{\sqrt{B^2 - y_1^2}} = \pm\int dx = \pm x + D \ \Rightarrow\ y_1(x) = B\sin(D \pm x)$$

$$\Rightarrow\ y_1(x) = B\sin(D)\cos(x) \pm B\cos(D)\sin(x),\ y_2 = y_1' = -B\sin(D)\sin(x) \pm B\cos(D)\sin(x)$$

$$\Rightarrow\ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = c_1\begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix} + c_2\begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix}$$

für beliebige Konstanten $c_1, c_2 \in \mathbb{R}$

Ein Fundamentalsystem zu (1.5.2.20) ist gegeben durch $\begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix}, \begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix}$, denn es gilt $w(x) = \det\begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} = \cos^2(x) + \sin^2(x) = 1 \neq 0$

2. Schritt Allgemeine Lösung von (1.5.2.20):

$$y_H(x) = c_1\begin{pmatrix} \cos(x) \\ -\sin(x) \end{pmatrix} + c_2\begin{pmatrix} \sin(x) \\ \cos(x) \end{pmatrix}$$

3. Schritt Spezielle Lösung zum inhomogenen Differentialgleichungssystem:

$$\text{Ansatz:}\quad \vec{y}_S(x) = c_1(x)\begin{pmatrix}\cos(x)\\ -\sin(x)\end{pmatrix} + c_2(x)\begin{pmatrix}\sin(x)\\ \cos(x)\end{pmatrix}$$

$$\Rightarrow\ \vec{y}_S'(x) = \begin{pmatrix}c_1'(x)\cos(x) + c_2'(x)\sin(x)\\ c_1'(x)(-\sin(x)) + c_2'(x)\cos(x)\end{pmatrix} + \begin{pmatrix}c_1(x)(-\sin(x)) + c_2(x)\cos(x)\\ c_1(x)(-\cos(x)) + c_2(x)(-\sin(x))\end{pmatrix}$$

$$= \begin{pmatrix}-c_1(x)\sin(x) + c_2(x)\cos(x) + 1\\ -(c_1(x)\cos(x) + c_2(x)\sin(x))\end{pmatrix} \Rightarrow \begin{array}{c} c_1'\cos(x) + c_2'\sin(x) = 1\\ -c_2'\sin(x) + c_2'\cos(x) = 0\\ \text{(linear Gleichungssystem für } c_1', c_2')\end{array}$$

$$\Rightarrow\ c_1' = \cos(x),\ c_2' = \sin(x),\ c_1 = \sin(x),\ c_2 = -\cos(x)\ \Rightarrow\ \vec{y}_S(x) = \begin{pmatrix}0\\ -1\end{pmatrix}$$

4. Schritt Allgemeine Lösung zum Ausgangs-Differentialgleichungssystem:

$$\vec{y}(x) = \begin{pmatrix}y_1(x)\\ y_2(x)\end{pmatrix} = \vec{y}_S(x) + y_H(x)$$

$$\Rightarrow\ y_1(x) = c_1\cos(x) + c_2\sin(x),\ y_2(x) = -c_1\sin(x) + c_2\cos(x) - 1$$

mit beliebigen Konstanten $c_1, c_2 \in \mathbb{R}$.

5. Schritt Einarbeitung der Anfangswerte:

$$1 = c_1\cos(0) + c_2\sin(0)\ \Rightarrow\ c_1 = 1$$
$$2 = -c_1\sin(0) + c_2\cos(0) - 1\ \Rightarrow\ c_2 = 3$$

$\Rightarrow$ Lösung des Anfangswertproblems:

$$\boxed{y_1(x) = \cos(x) + 3\sin(x), \quad y_2(x) = -\sin(x) + 3\cos(x) - 1}$$

Problem: Konstruktion eines Fundamentalsystems für das homogene Differentialgleichungssystem (1.5.2.16). (Dieses existiert gemäß Struktursatz, Teil a.)
Im allgemeinen Fall ist dies jedoch nicht explizit angebbar, es funktioniert aber bei Koeffizientenmatrizen A mit konstanten Einträgen.

1.5.2.2 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten

Ein homogenes lineares Differentialgleichungssystem 1. Ordnung mit konstanten Koeffizienten hat die Gestalt:

$$\vec{y}'(x) = A\vec{y}(x)\ \Leftrightarrow\ \begin{cases} y_1'(x) = a_{11}y_1(x) + \cdots + a_{1n}y_n(x)\\ y_2'(x) = a_{21}y_1(x) + \cdots + a_{2n}y_n(x)\\ \vdots\\ y_1n'(x) = a_{n1}y_1(x) + \cdots + a_{nn}y_n(x)\end{cases} \tag{1.5.2.21}$$

mit reellen konstanten Koeffizienten $a_{jk},\ j,k = 1,\dots,n,\ A = (a_{jk})$.
Durch Anwendung der Eigenwert-Theorie von Matrizen führt der Exponentialansatz für einen Lösungsvektor $\vec{y}$ zu (1.5.2.21)

$$\vec{y}(x) = e^{\lambda x}\vec{v} \tag{1.5.2.22}$$

mit zu bestimmenden Vektor $\vec{v} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ zum Ziel.

Behauptung 1.66. Ein Funktionenvektor vom Typ (1.5.2.22) ist genau dann Lösung des homogenen Differentialgleichungssystems (1.5.2.21), wenn λ ein Eigenwert zur Koeffizientenmatrix A und $\vec{v}$ ein zugehöriger Eigenvektor ist.

Beweis. $\vec{y}'(x) \underset{(1.5.2.22)}{=} \lambda e^{\lambda x}\vec{v}$ in (1.5.2.21) eingesetzt

$$(\lambda\vec{v})e^{\lambda x} = (A\vec{v})e^{\lambda x} \Leftrightarrow A\vec{v} = \lambda\vec{v}$$

Das ist Eigenvektor-Gleichung zum Eigenwert λ von A, d.h. λ ist Nullstelle des charakteristischen Polynoms $P_n(\lambda) = \det(A - \lambda E_n)$ zu A.

Konstruktion eines Fundamentalsystems zu (1.5.2.21):

1. Fall: Matrix A hat n linear unabhängige Eigenvektoren $\vec{v}_1,\dots,\vec{v}_n \in \mathbb{R}^n$ zu Eigenwerten $\lambda_1,\dots,\lambda_n \in \mathbb{R}$
(d.h. A ist diagonalisierbar, z.B. wenn $A = A^T$ (symmetrisch)).

   Dann bilden

$$\vec{y}_1(x) = e^{\lambda_1 x}\vec{v}_1,\dots,\vec{y}_n(x) = e^{\lambda_n x}\vec{v}_n$$

   ein Fundamentalsystem zu (1.5.2.21), denn $\vec{y}_j$ ist Lösung zu (1.5.2.21) nach obiger Behaauptung und

$$W[\vec{y}_1,\dots,\vec{y}_n](0) = \det(\vec{y}_1(0),\dots,\vec{y}_n(0) = \det(\vec{v}_1,\dots,\vec{v}_n) \neq 0$$

   wegen linearer Unabhängigkeit der Eigenvektoren.

2. Fall: Matrix A besitzt n linear unabhängige Eigenvektoren $\vec{v}_1,\dots,\vec{v}_n$ zu eventuell komplexen Eigenwerten $\lambda_1,\dots,\lambda_n \in \mathbb{C}$.

   Dann erhält man wie im ersten Fall ein *komplexes* Fundamentalsystem

$$\vec{y}_j(x) = e^{\lambda_j x}\vec{v}_j, \quad j = 1,\dots,n. \tag{1.5.2.23}$$

   Dabei ist ein komplexer Funktionenvektor $\vec{y}(x) = \vec{a}(x) + i\vec{b}(x)$ mit reellen Funktionenvektoren $\vec{a}(x), \vec{b}(x) \in \mathbb{C}^1(\mathbb{R},\mathbb{R}^n)$ genau dann eine Lösung des reellen Differentialgleichungssystems (1.5.2.21), wenn $\mathrm{Re}(\vec{y}(x)) = \vec{a}(x)$ und $\mathrm{Im}(\vec{y}(x)) = \vec{b}(x)$ Lösungen von (1.5.2.21) sind. Beachtet man außerdem, dass bei der reellen Matrix A neben $\lambda_1 \in \mathbb{C}\setminus\mathbb{R}$ auch $\overline{\lambda_1}$ Eigenwert und $\overline{\vec{v}_1}$ zugehöriger Eigenvektor ist, so ergibt sich aus (1.5.2.23) ein neues Fundamentalsystem

$$\mathrm{Re}\vec{y}_1,\ \mathrm{Im}\vec{y}_1,\ \vec{y}_3,\dots,\vec{y}_n$$

Wiederholte Anwendung dieser Trennung in Real- und Imaginärteil liefert am Ende ein reelles Fundamentalsystem zu (1.5.2.21).

3. Fall: (Allgemeiner Fall) Die Matrix A besitzt weniger als n linear unabhängige Eigenvektoren (d.h. A nicht diagonalisierbar).

Sei λ eine r-fache Nullstelle des charakteristischen Polynoms $P_n(\lambda)$ Dann gibt es r linear unabhängige Lösungen

$$\vec{y}_1 = \vec{p}_0(x)e^{\lambda x}, \ldots, \vec{y}_r = \vec{p}_{r-1}(x)e^{\lambda x} \tag{1.5.2.24}$$

des Differentialgleichungssystems (1.5.2.21), wobei $\vec{p}_l(x)$ ein Vektor aus n (komplexen) Polynomen höchstens l-ten Grades ist. Die Berechnung dieser Polynom-Vektoren erfolgt durch Einsetzen des Ansatzes (1.5.2.24) in (1.5.2.21).

Bemerkung 1.67. Hier ist r die algebraische Vielfachheit des Eigenwerts λ, welche größer als die geometrische Vielfachheit von λ ist.

Wird diese Konstruktion für jeden Eigenwert $\lambda = \lambda_k$ von A mit der algebraischen Vielfachheit r_k, $k = 1, \ldots, m$, $r_1 + \cdots + r_m = n$ ausgeführt, so erhält man n Lösungsvektoren, welche ein (komplexes) Fundamentalsystem bilden. Durch Trennung in Realteil und Imaginärteil dieser Funktionenvektoren entsteht wieder ein reelles Fundamentalsystem zu (1.5.2.21) (analog zu 2. Fall).

Beispiel 1.68. Gesucht ist Fundamentalsystem zu

$$\begin{array}{rcrcrl} y_1' = y_1 & - & 2y_2 & & & \\ y_2' = 2y_1 & & & - & y_3 & \Leftrightarrow \\ y_3' = 4y_1 & - & 2y_2 & - & y_3 & \end{array} \begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & -1 \\ 4 & -2 & -1 \end{pmatrix}}_{=A} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

1. Schritt: Eigenwerte von A:

$$\det(A-\lambda I) = \begin{vmatrix} 1-\lambda & -2 & 0 \\ 2 & -\lambda & -1 \\ 4 & -2 & -1-\lambda \end{vmatrix} = (1-\lambda)(2+\lambda+\lambda^2) = 0 \;\Leftrightarrow\; 2+\lambda+\lambda^2 = 0 \text{ oder } 1-\lambda = 0$$

$$\Leftrightarrow\; \lambda_{1,2} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} - 2} = -\frac{1}{2} \pm \alpha i,\; \alpha = \frac{\sqrt{7}}{2}, \quad \lambda_3 = 1$$

2. Schritt: Eigenvektoren von A:

2. i) $\lambda_1 = -\frac{1}{2} + \alpha i$: Eigenvektor-Gleichung:

$$A\vec{v}_1 = \left(-\frac{1}{2} + \alpha i\right)\vec{v}_1 \;\Leftrightarrow\; \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & -1 \\ 4 & -2 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = (-\frac{1}{2} + \alpha i \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

$$\Rightarrow\ b = \frac{1-\left(-\frac{1}{2}+\alpha i\right)}{2}a = \left(\frac{3}{4}-\frac{\alpha}{2}i\right)a, \quad c = 2a-\left(-\frac{1}{2}+\alpha i\right)b = \left(\frac{3}{2}-\alpha i\right)a.$$

Mit $a = \frac{3}{2}+i\alpha$ folgt $c = \left|\frac{3}{2}+i\alpha\right|^2 = \frac{9}{4}+\alpha^2 = 4$ und $b = \frac{1}{2}\left|\frac{3}{2}+i\alpha\right|^2 = 2$, also $\vec{v}_1 = \begin{pmatrix}\frac{3}{2}+i\alpha\\ 2\\ 4\end{pmatrix}$.

2. ii) $\lambda_2 = \overline{\lambda}_1 = -\frac{1}{2}-\alpha i$, $v_2 = \overline{v}_1 = \begin{pmatrix}\frac{3}{2}-i\alpha\\ 2\\ 4\end{pmatrix}$

2. iii) $\lambda_3 = 1$: Eigenvektor-Gleichung:

$$A\vec{v}_3 = 1\cdot\vec{v}_3 \ \Leftrightarrow\ \begin{pmatrix}1 & -2 & 0\\ 2 & 0 & -1\\ 4 & -2 & -1\end{pmatrix}\begin{pmatrix}a\\ b\\ c\end{pmatrix} = \begin{pmatrix}a\\ b\\ c\end{pmatrix}$$

$$\Rightarrow\ a-2b=a,\ \text{d.h. } b=0, \quad 2a-c=b,\ \text{also } 2a=c \Rightarrow\ \text{Eigenvektor } v_3 = \begin{pmatrix}1\\ 0\\ 2\end{pmatrix}$$

3. Schritt: Komplexes Fundamentalsystem:

$$\vec{y}_1(x) = e^{\lambda_1 x}\vec{v}_1 = e^{-\frac{x}{2}}(\cos(\alpha x)+i\sin(\alpha x))\left(\begin{pmatrix}\frac{3}{2}\\ 2\\ 4\end{pmatrix}+i\begin{pmatrix}\alpha\\ 0\\ 0\end{pmatrix}\right), \quad \vec{y}_2(x) = \overline{\vec{y}_1(x)}, \quad \vec{y}_3(x) = e^{\lambda_3 x}\vec{v}_3 = e^x\begin{pmatrix}1\\ 0\\ 2\end{pmatrix}$$

4. Schritt: Reelles Fundamentalsystem:

$$\mathrm{Re}(\vec{y}_1(x)) = e^{-\frac{x}{2}}\left(\begin{pmatrix}\frac{3}{2}\\ 2\\ 4\end{pmatrix}\cos(\alpha x)-\begin{pmatrix}\alpha\\ 0\\ 0\end{pmatrix}\sin(\alpha x)\right) = \begin{pmatrix}\frac{3}{2}e^{-\frac{x}{2}}(\cos(\alpha x)-\alpha\sin(\alpha x))\\ 2e^{-\frac{x}{2}}\cos(\alpha x)\\ 4e^{-\frac{x}{2}}\cos(\alpha x)\end{pmatrix},$$

$$\mathrm{Im}(\vec{y}_1(x)) = e^{-\frac{x}{2}}\left(\begin{pmatrix}\frac{3}{2}\\ 2\\ 4\end{pmatrix}\sin(\alpha x)+\begin{pmatrix}\alpha\\ 0\\ 0\end{pmatrix}\cos(\alpha x)\right) = \begin{pmatrix}\frac{3}{2}e^{-\frac{x}{2}}(\sin(\alpha x)+\alpha\cos(\alpha x))\\ 2e^{-\frac{x}{2}}\sin(\alpha x)\\ 4e^{-\frac{x}{2}}\sin(\alpha x)\end{pmatrix},$$

$$\vec{y}_3(x) = \begin{pmatrix}e^x\\ 0\\ 2e^x\end{pmatrix}.$$

1.5.2.3 Anwendung auf lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten a_j

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0 \tag{1.5.2.25}$$

Die Differentialgleichung (1.5.2.25) lässt sich äquivalent in ein lineares Differentialgleichungssystem 1. Ordnung umschreiben, indem man $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ mit

$$y_j(x) := y^{(j-1)}(x) \quad \text{für } j = 1, \ldots, n \tag{1.5.2.26}$$

setzt.

$$(1.5.2.25) \Leftrightarrow \begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad \vdots \\ y_{n-1}' = y_n \\ y_n' = -a_{n-1}y_n - \cdots - a_1 y_2 - a_0 y_1 \end{cases}$$

$$\Leftrightarrow \vec{y}\,' = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -a-0 & -a_1 & \ldots & \ldots & -a_{n-1} \end{pmatrix}}_{:=A} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{1.5.2.27}$$

Damit übertragen sich alle Ergebnisse aus Abschnitt 1.5.2.2 auf (1.5.2.25), d.h. insbesondere Superpositionsprinzip, Strukturaussagen zu den Lösungen der homogenen und inhomogenen Differentialgleichung sowie Existenz- und Eindeutigkeitsaussagen. Hier lauten wichtige Begriffe wie folgt:

i) p Funktionenvektoren $\vec{u}_1, ..., \vec{u}_p$ vom Typ (1.5.2.26) (also $u_{kj} = u_{k1}^{(j-1)}, k = 1, ..., p, j = 1, ..., n$) sind linear unabhängig (im Vektorraum $C(I, \mathbb{C}^n)$ genau dann, wenn die ersten Komponenten dieser Vektoren, d.h. die Funktionen $u_1, ..., u_p$ linear unabhängig (im Vektorraum $C(I, \mathbb{C})$) sind. (Beweis: Übung)

ii) n linear unabhängige Lösungen $u_1, ..., u_n$ der homogenen Differentialgleichung (1.5.2.25) bilden wieder ein Fundamentalsystem (oder auch Integralbasis oder Hauptsystem) zu (1.5.2.25)

Zur Bestimmung eines Fundamentalsystems gehen wir zurück zum äquivalenten Differentialgleichungssystem 1. Ordnung (1.5.2.27) und ermitteln die Eigenwerte der zugehörigen Koeffizientenmatrix A, d.h. die Nullstellen des charakteristischen Polynoms $P_n(\lambda) = \det(A - \lambda I)$

$$P_n(\lambda) = \begin{vmatrix} -\lambda & 1 & 0 & \ldots & 0 \\ 0 & -\lambda & 1 & \ldots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \ldots & 1 \\ -a-0 & -a_1 & \ldots & \ldots & -a_{n-1}-\lambda \end{vmatrix} = (-1)^n(a_0 + a_1\lambda + \cdots + a_{n-1}\lambda^{n-1} + \lambda^n)$$

Bemerkung 1.69. Oft bezeichnet man $\tilde{P}_n(\lambda)$ als charakteristisches Polynom zur Differentialgleichung (1.5.2.25). Die Nullstellen von P_n und $\tilde{P}_n$ sind dieselben.
Formal entsteht $\tilde{P}_n(\lambda)$ durch Ersetzung von $y^{(j)}$ durch λ^j für $j = 0, \ldots, n$. Setzt man $D := \frac{d}{dx}$, so gilt: $Dy = \frac{dy}{dx} = y', \quad D^j y = y^{(j)}$ und

$$(1.5.2.25) \;\Leftrightarrow\; \tilde{P}_n(D)[y] = 0$$

Zu einer r-fachen NST von $\tilde{P}_n$ gibt es nach Übergang zum äquivalenten Differentialgleichungssystem (1.5.2.27) r linear unabhängige Lösungen $\vec{y}_1, \ldots \vec{y}_r, \vec{y}_l = (y_{l1}, y'_{l1}, \ldots, y^{(n-1)}_{l1})$ mit den ersten Komponente:

$$y_1(x) = p_{01}(x)e^{\lambda x}, y_2(x) = p_{11}(x)e^{\lambda x}, \ldots, y_r(x) = p_{r-1,1}(x)e^{\lambda x} \tag{1.5.2.28}$$

mit $p_{l1}(x)$ als Polynom höchstens l-ten Grades in x.

$$y_l \in \left\{ f : \mathbb{R} \to \mathbb{C} \,\middle|\, f \in C(\mathbb{R}, \mathbb{C}), f(x) = \sum_{\nu=0}^{r-1} \alpha_\nu x^\nu e^{\lambda x} \right\} =: \mathcal{P}_{r-1}$$

Behauptung 1.70. $\mathcal{P}_{r-1}$ ist ein r-dimensionaler Unterraum des Vektorraum $C(\mathbb{R}, \mathbb{C})$ und hat die Basis

$$u_1(x) = x^0 e^\lambda, u_2(x) = x^1 e^{\lambda x}, \ldots, u_r(x) \cdot x^{r-1} e^{\lambda x} \tag{1.5.2.29}$$

Beweis. Übungsaufgabe

Folgerung 1.71. i) Die Funktionen $y_1, ..., y_r$ aus (1.5.2.28) sind linear unabhängige Funktionen aus $\mathcal{P}_{r-1}$ und bilden folglich auch eine Basis von $\mathcal{P}_{r-1}$. Somit kann jede Funktion aus $\mathcal{P}_{r-1}$ als Linearkombination der $y_1, ..., y_r$ dargestellt werden und ist nach Superpositionsprinzip eine Lösung von (1.5.2.25).

ii) Die Differentialgleichung (1.5.2.25) besitzt r linear unabhängige Lösungen vom Typ (1.5.2.29) (unter der Voraussetzung das λ r-fache Nullstelle von $\tilde{P}_n$ ist.)

iii) Werden zu n (komplexen) Nullstellen von $\tilde{P}_n$ gezählt mit ihrer Vielfachheit, gemäß ii) die entsprechenden Funktionen u_j konstruiert, so erhält man insgesamt ein komplexes Fundamentalsystem zu (1.5.2.25). Bei reellen Koeffizienten a_j erhält man daraus ein reelles Fundamentalsystem indem man $Re(u_j)$ und $Im(u_j)$ anstelle der komplexen u_j und $\overline{u_j}$ setzt

Beispiel 1.72. $y^{(5)} + 4y^{(4)} + 2y''' - 4y'' + 8y' + 16y = 0$
gesucht wird ein reelles Fundamentalsystem.
1.Schritt:
Nullstellen des charakteristischen Polynoms:
$\tilde{P}_5(\lambda) = \underbrace{\lambda^5 + 4\lambda^4 + 2\lambda^3 - 4\lambda^2 + 8\lambda + 16}_{(\lambda+2)^3(\lambda^2-2\lambda+2)} = 0$
$\lambda_{1/2/3} = -2, \lambda_{4/5} = 1 \pm \sqrt{1-2} = 1 \pm i$

2.Schritt:
komplexes Fundamentalsystem
$e^{-2x}, xe^{-2x}, x^2e^{-2x}, \underbrace{e^{(1+i)x}}_{e^x(\cos x + i \sin x)}, e^{(1-i)x}$

3.Schritt:
reelles Fundamentalsystem: $e^{-2x}, x, ^{-2x}, x^2e^{-2x}, e^x \cos x, e^x \sin x$
Allgemeine Lösung: $y(x) = (c_1 + c_2x + c_3x^2)e^{-2x} + (c_4 \cos x + c_5 \cos x)e^x, \quad c_j \in \mathbb{R}$

Kapitel 2

Differentialrechnung im $\mathbb{R}^N$

2.1 Differenzierbarkeit und Ableitung

2.1.1 Partielle Ableitung

2.1.1.1 Definition

Im Folgenden sei $\Omega \subseteq \mathbb{R}^N$ ein Gebiet. Wir betrachten die reellwertige oder skalare Funktion $f : \Omega \to \mathbb{R}$ und definieren die partiellen Ableitungen von f in $p \in \Omega$: Hält man alle Variablen bis auf x_1 fest, also $x_j = p_j$ für $j = 2, \ldots, N$ und setzt man $f_1(x_1) := f(x_1, p_2, \ldots, p_N)$, dann ist

$$f_1'(p) =: \frac{\partial f}{\partial x_1}(p)$$

Analog definiert man $\frac{\partial f}{\partial x_2}(p), \ldots, \frac{\partial f}{\partial x_N}(p)$, also

$$\frac{\partial f}{\partial x_i}(p) = f_{x_i}(p) := \lim_{h \to 0} \frac{f(p + he_i) - f(p)}{h} \qquad (\textit{partielle Ableitung} \text{ von } f \text{ nach } x_i)$$

für $i = 1, \ldots, N$, $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, 1 an i-ter Stelle.

Beispiel 2.1. i. $f(x_1, x_2) = \sin(x_1 x_2^2)$ ist definiert auf $\mathbb{R}^2$.

$$\frac{\partial f}{\partial x_1} = x_2^2 \cos(x_1 x_2^2)\,, \quad \frac{\partial f}{\partial x_2} = 2x_1 x_2 \cos(x_1 x_2^2)\,.$$

ii. $f(x, y, z, t) = \mathrm{e}^x + xyz + z^3 y$.

$$f_x = \mathrm{e}^x + yz\,, \quad f_y = xz + z^3\,, \quad f_z = xy + 3yz^2\,, \quad f_t = 0\,.$$

2.1.1.2 Geometrische Deutung

$\frac{\partial f}{\partial x_i}(p)$ liefert den Anstieg der Fläche $y = f(x)$ in x_i-Richtung, d.h. den Anstieg der *Tangentialebene* T an diese Fläche im Punkt $(p, f(p))$. Im $\mathbb{R}^2$ lautet die Gleichung von T:

$$y = f(p) + (x_1 - p_1)\frac{\partial f}{\partial x_i}(p) + (x_2 - p_2)\frac{\partial f}{\partial x_2}(p)$$

beziehungsweise allgemein:

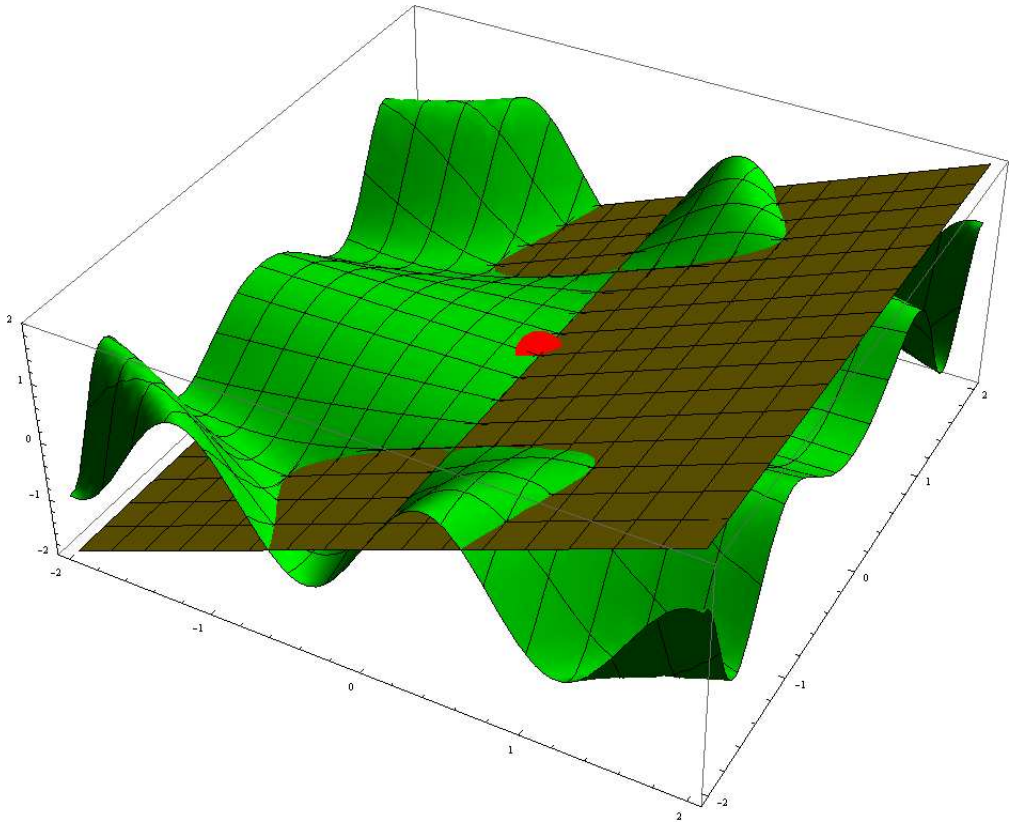
$$T : y = f(p) + \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(p)(x_i - p_i) = f(p) + \langle \text{grad } f(p), x - p \rangle$$

wobei der Vektor grad $f(p) := (f_{x_1}(p), \ldots, f_{x_N}(p))$ *Gradient* von f in p heißt und mit grad $f = \nabla f$ (*Nabla*[I] f) bezeichnet wird.

Beispiel 2.2. Tangentialebene an $y = f(x_1, x_2) = \sin(x_1 x_2^2)$ in $p = (0, 1)$.

$$\text{grad } f(0,1) = (f_{x_1}(0,1), f_{x_2}(0,1)) = (1, 0)$$

$$T : y = f(0,1) + \langle \text{grad } (0,1), x - (0,1) \rangle = 0 + 1(x_1 - 0) + 0(x_2 - 1) = x_1$$



[I] ∇ - Symbol für hebräisches Saiteninstrument

2.1.2 Partielle Ableitungen höherer Ordnung

Höhere partielle Ableitungen von f in p ergeben sich wie im $\mathbb{R}^1$:

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \left(f_{x_j}\right)_{x_i} = f_{x_j x_i}\,, \quad \frac{\partial^2 f}{\partial x_i^2} = f_{x_i x_i}$$

für $i, j = 1, \ldots, N$ *(partielle Ableitung 2. Ordnung)*.

Beispiel 2.3. (Fortsetzung von Beispiel 2.1.i.)

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_1^2} &= \frac{\partial f}{\partial x_1}(x_2^2 \cos(x_1 x_2^2)) = -x_2^4 \sin(x_1 x_2^2) \\
\frac{\partial^2 f}{\partial x_2 x_1} &= \frac{\partial}{\partial x_2}(x_2^2 \cos(x_1 x_2^2)) = 2x_2 \cos(x_1 x_2^2) - 2x_1 x_2^3 \sin(x_1 x_2^2) \\
\frac{\partial^2 f}{\partial x_1 x_2} &= \frac{\partial}{\partial x_1}(2x_1 x_2 \cos(x_1 x_2^2)) = 2x_2 \cos(x_1 x_2^2) - 2x_1 x_2^3 \sin(x_1 x_2^2) \\
\frac{\partial^2 f}{\partial x_2^2} &= \frac{\partial f}{\partial x_2}(2x_1 x_2 \cos(x_1 x_2^2)) = 2x_1 \cos(x_1 x_2^2) - 4x_1^2 x_2^2 \sin(x_1 x_2^2)
\end{aligned}$$

Bezüglich der Differentiationsreihenfolge gilt der Satz von Schwarz:

Satz 2.4. *(Satz von Schwarz)*
Sind alle partiellen Ableitungen 2. Ordnung für f in Ω vorhanden und in $p \in \Omega$ stetig, so gilt

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(p) = \frac{\partial^2 f}{\partial x_j \partial x_i}(p) \qquad \text{für } i, j = 1, \ldots, N.$$

Beweis. (nur für $N = 2$, $p = (0,0)$, $x_1 = x, x_2 = y$)
Seien f_{xy}, f_{yx} in $\Omega = (-r, r) \times (-r, r)$ vorhanden und in $(0,0) \in \Omega$ stetig.

(i). Anwendung des Mittelwertsatzes auf $\phi(x) := f(x,y) - f(x,0)$ für *festes* y, $(x,y) \in \Omega$.

$$\begin{aligned}
&\phi(x) - \phi(0) = \phi'(\xi)(x-0) \qquad \text{mit einem } \xi \in (-r,r),\ |\xi| < |x| \\
\Rightarrow\ & F(x,y) := (f(x,y) - f(x,0)) - (f(0,y) - f(0,0)) \\
&= \phi(x) - \phi(0) = \phi'(\xi)x = \underbrace{(f_x(\xi,y) - f_x(\xi,0))}_{=\psi(y)-\psi(0)} \cdot x\,,
\end{aligned}$$

wobei $\psi(y) := f_x(\xi, y)$ mit festem ξ ist. Anwendung des Mittelwertsatzes auf ψ

$$\begin{aligned}
&\psi(y) - \psi(0) = \psi'(\eta)(y-0) \qquad \text{mit einem } \eta \in (-r,r),\ |\eta| < |y| \\
\Rightarrow\ & F(x,y) = f_{xy}(\xi,\eta)xy\,.
\end{aligned}$$

(ii). Analog ergibt sich aus $F(x,y) = (f(x,y) - f(0,y)) - (f(x,0) - f(0,0))$ die Darstellung

$$F(x,y) = f_{yx}(\tilde{\xi},\tilde{\eta})xy \qquad \text{mit } \left|\tilde{\xi}\right| < |x|\,, |\tilde{\eta}| < |y|$$
$$\Rightarrow\ f_{xy}(\xi,\eta) = f_{yx}(\tilde{\xi},\tilde{\eta})\ .$$

Wegen Stetigkeit von f_{xy} und f_{yx} in $(0,0)$ folgt für $(x,y) \to (0,0)$:
$\xi, \tilde{\xi}, \eta, \tilde{\eta} \to 0$ und somit $f_{xy}(0,0) = f_{yx}(0,0)$.

Definition 2.5. Mit $C^m(\Omega)$ bezeichnen wir die Menge aller m-mal stetig differenzierbaren reellwertigen Funktionen, das heißt die Menge aller $f : \Omega \to \mathbb{R}$, für die alle partiellen Ableitungen bis zur Ordnung m existieren und stetig sind.

$$C^\infty(\Omega) = \bigcap_{m=0}^{\infty} C^m(\Omega) \quad (C^0(\Omega) := C(\Omega)).$$

Satz 2.6. *(Verallgemeinerter Satz von Schwarz)*
Für $f \in C^m(\Omega)$ sind die partiellen Ableitungen bis Ordnung m unabhängig von der Differentiationsreihenfolge.

2.1.2.1 Partielle Differenzierbarkeit und Stetigkeit

Der Zusammenhang zwischen partieller Differenzierbarkeit von f und Stetigkeit von f ist im $\mathbb{R}^N, N \geq 2$ anders als im $\mathbb{R}^1$.

Beispiel 2.7.

$$f(x_1,x_2) = \begin{cases} 0 & x_1 = 0 \\ 0 & x_2 = 0 \\ \mathrm{e}^{x_1} & x_1 \neq 0, x_2 \neq 0 \end{cases}$$

f ist unstetig in $p = (0,0)$, denn für $(x_n) = ((x_{1n}, x_{2n}))$ mit $\lim\limits_{n\to\infty} x_n = 0$ und $x_{1n}x_{2n} \neq 0$ ist $f(x_n) = \mathrm{e}^{x_{1n}} \overset{n\to\infty}{\longrightarrow} \mathrm{e}^0 = 1$, aber $f(0,0) = 0$.
Jedoch: $f_{x_1}(0,0) = \frac{d}{dx_1} f(x_1,0) = 0$, analog $f_{x_2}(0,0) = 0$.

Wie obiges Beispiel zeigt, liefert die Existenz der partiellen Ableitung von f in $p \in \Omega$ keine Information über das Änderungsverhalten von f in $U_\varepsilon(p)$. Wir benötigen dazu eine „stärkere“ Eigenschaft von f und orientieren uns an reellwertigen Funktionen f.

$$f(x+h) = f(x) + f'(x)h + R(h) \qquad \textit{(Zerlegungsformel)} \tag{2.1.2.1}$$

mit $\lim\limits_{h\to 0} \frac{R(h)}{h} = 0$, falls f differenzierbar in $x \in \mathbb{R}$ ist.

2.1.3 Totale Ableitung

2.1.3.1 Definition und Beispiel

Definition 2.8. Sei $f : \Omega \to \mathbb{R}$ und Ω ein Gebiet des $\mathbb{R}^N$. Dann heißt f *(total) differenzierbar* in $p \in \Omega$, falls die Zerlegungsformel (2.1.2.1)

$$f(p+h) = f(p) + \underbrace{\langle a, h\rangle}_{=a^T h} + R(h)$$

mit einem Vektor $a \in \mathbb{R}^N$ und $R : U_r(0) \to \mathbb{R}, \lim\limits_{h\to 0} \frac{R(h)}{\|h\|} = 0$ gilt.
Wir nennen $a := f'(p)$ *(totale) Ableitung* von f in p.

Beispiel 2.9. Sei eine quadratische Form $Q(x)$ auf $\mathbb{R}^N$ gegeben durch $Q(x) = x^T Ax$ mit $A \in \mathbb{R}^{N\times N}, A^T = A, A = (a_{ij})$. Gesucht ist $Q'(p)$. Gemäß (2.1.2.1) schreiben wir

$$Q(p+h) = (p+h)^T A(p+h) = p^T Ap + p^T Ah + \underbrace{h^T Ap}_{=(h^T Ap)^T = p^T A^T h} + h^T Ah$$

$$= Q(p) + \underbrace{2p^T Ah}_{=a^T h} + \underbrace{h^T Ah}_{=R(h)} ,$$

$$|R(h)| = \left|\sum_{i,j=1}^{N} a_{ij} h_i h_j\right| \le \underbrace{\left(\sum_{i,j=1}^{N} |a_{ij}|^2\right)^{\frac{1}{2}}}_{=:\|A\|_2} \left(\sum_{i,j=1}^{N} |h_i h_j|^2\right)^{\frac{1}{2}}$$

$$\le \|A\|_2 \left(\sum_{i=1}^{N} |h_i|^2 \sum_{j=1}^{N} |h_j|^2\right)^{\frac{1}{2}} = \|A\|_2 \, \|h\|^2 ,$$

$$\frac{|R(h)|}{\|h\|} \le \|A\|_2 \, \|h\| \;\Rightarrow\; \lim_{n\to 0} \frac{R(h)}{\|h\|} = 0 \;\Rightarrow\; Q'(p) = \underline{\underline{2Ap}} .$$

2.1.3.2 Folgerungen aus der totalen Differenzierbarkeit

Satz 2.10. Ist f (total) differenzierbar in p, so ist f auch stetig in p.

Beweis. Sei $(h_n) \subset U_r(0)$ mit $\lim\limits_{n\to\infty} h_n = 0$. Aus Definition (2.8) folgt

$$\langle a, h_n\rangle \to 0 , \quad R(h_n) = \frac{R(h_n)}{\|h_n\|} \|h_n\| \to 0 ,$$

also $\lim\limits_{n\to\infty} f(p+h_n) = f(p) + 0 = f(p)$.

Satz 2.11. Aus der totalen Differenzierbarkeit von f in p folgt die partielle Differenzierbarkeit in p und

$$a = f'(p) = \mathrm{grad} f(p) = (f_{x_1}(p), \ldots, f_{x_N}(p)) .$$

Beweis. Untersuche

$$\frac{f(p+te_j)-f(p)}{t} \underset{(2.1.2.1)}{=} \frac{1}{t}\Big(\underbrace{\langle a, te_j\rangle}_{=ta_j} + R(te_j)\Big) = a_j + \frac{R(te_j)}{t\left\|e_j\right\|}$$

Damit existiert $$\underbrace{\lim_{t\to 0}\frac{f(p+te_j)-f(p)}{t}}_{=\frac{\partial f}{\partial x_j}(p)} = a_j + \underbrace{\lim_{t\to 0}\frac{R(te_j)}{t\left\|e_j\right\|}}_{=0} .$$

Satz 2.12. *(Hauptkriterium der Differenzierbarkeit)*
Ist $f \in C^1(\Omega)$, so ist f (total) differenzierbar in Ω.

Beweis. (nur für $N = 2$) Sei $(0,0) \in \Omega$, so existiert $U_r(0,0) \subset \Omega$. Sei $(x,y) \in U_r(0,0)$.

$$f(x,y) - f(0,0) = \underbrace{(f(x,y) - f(x,0))}_{=f_y(x,\theta_2 y)y} + \underbrace{(f(x,0) - f(0,0))}_{=f_x(\theta_1 x,0)x}$$

unter Anwendung des Mittelwertsatzes auf $\psi(y) := f(x,y)$ für festes x und auf $\phi(x) := f(x,0)$ mit $\theta_1, \theta_2 \in (0,1)$.

$$\begin{aligned} f(x,y) - f(0,0) &= xf_x(\theta_1 x, 0) + yf_y(x, \theta_2 y) \\ &= xf_x(0,0) + yf_y(0,0) + x\underbrace{(f_x(\theta_1 x, 0) - f_x(0,0))}_{=R_1(x,y)} + y\underbrace{(f_y(x_1, \theta_2 y) - f_y(0,0))}_{=R_2(x,y)} . \end{aligned}$$

Wegen Stetigkeit von f_x und f_y ist $\lim\limits_{(x,y)\to(0,0)} R_l(x,y) = 0$ für $l = 1, 2$.
Mit $R(x,y) = xR_1(x,y) + yR_2(x,y)$ gilt

$$\frac{|R|}{\|(x,y)\|} = \frac{|R(x,y)|}{\sqrt{x^2+y^2}} \le \underbrace{\frac{|x|}{\sqrt{x^2+y^2}}}_{\le 1}|R_1(x,y)| + \underbrace{\frac{|y|}{\sqrt{x^2+y^2}}}_{\le 1}|R_2(x,y)| \overset{(x,y)\to(0,0)}{\longrightarrow} 0 .$$

Bemerkung 2.13. Fasst man den Vektor $f'(p) = \operatorname{grad} f(p) = \nabla f(p)$ als Matrix aus $\mathbb{R}^{1\times N}$ auf, so ist diese eine Darstellungsmatrix einer linearen Abbildung $\mathrm{d}f(p) : \mathbb{R}^N \to \mathbb{R}^1$, definiert durch

$$\mathrm{d}f(p)h := \langle f'(p), h\rangle = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(p)h_i .$$

Speziell:

$$\mathrm{d}x_j(p)h = \Big\langle \operatorname{grad} x_j(p), h\Big\rangle = \Big\langle e_j, h\Big\rangle = h_j$$

$$\Rightarrow \mathrm{d}f(p)h = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(p)\mathrm{d}x_j(p)h \quad \text{für alle } h \in \mathbb{R}^N$$

oder kurz:

$$\mathrm{d}f = \sum_{j=1}^{N} f_{x_j}\mathrm{d}x_j \quad (\textit{totales Differential} \text{ der Funktion } f).$$

Beispiel 2.14. Für $f(x_1, x_2) = \sin(x_1 x_2^2)$ ist

$$\begin{aligned} \mathrm{d}f = f_{x_1} dx_1 + f_{x_2} dx_2 &= x_2^2 \cos(x_1 x_2^2) dx_1 + 2x_1 x_2 \cos(x_1 x_2^2) dx_2 \\ &= x_2 \cos(x_1 x_2^2)(x_2 dx_1 + 2x_1 dx_2) \,. \end{aligned}$$

2.1.4 Differentiation von Vektorfunktionen

Wir betrachten die Abbildung $f : \Omega \to \mathbb{R}^M$ mit $\Omega \subseteq \mathbb{R}^N$ (Gebiet).

$$\Omega \ni x = (x_1, \ldots, x_N) \mapsto f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_M(x) \end{pmatrix} = f_1(p)e_1 + \ldots + f_M(x)e_M$$

Definition 2.15. f heißt *differenzierbar* in $p \in \Omega$, falls jede Koordinatenfunktion $f_i, i = 1, \ldots, M$ (total) differenzierbar in p ist.

Satz 2.16. $f : \Omega \to \mathbb{R}^M$ ist genau dann in $p \in \Omega$ differenzierbar, falls es eine lineare Abbildung $A : \mathbb{R}^N \to \mathbb{R}^M$, das heißt $A \in \mathbb{R}^{M \times N}$ und $R : U_r(0) \to \mathbb{R}^M$, gibt mit

$$f(p + h) = f(p) + Ah + R(h) \,, \text{ wobei } \lim_{h \to 0} \frac{R(h)}{||h||} = 0$$

gilt. $A =: f'(p)$ heißt *Ableitung* von f in p, und es ist

$$A = \left(\frac{\partial f_i}{\partial x_j}(p) \right)_{i=1,\ldots,M, j=1,\ldots,N}$$

die *Jacobi-Matrix* von f.

Beweis. f ist differenzierbar in p genau dann, wenn $f_i(p + h) = f_i(p) + \langle a_i, h \rangle + R_i(h)$ für $i = 1, \ldots, M$. Dabei ist

$$a_i = \operatorname{grad} f_i(p) \text{ und } \lim_{h \to 0} \frac{R_i(h)}{||h||} = 0$$

$$\Leftrightarrow \; f(p + h) = \begin{pmatrix} f_1(p + h) \\ \vdots \\ f_M(p + h) \end{pmatrix} = f(p) + \begin{pmatrix} a_1^T h \\ \vdots \\ a_M^T h \end{pmatrix} + \begin{pmatrix} R_1(h) \\ \vdots \\ R_M(h) \end{pmatrix} = f(p) + \begin{pmatrix} a_1^T \\ \vdots \\ a_M^T \end{pmatrix} h + R(h).$$

Somit ist

$$A = \begin{pmatrix} a_1^T \\ \vdots \\ a_M^T \end{pmatrix} = \left(\frac{\partial f_i}{\partial x_j}(p) \right) \text{ und } \lim_{h \to 0} \frac{R(h)}{||h||} = 0 \,.$$

Folgerung 2.17. Wegen der Äquivalenz der Differenzierbarkeit von f mit der koordinatenweise Differenzierbarkeit übertragen sich die Sätze und das Hauptkriterium von reellwertigen Funktionen auf vektorwertige Funktionen.

Beispiel 2.18.

$$f(x_1, x_2) = \begin{pmatrix} x_1 e^{x_2} \\ \ln(1 + x_1^2 x_2^2) \\ x_2 \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \\ f_3(x_1, x_2) \end{pmatrix}$$

$$f'(x) = \begin{pmatrix} f_{1_{x_1}} & f_{1_{x_2}} \\ f_{2_{x_2}} & f_{2_{x_2}} \\ f_{3_{x_1}} & f_{3_{x_2}} \end{pmatrix} = \begin{pmatrix} e^{x_2} & x_1 e^{x_2} \\ \frac{2x_1 x_2^2}{1+x_1^2 x_2^2} & \frac{2x_2 x_1^2}{1+x_1^2 x_2^2} \\ 0 & 1 \end{pmatrix} \quad \text{für alle } x \in \mathbb{R}^2.$$

Wir schreiben $f \in C^m(\Omega, \mathbb{R}^M)$, falls jede Komponente $f_i \in C^m(\Omega)$ ist für $i = 1, \ldots, M$. f heißt dann eine C^m-Abbildung auf Ω.

2.2 Die Kettenregel

2.2.1 Herleitung

Sei $f : \Omega \to \mathbb{R}^M$ mit $\Omega \subseteq \mathbb{R}^N$ (Gebiet), f differenzierbar in $p \in \Omega$, $g : \Sigma \to \mathbb{R}^N$, $\Sigma \subseteq \mathbb{R}^L$ (Gebiet), $g(\Sigma) \subseteq \Omega$, g differenzierbar in $q \in \Sigma$, $g(q) = p$.

$$\begin{array}{ccccc} \Sigma & \xrightarrow{g} & \Omega & \xrightarrow{f} & \mathbb{R}^M \\ \cup & & \cup & & \cup \\ u & \mapsto & x = g(u) & \mapsto & f(g(u)) = (f \circ g)(u) \end{array}$$

Behauptung 2.19. Dann ist $f \circ g$ in q differenzierbar mit

$$(f \circ g)'(q) = f'(g(q)) \cdot g'(q) .$$

Ausführlich:

$$\begin{aligned} (f \circ g)' = (f' \circ g)g' \;&\Leftrightarrow\; \left(\frac{\partial f_i(g(q))}{\partial q_k}\right) = \left(\frac{\partial f_i}{\partial x_j}(g(q))\right) \cdot \left(\frac{\partial g_i(q)}{\partial q_k}\right) \\ &\Leftrightarrow\; \frac{\partial f_i(g(q))}{\partial q_k} = \sum_{j=1}^{N} \frac{\partial f_i}{\partial x_j} g(q) \frac{\partial g_j}{\partial q_k}(q) \end{aligned}$$

für $i = 1, \ldots, M, k = 1, \ldots, L$.

Beweis. (i). Für g gilt nach Vorraussetzung die Zerlegungsformel (2.1.2.1), d.h.

$$g(q + \eta) = g(q) + B\eta + S(\eta) \;\text{ mit }\; \lim_{\eta \to 0} \frac{S(\eta)}{\|\eta\|} = 0, \; B = g'(q),$$

$$\begin{aligned} \Rightarrow\; (f \circ g)(q + \eta) &= f(g(q + \eta)) = f(\underbrace{g(q)}_{=p} + \underbrace{B\eta + S(\eta)}_{=h}) \\ &\underset{(2.1.2.1)}{=} f(p) + A(B\eta + S(\eta)) + R((B\eta) + S(\eta)) \\ &= (f \circ g)(q) + AB\eta + T(\eta) . \end{aligned}$$

(ii). Wir müssen zeigen: $\lim_{\eta\to 0} \frac{T(\eta)}{\|\eta\|} = 0$ für $T(\eta) = AS(\eta) + R(B\eta + S(\eta))$.
Sei $\eta_n \in U_r(0) \subset \mathbb{R}^L$ mit $\lim_{n\to\infty} \eta_n = 0$, so

$$\frac{1}{\|\eta_n\|} AS(\eta_n) = A\frac{S(\eta_n)}{\|\eta_n\|} \to 0 \quad \text{für } n \to \infty,$$

$$\frac{1}{\|\eta_n\|} R(B\eta_n + S(\eta_n)) = \underbrace{\frac{R(B\eta_n + S(\eta_n))}{\|B\eta_n + S(\eta_n)\|}}_{\to 0} \frac{\|B\eta_n + S(\eta_n)\|}{\|\eta_n\|} .$$

Weiterhin ist

$$\frac{1}{\|\eta_n\|}\|B\eta_n + S(\eta_n)\| \le \frac{1}{\|\eta_n\|}(\|B\eta_n\| + \|S(\eta_n)\|) \le \|B\|_2 + \frac{\|S(\eta_n)\|}{\|\eta_n\|} ,$$

also

$$\frac{1}{\|\eta_n\|}\|B\eta_n + S(\eta_n)\| \le \|B\|_2 + 1 \quad \text{für } n \ge n_0(1).$$

(iii). Ergebnis:

$$\lim_{n\to\infty} \frac{1}{\|\eta_n\|} T(\eta_n) = 0$$

und folglich

$$(f \circ g)'(q) = AB \quad \text{mit } A = f'(p) = f'(g(q)) .$$

2.2.2 Anwendung auf skalare Funktionen

2.2.2.1 Summenregel, Produktregel, Quotientenregel

Produktregel für differenzierbare skalare Funktionen $f_1, f_2 : \Omega \to \mathbb{R}$:

$$\begin{aligned}(f_1 f_2)' = f_1 f_2' + f_1' f_2 \;&\Leftrightarrow\; \text{grad}\,(f_1 f_2) = f_1 \text{grad}\, f_2 + f_2 \text{grad}\, f_1 \\ &\Leftrightarrow\; \nabla(f_1 f_2) = f_1(\nabla f_2) + f_2(\nabla f_1) .\end{aligned}$$

Beweis. $G(u_1, u_2) := u_1 u_2$, $G : \mathbb{R}^2 \to \mathbb{R} \;\Rightarrow\; G(f_1, f_2) = f_1 f_2$. Die verkette Funktion $G \circ (f_1, f_2)$ hat die Ableitung $(G \circ (f_1, f_2))' = G'(f_1, f_2) \cdot (f_1, f_2)'$. Dabei ist

$$G'(u_1, u_2) = (G_{u_1}, G_{u_2}) = (u_2, u_1) \;\Rightarrow\; (G \circ (f_1, f_2))' = (f_2, f_1)\begin{pmatrix} f_1' \\ f_2' \end{pmatrix} = f_2 f_1' + f_1 f_2' .$$

Analog beweist man Summen- und Quotientenregel.

2.2.2.2 Spezialfälle der Kettenregel

(a). $M = N = L = 1$: Kettenregel für reelle Funktionen

(b). $M = L = 1$, das heißt $g : I \to g(I) \subseteq \Omega \subseteq \mathbb{R}^N$, f ist eine skalare Funktion $f : \Omega \to \mathbb{R}$, $I = (a, b) \subset \mathbb{R}$. Betrachte die neue Funktion $\phi(t) := f(g(t))$

$$\phi'(t) = \frac{\mathrm{d}f(g(t))}{\mathrm{d}t} = \langle \operatorname{grad} f(g(t)), g'(t) \rangle = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(g(t)) \cdot \frac{\mathrm{d}g_j}{\mathrm{d}t} .$$

Dabei ist

$$g'(t) = \begin{pmatrix} g_1'(t) \\ \vdots \\ g_n'(t) \end{pmatrix} = \lim_{\tau \to 0} \frac{g(t+\tau) - g(t)}{\tau}$$

der *Tangentenvektor* im Punkt $p = g(t)$ an die Kurve $x = g(t), t \in I$.

Bemerkung 2.20. Wenn $x = g(t)$ die Trajektorie oder Bahnkurve eines Massenpunktes zur Zeit t ist, so ist $g'(t)$ der Geschwindigkeitsvektor des Massenpunktes. Bezeichnung: $g'(t) = \dot{g}(t)$.

(c). $M = 1$. $\phi(u) := f(g(u)), u \in \Sigma$.

$$\frac{\partial \phi}{\partial u_k} = \left\langle \operatorname{grad} f(g(u)), \frac{\partial g}{\partial u_k} \right\rangle = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(g(u)) \cdot \frac{\partial g_j}{\partial u_k} , \quad \frac{\partial g}{\partial u_k} = \left(\frac{\partial g_1}{\partial u_k}, \ldots, \frac{\partial g_N}{\partial u_k} \right)$$

Beispiel 2.21. Einführung von *Polarkoordinaten* im $\mathbb{R}^2$. Sei $f \in C^1(\mathbb{R}^2)$.

$$(r, \phi) \mapsto x = (r\cos\phi, r\sin\phi) \mapsto f(x) = f(r\cos\phi, r\sin\phi) = \tilde{f}(r, \phi)$$

mit $r \in (0, \infty), \phi \in (0, 2\pi)$.

$$\frac{\partial \tilde{f}}{\partial r} = \left\langle \operatorname{grad} f, \left(\frac{\partial x_1}{\partial r}, \frac{\partial x_2}{\partial r} \right) \right\rangle = \langle \operatorname{grad} f, (\cos\phi, \sin\phi) \rangle = \left\langle \operatorname{grad} f, \frac{x}{r} \right\rangle ,$$
$$\frac{\partial \tilde{f}}{\partial \phi} = \left\langle \operatorname{grad} f, \left(\frac{\partial x_1}{\partial \phi}, \frac{\partial x_2}{\partial \phi} \right) \right\rangle = \left\langle \operatorname{grad} f, \underbrace{(-r\sin\phi, r\cos\phi)}_{=x^\perp} \right\rangle$$

mit $\|x^\perp\| = \|x\|$, $\langle x^\perp, x \rangle = 0$.

2.2.2.3 Richtungsableitung

Definition 2.22. Sei f eine skalare Funktion mit $f : \Omega \to \mathbb{R}, \Omega \subseteq \mathbb{R}^N$ ein Gebiet, $a \in \mathbb{R}^N$, $\|a\| = 1$. Dann heißt

$$\frac{\partial f}{\partial a}(p) = \lim_{t \to 0} \frac{f(p + ta) - f(p)}{t} = \left. \frac{\mathrm{d}}{\mathrm{d}t}(f(p + ta)) \right|_{t=0}$$

Richtungsableitung von f in $p \in \Omega$ bez. der Richtung a.

Sei f total differenzierbar in p, dann ist $\phi(t) := f(p + ta)$ ist eine verkettete Funktion

$$\begin{array}{rccc} \phi: & t & \mapsto & f(p+ta) \\ & \cup & & \cup \\ & \mathbb{R} & & \mathbb{R} \end{array}$$

$$\text{mit } \phi'(t) = \langle \operatorname{grad} f(p+ta), a\rangle \Rightarrow \frac{\partial f}{\partial a}(p) = \phi'(0) = \langle \operatorname{grad} f(p), a\rangle .$$

Speziell: Wenn $a = e_i$, dann $\dfrac{\partial f}{\partial e_i}(p) = \dfrac{\partial f}{\partial x_i}(p)$.

Bemerkung 2.23. Ist f total differenzierbar in p, so existiert die Richtungsableitung von f in p bezüglich jeder Richtung. Ist f partiell differenzierbar, so existiert die Richtungsableitung nur in jede Koordinatenrichtung.

Beispiel 2.24. $f(x_1, x_2) = \sin(x_1 x_2^2)$, $p = (0, 1)$, $a = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$.

$$\frac{\partial f}{\partial a}(0,1) = \left\langle \operatorname{grad} f(0,1), \left(\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}\right)\right\rangle = \left\langle (1,0), \left(\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2}\right)\right\rangle = \tfrac{\sqrt{2}}{2} .$$

2.2.2.4 Geometrische Deutung von ∇f = **grad** f

1)
$$\frac{\partial f}{\partial a}(p) = \langle \operatorname{grad} f(p), a\rangle = \|\nabla f(p)\| \underbrace{\|a\|}_{=1} \cos(\sphericalangle(\nabla f(p), a))$$

$\cos(\sphericalangle(\nabla f(p), a))$ wird maximal, falls $\sphericalangle(\nabla f(p), a) = 0$.

Folgerung 2.25. $\nabla f(p)$ gibt die Richtung des steilsten Anstieges von f an.

2) $\mathcal{N}_c := \{x \in \Omega | f(x) = c\}$ heißt *Niveau-* oder *Höhenlinie* von $y = f(x)$. Sei $\mathcal{N}_c$ in $U_r(p) \subseteq \Omega$ gegeben durch die Gleichung $x = \phi(t)$, $\phi \in C^1((a, b), \mathbb{R}^N)$. Auf $\mathcal{N}_c$ gilt

$$f(\phi(t)) = c \Rightarrow \frac{\mathrm{d}}{\mathrm{d}t}(f(\phi(t))) = 0 \Rightarrow \langle \nabla f(\phi(t)), \phi'(t)\rangle = 0 \text{ für alle } t \in (a, b),$$

das heißt, $\nabla f(p)$ steht senkrecht zum Tangentialvektor $\phi'(t)$ in p, oder ∇f steht senkrecht zur Kurve $\mathcal{N}_c$ in jedem ihrer Punkte.

2.3 Vektorfelder

2.3.1 Potentialfelder

Definition 2.26. Ein *Vektorfeld* auf $\Omega \subseteq \mathbb{R}^N$ ist eine Abbildung

$$\begin{array}{ccccc} v & : & \Omega & \to & \mathbb{R}^N \\ & & \cup & & \cup \\ & & x & \mapsto & v(x) \quad = \quad (v_1(x), \ldots, v_N(x)) \end{array}$$

Speziell heißt v ein C^r-*Vektorfeld* auf Ω, falls $v \in C^r(\Omega, \mathbb{R}^N)$.

Beispiel 2.27. i. In der Physik: Geschwindigkeitsfelder, Gravitationsfelder, elektrische Felder, meist noch abhängig von der Zeit t.

ii. Das einfachste Vektorfeld ist das *Gradientenfeld* $v = \nabla f$ einer skalaren Funktion $f \in C^1(\Omega)$. f heißt *Stammfunktion* zu v, $U = -f$ ($v = -\nabla U$) ein *Potential* von v sowie v ein *Potentialfeld*.

Satz 2.28. Zwei Stammfunktionen zum Potentialfeld v auf einem Gebiet Ω unterscheiden sich nur durch eine Konstante.

Beweis. Seien f_1, f_2 Stammfunktionen zu v, so ist $\nabla f_1 = v$, $\nabla f_2 = v$, also gilt für die Differenzfunktion $f := f_1 - f_2$: $\nabla f = 0$. Somit ist f konstant in jeder Umgebung $U_\varepsilon(p) \subset \Omega$ für $p \in \Omega$. Da Ω zusammenhängend ist, gibt es einen stetigen Weg von p zu jedem anderen $\tilde{p} \in \Omega$ und die möglichen Konstanten müssen alle gleich sein, d.h. $f = \text{const.}$ in Ω.

Beispiel 2.29. Nach dem Newtonschen Gravitationsgesetz gilt für die Anziehungskraft $F(x)$ einer Masse M in $0 \in \mathbb{R}^3$ auf einen Probekörper der Masse m in $x \in \mathbb{R}^3$

$$F(x) = -\gamma \frac{Mm}{\|x\|^3} x \text{ , d.h. } F_i(x) = -\gamma \frac{Mm}{\|x\|^3} x_i \text{ für } i = 1, 2, 3 \text{ .}$$

Dann ist $U(x) = -\gamma \frac{Mm}{\|x\|}$ ein Potential zu F in $\mathbb{R}^3 \setminus \{0\}$.

2.3.1.1 Integrabilitätsbedingungen

Frage: Wann ist ein gegebenes C^1-Vektorfeld auf $\Omega \subseteq \mathbb{R}^N$ ein Gradientenfeld?

Satz 2.30. *(Integrabilitätsbedingungen)*

$$v = \nabla f \quad \text{in } \Omega \quad \Rightarrow \quad \frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i} \quad \text{in } \Omega \text{ für alle } i, j = 1, \ldots, N \text{ .}$$

Beweis. Da $v_i = f_{x_i}$, ist $\frac{\partial v_i}{\partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial v_j}{\partial x_i}$ nach dem Satz von Schwarz.

Umgekehrt gilt auch die Hinlänglichkeit dieser Bedingungen in der folgenden Form:

Satz 2.31. Erfüllt ein Vektorfeld $v \in C^1(\Omega, \mathbb{R}^N)$ die Integrabilitätsbedingungen in Ω, so besitzt v in jeder Umgebung $U_\varepsilon(p) \subseteq \Omega$ eine Stammfunktion f.

Achtung: Die lokalen Potentiale lassen sich bei beliebigem Ω nicht immer widerspruchsfrei zu einem Potential auf ganz Ω ergänzen (s. Kurvenintegrale).

Beweis. (nur für $N = 2$).

$$v(x_1, x_2) = (P(x, y), Q(x, y)),\ x_1 = x, x_2 = y$$

Integrabilitätsbedingungen: $\dfrac{\partial v_1}{\partial x_2} = \dfrac{\partial v_2}{\partial x_1} \Leftrightarrow P_y = Q_x$.
Berechnung von f durch achsenparallele Integration:
Aus der Bedingung $f_x = P$ folgt $f(x, y) = g(x, y) + \phi(y)$ für festes y mit $g(x, y) = \int P(x, y)\mathrm{d}x$ und einer zu bestimmenden Funktion $\phi(y)$.
Aus der Bedingung $f_y = Q$ folgt $Q = \frac{\partial g}{\partial y} + \phi'(y)$, also $\phi(y) = \int \left(Q - \frac{\partial g}{\partial y}\right)\mathrm{d}y$.
Beachte: $Q - \frac{\partial g}{\partial y}$ hängt nicht von x ab, weil gilt

$$\frac{\partial}{\partial x}\left(Q(x, y) - \frac{\partial g}{\partial y}(x, y)\right) = Q_x - \frac{\partial^2 g}{\partial x \partial y} = Q_x - \frac{\partial}{\partial y}P = Q_x - P_y = 0\ .$$

Beispiel 2.32. Gesucht ist die Stammfunktion f zu $v(x, y) = (\underbrace{12xy + 3}_{=P}, \underbrace{6x^2 + 2y}_{=Q})$

1. Schritt: Test auf Exaktheit (Integrabilitätsbedingungen):

$$P_y = 12x, Q_x = 12x \Rightarrow P_y = Q_x$$

d.h. die Integrabilitätsbedingungen sind erfüllt.

2. Schritt: Integration bezüglich x:

$$f_x = P \Rightarrow f = \int P\mathrm{d}x + \phi(y) = \int (12xy + 3)\mathrm{d}x + \phi(y) = 6x^2y + 3x + \phi(y)$$

3. Schritt: Integration bezüglich y:

$$f_y = Q \Rightarrow \frac{\partial}{\partial y}\left(6x^2y + 3x + \phi(y)\right) = 6x^2 + 2y$$

$$\Rightarrow 6x^2 + \phi'(y) = 6x^2 + 2y \Rightarrow \phi'(y) = 2y \Rightarrow \phi(y) = y^2 + C$$

$$\Rightarrow \underline{\underline{f(x, y) = 6x^2y + 3x + y^2 + C}}\ .$$

2.3.1.2 Exakte Differentialgleichungen

Definition 2.33. Die Differentialgleichung

$$P(x, y) + Q(x, y)\frac{dy}{dx} = 0 \tag{2.3.1.2}$$

heißt *exakt* auf dem Rechteck $R := (a, b) \times (c, d) \subseteq \mathbb{R}^2$, falls (P, Q) ein C^1-Potentialfeld auf R ist, d.h. falls eine Funktion $f \in C^2(R)$ existiert mit $f_x = P$, $f_y = Q$ in R bzw. falls $P_y = Q_x$ in R gilt.

Bemerkung 2.34. Die Differentialgleichung (2.3.1.2) schreibt man häufig als Gleichung zwischen den Differentialen dx, dy in der Form

$$Pdx + Qdy = 0$$

Diese ist genau dann exakt, wenn $Pdx+Qdy$ das totale Differential df einer Funktion $f : R \to \mathbb{R}$ ist, d.h. $Pdx + Qdx = df = f_x dx + f_y dy$.

Folgerung 2.35. Sei (2.3.1.2) exakt auf R. Dann sind ihre Lösungskurven $y = y(x)$ für $x \in (a, b)$ genau die Äquipotentiallinien $f(x, y(x)) = \text{const.}$ von f, denn

$$f(x, y(x)) = \text{const} \;\Leftrightarrow\; \frac{d}{dx} f(x, y(x)) = 0 \;\Leftrightarrow\; f_x + f_y y' = P(x, y(x)) + Q(x, y(x))\frac{dy}{dx} = 0$$

für alle $x \in (a, b)$.

Beispiel 2.36. Gesucht: Lösung des Anfangswertproblems $(12xy+3)+(6x^2+2y)y' = 0\,,\ y(0) = 1$

1. Schritt: Test auf Exaktheit (s. Beispiel 2.32)

2. Schritt: Bestimmung einer Stammfunktion $f = f(x, y)$ (s. Beispiel 2.32)

3. Schritt: Einarbeitung der Anfangswerte (falls gegeben):
Allgemeine Lösung ist $f(x, y) = 6x^2y + 3x + y^2 = \text{const}$. Aus Anfangswerten folgt

$$f(x, y) = f(0, 1) \;\Rightarrow\; 6x^2y + 3x + y^2 = 0 + 0 + 1 = 1\ .$$

4. Schritt: Auflösung nach y (falls möglich, sonst nur implizite Lösungsdarstellung):

$$y^2 + 6x^2y + 3x - 1 = 0 \;\Rightarrow\; y_{1/2} = -3x^2 \pm \sqrt{9x^4 - 3x + 1} \;\Rightarrow\; y(x) = -3x^2 + \sqrt{9x^4 - 3x + 1}$$

wieder unter Verwendung der Anfangswerte.

2.3.2 (Differential-)Operationen mit Vektorfeldern

2.3.2.1 Das Vektorprodukt

Neben den algebraischen Operatoren $\lambda v, v + w, \langle v, w\rangle = vw = v \cdot w = v^\top w$ im euklidischen Vektorraum $\mathbb{R}^N$ für $\lambda \in \mathbb{R}$ und $v, w \in \mathbb{R}^N$ gibt es speziell im $\mathbb{R}^3$ das *Kreuz-* oder *Vektorprodukt*:

$$v \times w := \begin{vmatrix} e_1 & e_2 & e_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = (v_2w_3 - v_3w_2)e_1 + (v_3w_1 - v_1w_3)e_2 + (v_1w_2 - v_2w_1)e_3$$

(Entwicklung der formalen Determinante nach der ersten Zeile).

Eigenschaften des Vektorproduktes:

a. Das Vektorprodukt ist linear in jedem Faktor.

b. Das Vektorprodukt ist *schiefsymmetrisch*: $v \times w = -w \times v \quad (\Rightarrow\ v \times v = 0)$.

c. $\langle u, v \times w\rangle = \det \begin{pmatrix} u \\ v \\ w \end{pmatrix}$, ($u, v, w$ sind hier Zeilenvektoren) heißt auch *Spatprodukt* der drei Vektoren u, v, w (s. Übungsaufgabe).

Folgerung 2.37. $\langle v, v \times w\rangle = \det \begin{pmatrix} v \\ v \\ w \end{pmatrix} = 0$, $\langle w, v \times w\rangle = \det \begin{pmatrix} w \\ v \\ w \end{pmatrix} = 0$. Das heißt, $v \times w$ steht senkrecht auf der von v und w aufgespannten Ebene. v, w und $v \times w$ bilden ein *positiv orientiertes Dreibein* im $\mathbb{R}^3$ (3-Finger-Regel).

2.3.2.2 Rotation und Divergenz

Sei v ein C^1-Vektorfeld auf Ω und $\Omega \subseteq \mathbb{R}^N$ ein Gebiet.

1. Die *Divergenz* von v ist definiert durch $\quad \operatorname{div} v : \Omega \to \mathbb{R}$

$$\operatorname{div} v := \sum_{i=1}^{N} \frac{\partial v_i}{\partial x_i} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \ldots + \frac{\partial v_N}{\partial x_N} .$$

   Ist $\operatorname{div} v = 0$, so heißt das Vektorfeld *quellenfrei*.

2. Speziell für $N = 3$: Die *Rotation* von v ist definiert durch $\quad \operatorname{rot} v : \Omega \to \mathbb{R}^3$

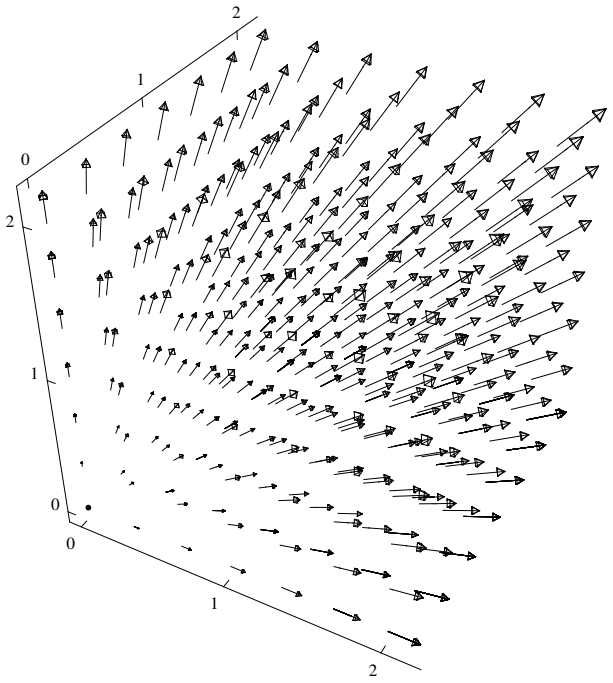
$$\operatorname{rot} v := \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}\right) e_1 + \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}\right) e_2 + \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) e_3 .$$

   Ist $\operatorname{rot} v = 0$, so heißt das Vektorfeld *wirbelfrei*.

Beispiel 2.38. i. $v(x) = x \in \mathbb{R}^3$

$$v_i(x) = x_i \ \Rightarrow\ \frac{\partial v_i}{\partial x_j} = \delta_{ij} \quad \text{für } i = 1, 2, 3 \ \Rightarrow\ \operatorname{div} v = 3,\ \operatorname{rot} v = 0.$$

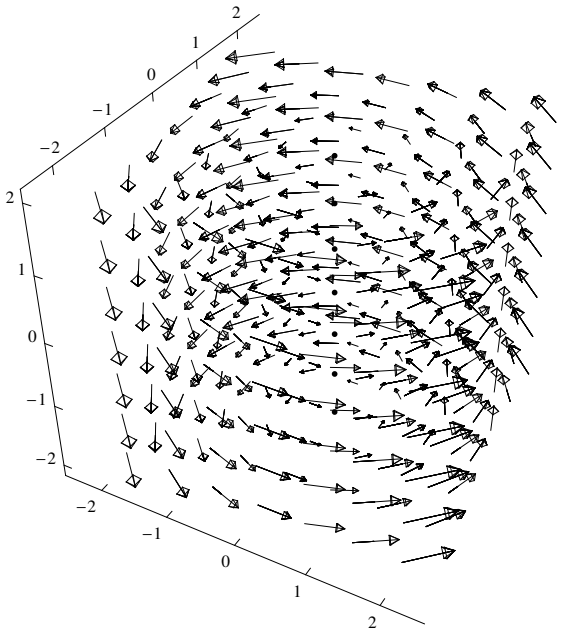
Dieses Vektorfeld vermittelt das Bild einer Strömung mit einer Quelle in 0.
$\operatorname{div} v$ ist ein Maß für die „Quellstärke“ des Vektorfeldes v.



ii. $v(x) = (-x_2, x_1, 0)$.

$$\frac{\partial v_1}{\partial x_2} = -1,\ \frac{\partial v_2}{\partial x_1} = 1,\ \text{sonst}\ \frac{\partial v_j}{\partial x_j} = 0\ \Rightarrow\ \operatorname{div} v = 0,\ \operatorname{rot} v = (0,0,2).$$

Dieses Vektorfeld vermittelt das Bild einer Strömung mit einem Wirbel um die x_3-Achse.
$\|\operatorname{rot} v\|$ ist ein Maß für die „Wirbelstärke“ des Vektorfeldes v und $\operatorname{rot} v$ gibt die Richtung der Rotationsachse an.



2.3.2.3 Der Nablaoperator

Mit Hilfe des *Nablaoperators* $\nabla := \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_N}\right)$ lassen sich die Begriffe Gradient, Divergenz und Rotation einheitlich formulieren *(Nabla-Kalkül)*.

$$
\begin{aligned}
f : \Omega \to \mathbb{R}\,, \quad & \nabla f := \left(\frac{\partial}{\partial x_1} f, \frac{\partial}{\partial x_2} f, \ldots, \frac{\partial}{\partial x_N} f\right) = \operatorname{grad} f\,, \\
v : \Omega \to \mathbb{R}^N\,, \quad & \nabla v := \langle \nabla, v \rangle = \frac{\partial}{\partial x_1} v_1 + \frac{\partial}{\partial x_2} v_2 + \ldots + \frac{\partial}{\partial x_N} v_N = \operatorname{div} v\,, \\
v : \mathbb{R}^3 \supseteq \Omega \to \mathbb{R}^3\,, \quad & \nabla \times v := \begin{vmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ v_1 & v_2 & v_3 \end{vmatrix} = \operatorname{rot} v\,.
\end{aligned}
$$

2.3.2.4 Rechenregeln

Sei $f \in C^2(\Omega)$, $v, w \in C^2(\Omega, \mathbb{R}^3)$, $\Omega \subseteq \mathbb{R}^3$.

a. $\operatorname{rot}(\operatorname{grad} f) = \nabla \times (\nabla f) = (\nabla \times \nabla) f = 0$

b. $\operatorname{div}(\operatorname{rot} v) = \nabla(\nabla \times v) = \begin{vmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \\ v_1 & v_2 & v_3 \end{vmatrix} = 0$

c. $\operatorname{div}(\operatorname{grad} f) = \nabla(\nabla f) = \nabla\nabla f = \nabla^2 f = \sum_{i=1}^{N} \frac{\partial^2}{\partial x_i^2} f = \left(\frac{\partial^2}{\partial x_1^2} + \ldots + \frac{\partial^2}{\partial x_N^2}\right) f =: \Delta f.$
Δ ist der *Laplace-Operator*.

d. $\operatorname{div}(fv) = \nabla(fv) = (\nabla f)v + f\nabla v = \langle \operatorname{grad} f, v \rangle + f \operatorname{div} v$

e. $\operatorname{rot}(fv) = \nabla \times (fv) = (\nabla f) \times v + f(\nabla \times v) = \operatorname{grad} f \times v + f \operatorname{rot} v$

f. $\operatorname{div}(v \times w) = \nabla(v \times w) = (\nabla \times v)w - v(\nabla \times w) = \langle \operatorname{rot} v, w \rangle - \langle v, \operatorname{rot} w \rangle$

(Beweise s. Übungsaufgaben)

Bemerkung 2.39.

1. Die Integrabilitätsbedingungen für $v \in C^1(\Omega, \mathbb{R}^3)$ sind äquivalent zu $\operatorname{rot} v = 0$

2. Alle Potentialfelder sind wirbelfrei für $v \in C^1(\Omega, \mathbb{R}^N)$ (nach Rechenregel a).

2.4 Taylorscher Lehrsatz und Anwendungen

2.4.1 Der Taylorsche Lehrsatz im $\mathbb{R}^N$

2.4.1.1 Herleitung der Taylorformel

Zur Taylorentwicklung einer skalaren Funktion $f : \Omega \to \mathbb{R}^N$ ($\Omega \subseteq \mathbb{R}^N$ ist Gebiet) studiert man f in $U_\delta(p) \subset \Omega$ längs aller Geraden: $x = p + th$, $h \in \mathbb{R}^N$ *(Methode der eindimensionalen Schnitte)*. Das heißt, wir untersuchen

$$g(t) := f(p + th) \quad \text{für } \|h\| < \delta,\ t \in [-1, 1] \quad (\Rightarrow\ p + th \in U_\delta(p)).$$

Taylorentwicklung der reellen Funktion g an $t = 0$:

$$g(1) = \underbrace{g(0) + \frac{g'(0)}{1!} + \frac{1}{2!}g''(0) + \ldots + \frac{1}{n!}g^{(n)}(0)}_{=T_n(0,1)} + \underbrace{\frac{1}{(n+1)!}g^{(n+1)}(\theta)}_{=R_n(0,1)} \tag{2.4.1.3}$$

mit $\theta \in (0, 1)$.

$$g'(t) = \langle \operatorname{grad} f(p + th), h\rangle = \langle h, \nabla f(p + th)\rangle = ((h\nabla)f)(p + th) = \sum_{i=1}^{N} h_i \frac{\partial}{\partial x_i} f(p + th)\,,$$

$$g''(t) = \frac{\mathrm{d}}{\mathrm{d}t}((h\nabla)f)(p + th) = (h\nabla)\frac{\mathrm{d}}{\mathrm{d}t}\underbrace{f(p + th)}_{=g(t)} = (h\nabla)((h\nabla)f)(p + th) = \left((h\nabla)^2 f\right)(p + th)\,,$$

$$\vdots$$

$$g^{(n)}(t) = ((h\nabla)^n f)(p + th)$$

mit $(h\nabla)^n = \left(h_1\frac{\partial}{\partial x_1} + h_2\frac{\partial}{\partial x_2} + \ldots + h_N\frac{\partial}{\partial x_N}\right)^n$. Durch Einsetzen dieser Terme in (2.4.1.3) ergibt sich

Satz 2.40. *(Taylorformel)*
Sei $f \in C^{n+1}(\Omega)$, $\Omega \subseteq \mathbb{R}^N$ ein Gebiet, $U_\delta(p) \subset \Omega$, $h \in \mathbb{R}^N$ und $\|h\| < \delta$. Dann existiert ein $\theta \in (0, 1)$:

$$f(p + h) = \underbrace{\sum_{k=0}^{N} \frac{1}{k!}\left((h\nabla)^k f\right)(p)}_{\text{Taylorpolynom } T_n(h)} + \underbrace{\frac{1}{(n+1)!}\left((h\nabla)^{n+1} f\right)(p + \theta h)}_{\text{Restglied } R_n(h)} \tag{2.4.1.4}$$

2.4.1.2 Spezialfälle

a. $n = 0$:

$$f(p+h) = f(p) + \langle h, \operatorname{grad} f(p+\theta h)\rangle$$

(entspricht dem Mittelwertsatz der Differentialrechnung).

b. $n = 1$:

$$\begin{aligned}(h\nabla)^2 f &= \left(\sum_{i=1}^{N} h_i \frac{\partial}{\partial x_i}\right)^2 f = \left(\sum_{i=1}^{N} h_i \frac{\partial}{\partial x_i}\right)\left(\sum_{j=1}^{N} h_j \frac{\partial}{\partial x_j}\right) f \\ &= \sum_{i,j=1}^{N} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j} f = h^\top f'' h\end{aligned}$$

wobei h als Spaltenvektor aufzufassen ist, $f''(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j=1,\dots,N}$ die Ableitung von $f' = \operatorname{grad} f$ ist und *Hesse-Matrix* (Hesse 1811 - 1874) von f and der Stelle x heißt.

Beispiel 2.41. Anwendung der Taylorformel (2.4.1.4) mit $n = 1$ auf $f(x, y) = \sin(x + y)$ an der Stelle $p = (0, 0)$. Setze $h = (h_1, h_2) = (x, y) \in \mathbb{R}^2$.

$$\begin{aligned}\sin(x+y) = &\sin(0+0) + \langle (x,y), \overbrace{(\cos 0, \cos 0)}^{=\operatorname{grad} f(0,0)}\rangle \\ &+ \frac{1}{2}\left\langle (x,y), \begin{pmatrix} -\sin(\theta(x+y)) & -\sin(\theta(x+y)) \\ -\sin(\theta(x+y)) & -\sin(\theta(x+y)) \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle \\ = &x + y - \frac{1}{2}(x^2 + 2xy + y^2)\sin(\theta(x+y)) \qquad \text{mit } \theta \in (0,1).\end{aligned}$$

2.4.2 Berechnung lokaler Extremstellen skalarer Funktionen

Definition 2.42. $f \in C^2(\Omega)$ hat in $p \in \Omega$ ein *lokales Minimum*, wenn es $U_r(p), r > 0$ gibt, so dass $f(x) \geq f(p)$ für alle $x \in U_r(p) \cap \Omega$. $f \in C^2(\Omega)$ hat in $p \in \Omega$ ein *lokales Maximum*, wenn es $U_r(p), r > 0$ gibt, so dass $f(x) \leq f(p)$ für alle $x \in U_r(p) \cap \Omega$. Lokale Minima/Maxima bezeichnet man auch als *lokale Extrema* von f im Gegensatz zu *globalen Extrema* von f auf Ω wenn die Ungleichungen für alle $x \in \Omega$ gelten. p heißt dann auch *lokale* bzw. *globale Extremalstelle*.

Wir betrachten im Folgenden nur lokale Minima. Für lokale Maxima muss man $-f$ untersuchen.

2.4.2.1 Notwendige Bedingung

Satz 2.43. Hat f in $p \in \Omega$ ein lokales Minimum, so gilt

$$f'(p) = 0 \quad \text{und} \quad \langle h, f''(p)h\rangle \geq 0 \ \text{ für alle } h \in \mathbb{R}^N$$

(d.h., die Hesse-Matrix $f''(p)$ muss positiv semidefinit sein).

Beweis. (i). Betrachte die reelle Funktion $f_1(x_1) := f(x_1, p_2, \ldots, p_N)$. Wegen $f(x) \geq f(p)$ hat f_1 ein lokales Minimum. Somit ist $f_1'(p_1) = \frac{\partial f}{\partial x_1}(p) = 0$. Analoges gilt somit auch für die anderen Koordinatenrichtungen $\Rightarrow$ grad $f(p) = 0$.

(ii). Aus der Taylorformel (2.4.1.4) mit $n = 1$ folgt

$$f(p+h) = f(p) + \underbrace{\langle h, f'(p)\rangle}_{=0} + \frac{1}{2}\langle h, f''(p+\theta h)\rangle \text{ für alle } h \subset \mathbb{R}^N, \|h\| < \delta$$

$$\text{Wegen } f(x) \geq f(p) \text{ ist} \qquad f(p+h) \geq f(p) \Rightarrow \langle h, f''(p+\theta h)h\rangle \geq 0$$

Setze $h := \lambda a$ mit festem Vektor $a \in \mathbb{R}^N$, $\|a\| = 1$, $\lambda \in (0, \delta)$.

$$\Rightarrow \lambda^2\langle a, f''(p+\theta\lambda a)a\rangle \geq 0 \Rightarrow \langle a, f''(p+\theta\lambda a)a\rangle \geq 0$$

Der Grenzübergang $\lambda \to 0+$ liefert wegen Stetigkeit von f'': $\langle a, f''(p)a\rangle \geq 0$ für beliebige a mit $\|a\| = 1$. Für beliebige $h \in \mathbb{R}^N$ schreibe $h = \|h\|a$ mit $a = \frac{1}{\|h\|}h$ ($\Rightarrow \|a\| = 1$).

$$\langle h, f''(p)h\rangle = \|h\|^2\langle a, f''(p)a\rangle \geq 0\,.$$

2.4.2.2 Hinreichende Bedingung

Satz 2.44. Gilt für $p \in \Omega$: $f'(p) = 0$ und ist die Hesse-Matrix $f''(p)$ positiv definit, dann hat f in p ein lokales Minimum.

Beweis. Auch hier wird wieder die Taylorformel (2.4.1.4) mit $n = 1$ benutzt. Um $f(x) \geq f(p)$ nachzuweisen, braucht man nur das Vorzeichen von $R_1(h)$ zu untersuchen. Nach Vorraussetzung ist

$$\langle h, f''(p)h\rangle > 0 \quad \text{für alle } h \in \mathbb{R}^N, h \neq 0$$

Wegen Stetigkeit von f'' ist $\langle h, f''(x)h\rangle > 0$ für dieselben h und $\|x - p\| < \delta \Rightarrow f(x) > f(p)$ für alle $x \in U_\delta(p)$.

2.4.2.3 Ergänzende Betrachtungen

Bemerkung 2.45.

1. Punkte $p \in \Omega$ mit grad $f(p) = 0$ heißen *stationäre* oder *kritische Punkte* von f
2. Aus dem soeben geführten Beweis folgt, dass die Minimalstelle p in $U_\delta(p)$ eindeutig bestimmt ist.
3. Ist $f''(p)$ indefinit, so kann in p kein lokales Extremum von f vorliegen.

Der Test auf positive Definitheit von $f''(p)$ erfolgt mittels

Satz 2.46. (*Trägheitssatz von Sylvester*, Sylvester 1814 - 1897)

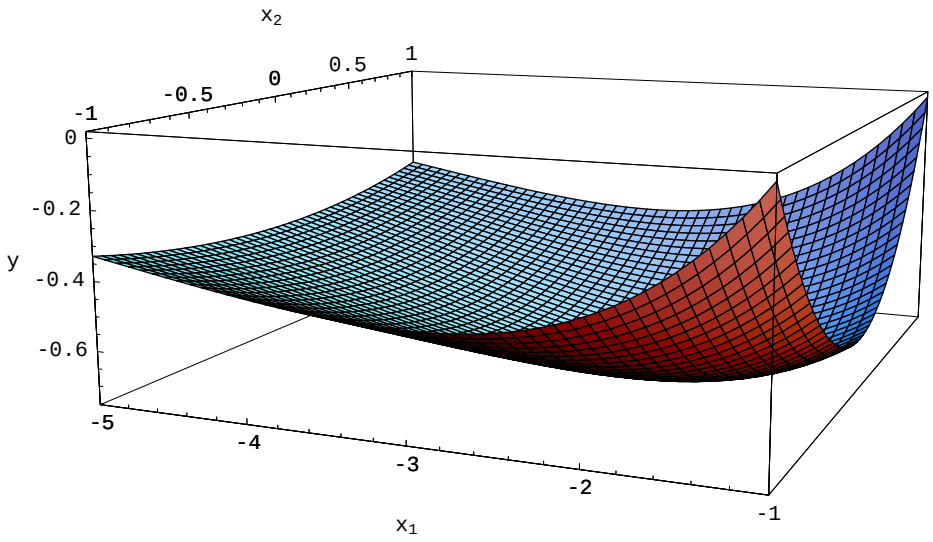
$$f''(p) \text{ ist positiv definit} \quad \Leftrightarrow \quad \text{Alle Eigenwerte von } f''(p) \text{ sind positiv}$$
$$\Leftrightarrow \quad \text{Alle Hauptminoren der Matrix } f''(p) \text{ sind positiv}$$

Folgerung 2.47. Hinreichende Bedingung für $f := f(x_1, x_2)$:
Sei $f \in C^2(\Omega)$, $\Omega \subseteq \mathbb{R}^2$ ein Gebiet, $f_{x_1}(p) = f_{x_2}(p) = 0$. Dann hat f in p ein lokales Minimum, falls $f_{x_1x_1}(p) > 0$ und $\det(f_{x_ix_j}(p)) > 0$. f hat ein lokales Maximum, falls $f_{x_1x_1}(p) < 0$ und $\det(f_{x_ix_j}(p)) > 0$. Es liegt kein lokales Extremum vor, falls $\det(f_{x_ix_j}(p)) < 0$.

2.4.2.4 Beispiele

Beispiel 2.48.

$$y = f(x_1, x_2) = (x_1 + x_2^2)e^{\frac{1}{2}x_1}, \quad x = (x_1, x_2) \in \mathbb{R}^2$$



i. Notwendige Bedingung: Bestimmung von $x \in \mathbb{R}^2$, so dass grad $f = 0$.

$$\frac{\partial f}{\partial x_1} = \left(1 + \tfrac{1}{2}(x_1 + x_2^2)\right)e^{\frac{1}{2}x_1}, \quad \frac{\partial f}{\partial x_2} = 2x_2e^{\frac{1}{2}x_1}$$

$$\text{grad } f(x) = 0 \;\Leftrightarrow\; \begin{Bmatrix} \left(1 + \frac{1}{2}(x_1 + x_2^2)\right)e^{\frac{1}{2}x_1} & = & 0 \\ 2x_2e^{\frac{1}{2}x_1} & = & 0 \end{Bmatrix} \;\Leftrightarrow\; \begin{matrix} x_1 & = & -2 \\ x_2 & = & 0 \end{matrix}$$

Somit ist $p = (-2, 0)$ einziger stationärer Punkt von f.

ii. Hinreichende Bedingung: Definitheit der Hesse-Matrix

$$\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{i,j=1,2} = \begin{pmatrix} \left(1 + \frac{x_1}{4} + \frac{x_2^2}{4}\right)e^{\frac{x_1}{2}} & x_2e^{\frac{x_1}{2}} \\ x_2e^{\frac{x_1}{2}} & 2e^{\frac{x_1}{2}} \end{pmatrix}$$

$$\Rightarrow \frac{\partial^2 f}{\partial x_1^2}(-2, 0) = \frac{1}{2e} > 0, \quad \det\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(-2, 0)\right) = \begin{vmatrix} \frac{1}{2e} & 0 \\ 0 & \frac{2}{e} \end{vmatrix} = \frac{1}{e^2} > 0$$

$\Rightarrow$ lokales Minimum in p.

Ergebnis: $\min_{x\in\mathbb{R}^2} f(x) = f(p) = -\frac{2}{e}$. Es existiert kein lokales und somit auch kein globales Maximum.

Bemerkung 2.49. Zur Bestimmung globaler Extrema einer reellwertigen Funktion $f = f(x)$ auf einer beliebigen Menge M des $\mathbb{R}^N$ bestimmt man mittels Differentialrechnung zunächst alle lokalen Extrema in $\operatorname{int} M$ und vergleicht dann deren Funktionswerte von f an diesen Stellen. Anschließend muss noch das Verhalten von f in der Umgebung von ∂M untersucht werden.

Beispiel 2.50.

$$f(x_1, x_2) = e^{x_1+x_2}$$

Gesucht sind die Extrema von f auf $\quad M = \{(x_1, x_2) | \, x_1 \geq 0, \; x_2 \geq 0\}$.

$$f_{x_1} = f_{x_2} = e^{x_1+x_2} \neq 0 \quad \text{für alle } x \in \mathbb{R}^2$$

$\Rightarrow$ keine lokalen Extrema auf $\operatorname{int} M$. Aber:

$$f(x) \geq 1 = e^{0+0} = f(0) \quad \text{für alle } x \in M \;\Rightarrow\; \min_{x\in M} f(x) = f(0)$$

Das heißt, $f(0) = f(0, 0) = 1$ ist globales Minimum auf M. Es existiert kein globales Maximum von f auf M wegen Grenzverhalten von f auf den Koordinatenachsen.

2.5 Implizite Funktionen

2.5.1 Der Hauptsatz über implizite Funktionen

2.5.1.1 Problemstellung

Sei $F = F(x, y)$, $x, y \in \mathbb{R}$ gegeben, so fragt man nach der Auflösbarkeit der Gleichung $F(x, y) = 0$ nach der Variablen y, d.h., man sucht die reelle Funktion f :

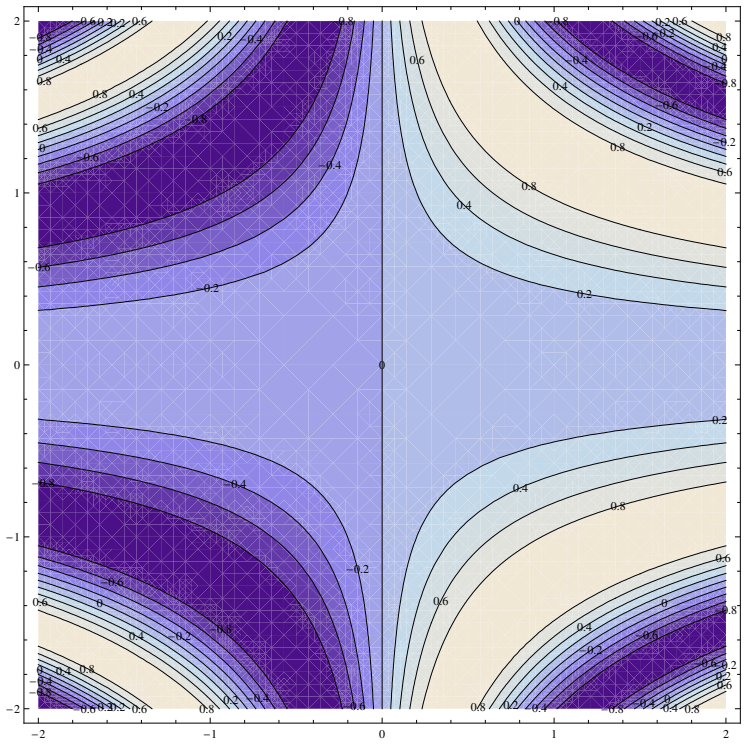
$$F(x, y) = 0 \;\Leftrightarrow\; y = f(x).$$

Mit anderen Worten: Die Niveaulinie $\mathcal{N}_0 := \{(x, y) \in \mathbb{R}^2 | F(x, y) = 0\}$ soll dargestellt werden durch eine Kurve, die durch den Graphen einer Funktion f bestimmt wird.
Die Fragestellung lautet also:

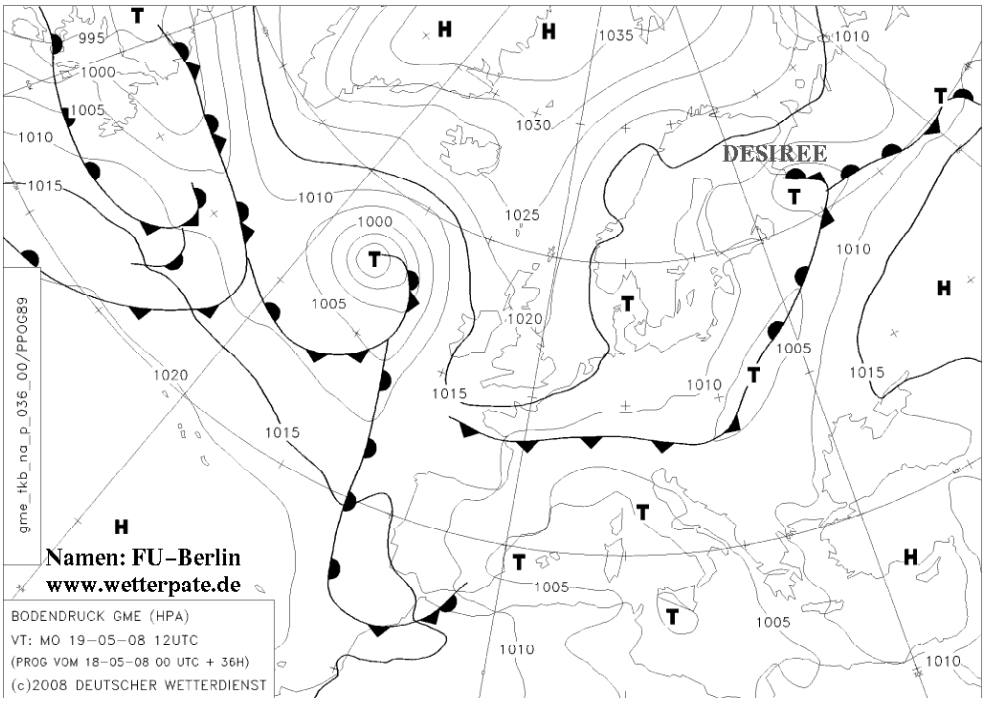
$$\text{Ist} \qquad \mathcal{N}_0 = \left\{(x, y) \in \mathbb{R}^2 \middle| \; y = f(x)\right\}?$$

Beispiel 2.51. Niveaulinien der reellwertigen Funktion

$$F(x, y) := \sin\left(xy^2\right) \quad \text{für} \quad (x, y) \in \mathbb{R}^2$$



Niveaulinien von reellwertigen Funktionen sind z.B. die Höhenlinien eines Geländes, dargestellt auf einer Wanderkarte und die Isobaren auf einer Wetterkarte.

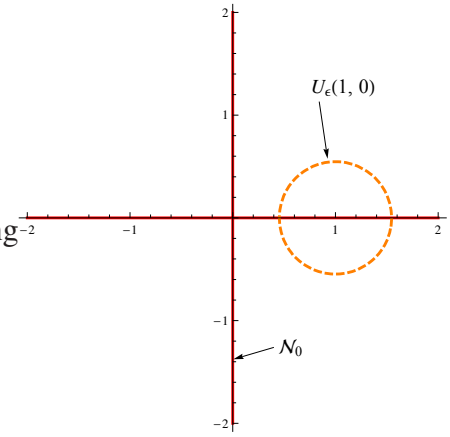


Beispiel 2.52. i. $F(x,y) = x^2y - 3x + y$ Also ist

$$\mathcal{N}_0 = \left\{(x,y)\,\middle|\, x^2y - 3x + y = 0\right\} = \left\{(x,y)\,\middle|\, y = \tfrac{3x}{1+x^2}\right\} \;\Rightarrow\; f(x) = \frac{3x}{1+x^2}\,.$$

ii. Aber nicht für jedes F ist $\mathcal{N}_0$ durch einen Graphen bestimmt, zum Beispiel

$$\begin{aligned} F(x,y) &= xy \\ \mathcal{N}_0 &= \{(x,y)|\, xy = 0\} \\ &= \{(x,0)|\, x \in \mathbb{R}\} \cup \{(0,y)|\, y \in \mathbb{R}\} \end{aligned}$$



Jedoch ist $\mathcal{N}_0$ ein Graph in einer Umgebung $U_\varepsilon((1,0)), \varepsilon \in (0,1)$:

$$\begin{aligned} \mathcal{N}_0 \cap U_\varepsilon((1,0)) &= \{(x,0)|\, |x-1| < \varepsilon\} \\ &= \{(x,y)|\, y \equiv 0, |x-1| < \varepsilon\} \end{aligned}$$

Deshalb betrachten wir die lokale Auflösbarkeit von $F(x,y) = 0$ nach y, wobei $F : \Omega \times \Sigma \to \mathbb{R}^M$ mit $\Omega \subseteq \mathbb{R}^N$, $\Sigma \subseteq \mathbb{R}^M$. Ω und Σ sind Gebiete in den Räumen. Das heißt, die M Gleichungen

$$\left.\begin{pmatrix} F_1(x_1,\ldots,x_N,y_1,\ldots,y_M) &=& 0 \\ \vdots & & \\ F_M(x_1,\ldots,x_N,y_1,\ldots,y_M) &=& 0 \end{pmatrix}\right. \Leftrightarrow F(x,y) = 0 \qquad (2.5.1.5)$$

sollen nach den M Variablen $y_1,\ldots,y_M$ aufgelöst werden in Abhängigkeit von $x_1,\ldots,x_N$.

2.5.1.2 Hauptsatz

Satz 2.53. *(Hauptsatz über implizite Funktionen)*
Sei $F \in C^1(\Omega \times \Sigma, \mathbb{R}^M)$. Für ein $(\xi,\eta) \in \Omega$ gelte

$$F(\xi,\eta) = 0 \quad \text{sowie die } \textit{Auflösebedingung} \quad \det\left(\frac{\partial F_i}{\partial y_k}(\xi,\eta)\right) \neq 0\,.$$

Dann gibt es $U := U_\delta(\xi) \subset \Omega$ und $V := U_\varepsilon(\eta) \subset \Sigma$, so dass genau eine stetige Funktion $f : U \to V$ existiert mit $f(\xi) = \eta$ und $F(x, f(x)) = 0$ für alle $x \in U$.

Beweis. (Beweisidee). Wir schreiben das nichtlineare Gleichungssystem (2.5.1.5) (bei festgehaltenem $x \in \Omega$) um in die Fixpunktgleichung $T_0(y) = y$ mit der Abbildung

$$T_0(y) := y - F_y(\xi,\eta)^{-1} F(x,y)$$

Lässt man $x \in U$ wieder variabel zu, so vermittelt T_0 eine Abbildung von $C(U, \mathbb{R}^M)$ in sich. Mittels Banachschem Fixpunktsatz ergibt sich dann die Existenz eines eindeutigen $f \in C(U, \mathbb{R}^M)$ mit

$$T_0(f) = f \;\Leftrightarrow\; f(x) - F_y(\xi,\eta)^{-1} F(x, f(x)) = f(x) \;\Leftrightarrow\; F(x, f(x)) = 0 \text{ für alle } x \in U.$$

Die nötigen Abschätzungen zum Nachweis der Kontraktivität der Abbildung T_0 folgen dabei aus dem Taylorschem Lehrsatz für F.

2.5.1.3 Spezialfall

Spezialfall: $M = N = 1$. $F(x, y) = 0$ ist auflösbar nach y in $U_\delta(\xi) \times U_\varepsilon(\eta)$ mit $F(\xi, \eta) = 0$, falls

$$F_y(\xi,\eta) = \frac{\partial F}{\partial y}(\xi,\eta) \neq 0$$

ist. Der Hauptsatz liefert $y = f(x)$ mit einer stetigen rellen Funktion f.

Beispiel 2.54.

$$\begin{aligned} F(x,y) &= \mathrm{e}^{\sin(xy)} + x^2 - 2y - 1 \\ F(0,0) &= \mathrm{e}^{\sin 0} + 0 - 2 \cdot 0 - 1 = 0 \\ \frac{\partial F}{\partial y} &= x\cos(xy)\mathrm{e}^{\sin(xy)} - 2 \\ \frac{\partial F}{\partial y}(0,0) &= -2 \neq 0\,. \end{aligned}$$

Es existiert $y = f(x)$ mit $f(0) = 0$ und $F(x, f(x)) = 0$.

2.5.1.4 Korollar zum Hauptsatz 2.53

a. Für genügend kleines $\delta > 0$ ist $f \in C^1(U, \mathbb{R}^M)$. Es gilt

$$f'(x) = -\Big(F_y(x, f(x))\Big)^{-1} F_x(x, f(x))\,, \tag{2.5.1.6}$$

denn man kann die Ableitung von f berechnen, ohne f selbst zu kennen. Es gilt:

$$F(x, f(x)) = 0 \;\Leftrightarrow\; \left\{ \begin{array}{l} F_i(x_1, \ldots, x_N, f_1(x), \ldots, f_M(x)) = 0 \\ \text{für alle } \; x = (x_1, \ldots, x_N) \in U,\; i = 1, \ldots, M. \end{array} \right. \tag{2.5.1.7}$$

$f'(x) = \left(\dfrac{\partial f_i}{\partial x_j}\right)_{i=1,\ldots,M,\, j=1,\ldots,N}$ berechnet sich wie folgt:

$$\begin{aligned} (2.5.1.7) \;&\Rightarrow\; \frac{\partial}{\partial x_j}(F_i(x, f_1(x), \ldots, f_M(x))) = 0 \\ &\Rightarrow\; \frac{\partial F_i}{\partial x_j} + \frac{\partial F_i}{\partial y_1}(x, f(x))\frac{\partial f_1}{\partial x_j} + \ldots + \frac{\partial F_i}{\partial y_M}(x, f(x))\frac{\partial f_M}{\partial x_j} = 0 \\ &\Leftrightarrow\; \frac{\partial F_i}{\partial x_j} + \operatorname{grad}_y F_i(x, f(x))\frac{\partial f}{\partial x_j} = 0 \end{aligned}$$

für $i = 1, \ldots, M$. Lineares Gleichungssystem für $\frac{\partial f}{\partial x_j}$:

$$\frac{\partial F}{\partial x_j} + F_y(x, f(x))\frac{\partial f}{\partial x_j} = 0 \, .$$

Das Gleichungssystem ist auflösbar wegen $\det F_y(x, f(x)) \neq 0$.

$$\Rightarrow \ \frac{\partial f}{\partial x_j} = -\Big(F_y(x, f(x))\Big)^{-1} \frac{\partial F}{\partial x_j}(x, f(x)) \ \Leftrightarrow \ (2.5.1.6).$$

b. Ist F eine C^r-Abbildung auf $\Omega \times \Sigma$, so ist $f \in C^r(U, \mathbb{R}^M)$.

In Beispiel 2.54 ergibt sich somit durch direktes Nachrechnen (*ohne* Anwendung der Formel (2.5.1.6)):

$$\begin{aligned}
0 &= F(x, f(x)) = \mathrm{e}^{\sin(x \cdot f(x))} + x^2 - 2f(x) - 1 \\
0 &= \frac{\mathrm{d}}{\mathrm{d}x} F(x, f(x)) = F_x(x, f(x)) + F_y(x, f(x)) f'(x) \\
&= \Big(y\cos(xy)\mathrm{e}^{\sin(xy)} + 2x\Big) + \Big(x\cos(xy)\mathrm{e}^{\sin(xy)} - 2\Big) f'(x)\Big|_{y=f(x)} \\
0 &= f(x)\cos(x \cdot f(x))\mathrm{e}^{\sin(x \cdot f(x))} + 2x - 2f'(x) + x \cdot f'(x) \, .
\end{aligned}$$

Berechnung von $f'(0)$:

$$0 = 0 + 2 \cdot 0 - 2f'(0) \ \Rightarrow \ f'(0) = 0 \, .$$

Taylorentwicklung von f:

$$f(x) = f(0) + f'(0) + R_1(x) = f''(\vartheta x)\frac{x^2}{2} \, , \ \vartheta \in (0, 1).$$

2.5.2 Extrema von Funktionen unter Nebenbedingungen

2.5.2.1 Problemstellung

Beispiel 2.55. Sei $g(x_1, x_2)$ eine Funktion, welche die Verteilung des Luftdrucks in einem Gebiet $\Omega \subseteq \mathbb{R}^2$ beschreibt. Gesucht ist eine Minimalstelle von g unter der Nebenbedingung, dass (x_1, x_2) auf Meereshöhe liegt, das heißt $f(x_1, x_2) = 0$.

Allgemein: Suche lokale Minima einer Funktion $g \in C^1(\Omega), \Omega \subseteq \mathbb{R}^N$ unter den M Nebenbedingungen

$$\left.\begin{array}{rcl} f_1(x) &=& 0 \\ \vdots & & \\ f_M(x) &=& 0 \end{array}\right\} \ \Leftrightarrow \ f(x) = 0 \, .$$

Das heißt: Suche $\quad p \in \mathcal{N}_0 := \{x \in \Omega |\ f(x) = 0\}$

$$\text{mit} \quad g(x) \geq g(p) \quad \text{für alle} \ \ x \in \mathcal{N}_0 \cap U_r(p) \quad \text{für ein} \ \ r > 0.$$

1. Möglichkeit: Löse das Gleichungssystem $f(x) = 0$ nach gewissen Variablen auf, setze die Auflösefunktion in g ein und betrachte dann die Extremwertbestimmung für eine Funktion der restlichen Variablen ohne Nebenbedingungen.

2. Möglichkeit: *Methode der Lagrangeschen Multiplikatoren* (unter Vermeidung der expliziten Auflösung der Nebenbedingungen).

2.5.2.2 Lagrangesche Multiplikatoren

Satz 2.56. Sei $g \in C^1(\Omega)$, $f \in C^1(\Omega, \mathbb{R}^M)$, $M < N$ und habe $f'(x) \in \mathbb{R}^{M\times N}$ maximalen Rang für alle $x \in \mathcal{N}_0$. Hat g in $p \in \mathcal{N}_0$ ein lokales Extremum unter den Nebenbedingungen $f(x) = 0$, so gibt es eindeutig bestimmte reelle Zahlen (*Lagrangesche Multiplikatoren*) $\lambda_1, \dots, \lambda_M$ mit

$$\nabla g(p) + \sum_{i=1}^{M} \lambda_i \nabla f_i(p) = 0 \tag{2.5.2.8}$$

Beweis. (i). Wegen der Voraussetzung Rang $f'(x) = M$ sind die Zeilen

$$\nabla f_1(x), \nabla f_2(x), \dots, \nabla f_M(x) \in \mathbb{R}^N$$

von $f'(x)$ sind linear unabhängig für jedes $x \in \mathcal{N}_0$. Durch Umnummerierung der Variablen $x_1, \dots, x_N$ können wir erreichen, dass

$$x = (x_1, \dots, x_L, x_{L+1}, \dots, x_N)\ ,\ L = N - M \qquad \text{und} \qquad \det\left(\left(\frac{\partial f_i}{\partial x_{L+k}}(\xi)\right)_{i,k=1,\dots,M}\right) \neq 0$$

ist. Nach dem Hauptsatz 2.53 über implizite Funktionen kann damit für $x \in U_r(p)$ das nichtlineare Gleichungssystem $f(x) = 0$ nach den Variablen $x_{L+1}, \dots, x_{L+M}$ aufgelöst werden, d.h.

$$f(x) = 0 \ \Leftrightarrow\ x_{L+i} = \phi_i(\underbrace{x_1, \dots, x_L}_{:=u\in\mathbb{R}^L})$$

für $i = 1, \dots, M$, $\phi_i \in C^1(U_r(q))$, $q := (p_1, \dots, p_L)$,

$$\Phi(u) := (u_1, \dots, u_L, \phi_1(u), \dots, \phi_M(u)), \quad \Phi : U_r(q) \subset \mathbb{R}^L \to \mathbb{R}^N.$$

Φ ist eine C^1-Abbildung, $\Phi(q) = p$,

$$\frac{\partial \Phi}{\partial u_j} = (\underbrace{0, \dots, 1, 0, \dots}_{L}, \underbrace{\frac{\partial \phi_1}{\partial u_j}, \dots, \frac{\partial \phi_M}{\partial u_j}}_{M}) \quad \text{bzw.} \quad \frac{\partial \Phi}{\partial u_j}(q) = e_j + \sum_{k=1}^{M} \frac{\partial \phi_k}{\partial u_j}(q) e_{L+k} \in \mathbb{R}^N.$$

$$\frac{\partial \Phi}{\partial u_1}(q), \dots, \frac{\partial \Phi}{\partial u_L}(q) \text{ sind linear unabhängig} \quad \Rightarrow\ T := \operatorname{lin}\left\{\frac{\partial \Phi}{\partial u_1}(q), \dots, \frac{\partial \Phi}{\partial u_L}(q)\right\}$$

ist ein linearer Unterraum des $\mathbb{R}^N$, $\dim T = L$.

(ii).
$$f(\Phi(u)) = f(x) = 0 \quad \text{für alle } x \in U_r(p)$$
beziehungsweise
$$u \in U_r(q) \Rightarrow (f'(\Phi(u)))\Phi'(u) = 0 \Rightarrow \left\langle \nabla f_i(p), \frac{\partial \Phi}{\partial u_j}(q) \right\rangle = 0$$
für alle $i = 1, \ldots, M$, $j = 1, \ldots, L \quad \Rightarrow \nabla f_i(p) \in T^\perp$.
$\dim T^\perp = N - \dim T = N - L = M$, wegen linearer Unabhängigkeit von $\nabla f_1, \ldots, \nabla f_M$ bilden diese Vektoren eine Basis von $T^\perp$.

(iii). Nach Vorraussetzung hat $h(u) := g(\Phi(u))$ ein lokales Extremum in q (ohne Nebenbedingungen):
$$\Rightarrow 0 = h'(q) = g'(p)\Phi'(q) \Leftrightarrow \left\langle \nabla g(p), \frac{\partial \Phi}{\partial u_j}(q) \right\rangle = 0$$
Also $\nabla g(p) \in T^\perp \Rightarrow$ Behauptung (2.5.2.8).

Folgerung 2.57. Geometrische Deutung der Langrangeschen Multiplikatoren für $M = 1$ ($f = f_1$):
$\nabla g(p)$ muss parallel zu $\nabla f(p)$ sein, das heißt p muss ein Punkt sein, wo sich die Niveaulinien $g(x) = c_1$ und $f(x) = c_2$ berühren.

2.5.2.3 Lagrangefunktion

Praktische Vorgehensweise zur Bestimmung möglicher lokaler Extrema von g unter der Nebenbedingung $f(x) = 0$: Die Gleichungen (2.5.2.8) für die Lagrangemultiplikatoren und die Bedingung $p \in \mathcal{N}_0$ bilden die Gleichungen für einen stationären Punkt der *Lagrangefunktion*
$$L(x, \mu) := g(x) + \sum_{i=1}^{M} \mu_i f_i(x)$$
ohne Nebenbedingung ($L : \Omega \times \mathbb{R}^M \to \mathbb{R}$), denn
$$L'(p, \lambda) = 0 \Leftrightarrow \begin{cases} \dfrac{\partial L}{\partial x_j}(p, \lambda) = 0, j = 1, \ldots, N & \Leftrightarrow \dfrac{\partial g}{\partial x_j}(p) + \sum\limits_{i=1}^{M} \lambda_i \frac{\partial f_i}{\partial x_i} = 0 \Leftrightarrow (2.5.2.8) \\ \dfrac{\partial L}{\partial \mu_i}(p, \lambda) = 0, i = 1, \ldots M & \Leftrightarrow f_i(p) = 0 \Leftrightarrow \text{Nebenbedingungen.} \end{cases}$$

Beispiel 2.58. Bestimmung der Extrema der Funktion $g(x, y, z) = 5x + y - 3z$
auf den Schnitt der Ebene $x + y + z = 0$ mit der Kugeloberfläche $x^2 + y^2 + z^2 = 1$.
Nebenbedingung:
$$f(x, y, z) = 0 \Leftrightarrow \begin{cases} f_1(x, y, z) := x + y + z = 0 . \\ f_2(x, y, z) := x^2 + y^2 + z^2 - 1 = 0 . \end{cases}$$

$$f'(x,y,z) = \begin{pmatrix} \nabla f_1 \\ \nabla f_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}$$

hat den Rang 2 für alle $(x,y,z) \in \mathcal{N}_0 = \{(x,y,z) | \, f(x,y,z) = 0\}$.

Lagrangefunktion $\qquad L(x,y,z,\mu,\nu) := g(x,y,z) + \mu f_1(x,y,z) + \nu f_2(x,y,z)$

Löse das Gleichungssystem:

$$\begin{array}{lcrcrcrcrcl} L_x & = & 5 & + & \mu & + & 2\nu x & & & = & 0 \\ L_y & = & & & \mu & + & 2\nu y & & & = & 0 \\ L_z & = & -3 & + & \mu & + & 2\nu z & & & = & 0 \\ L_\mu & = & & & x & + & y & + & z & = & 0 \\ L_\nu & = & -1 & + & x^2 & + & y^2 & + & z^2 & = & 0 \end{array}$$

Man erhält $x = \pm\frac{\sqrt{2}}{2}$, $y = 0$, $z = \mp\frac{\sqrt{2}}{2}$, $\mu = -1$, $\nu = \mp 2\sqrt{2}$.

Da $\mathcal{N}_0$ beschränkt und abgeschlossen ist, ist $\mathcal{N}_0$ kompakt und die stetige Funktion g hat nach dem Satz vom Weierstrass 1.45 auf $\mathcal{N}_0$ ein Minimum und ein Maximum. Die Extremstellen können nur in $p_{1,2} = \pm\frac{1}{2}(\sqrt{2}, 0, \sqrt{2})$ liegen:

$$\Rightarrow \; g(p_1) = 4\sqrt{2} = \max_{x \in \mathcal{N}_0} g(x) \, , \; g(p_2) = -4\sqrt{2} = \min_{x \in \mathcal{N}_0} g(x) \, .$$

Kapitel 3

Integralrechnung im $\mathbb{R}^N$

3.1 Ergänzungen zum Riemann-Integral im $\mathbb{R}^1$

3.1.1 Uneigentliche Integrale

3.1.1.1 Integrale über unbeschränkte Intervalle

Definition 3.1. Existiert $F(b) := \int_a^b f(x)\mathrm{d}x$ für jedes $b > a$ und $\lim_{b\to\infty} F(b) := g$ so sagt man: Das *uneigentliche Integral* $\int_a^\infty f(x)\mathrm{d}x$ existiert oder konvergiert. Bezeichnung:

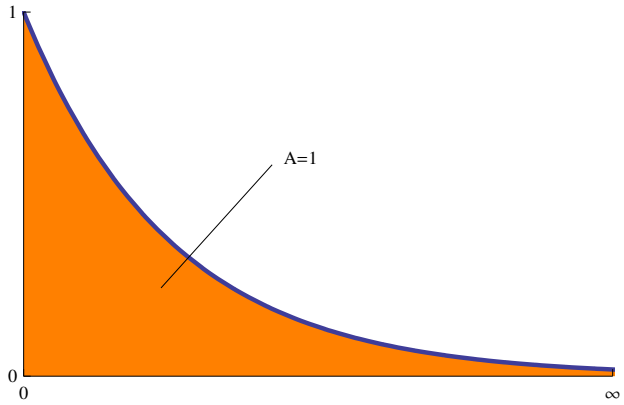
$$\int_a^\infty f(x)\mathrm{d}x := g = \lim_{b\to\infty} \int_a^b f(x)\mathrm{d}x\,.$$

Existiert g nicht, so heißt $\int_a^\infty f(x)\mathrm{d}x$ divergent. (Analog für $\int_{-\infty}^a f(x)\mathrm{d}x$)

Beispiel 3.2. i. Kosmische Fluchtgeschwindigkeit (vgl. Übungsaufgabe 19).

ii. Inhalt A einer unendlichen Fläche

$$\begin{aligned}\int_0^\infty \mathrm{e}^{-x}\mathrm{d}x &= \lim_{b\to\infty} \int_0^b \mathrm{e}^{-x}\mathrm{d}x \\ &= \lim_{b\to\infty} \left(-\mathrm{e}^{-x}\Big|_0^b\right) \\ &= \lim_{b\to\infty}(1-\mathrm{e}^{-b}) = 1\,.\end{aligned}$$



iii. $\int_{10}^{\infty} \cos x \mathrm{d}x$ divergiert, denn

$$\int_{10}^{b} \cos x \mathrm{d}x = \sin x|_{10}^{b} = \sin b - \sin 10$$

konvergiert nicht für $b \to \infty$.

3.1.1.2 Integrale über unbeschränkte Funktionen

Definition 3.3. Ist f unbeschränkt in $[a, b)$ und existiert jedoch $F(t) = \int_a^t f(x)\mathrm{d}x$ für alle $t \in [a, b)$, sowie $\lim_{t\to b-} F(b) =: h$, so sagt man: Das *uneigentliche Integral* $\int_a^b f(x)\mathrm{d}x$ existiert oder konvergiert mit

$$\int_a^b f(x)\mathrm{d}x = h = \lim_{t\to b-} \int_a^t f(x)\mathrm{d}x \,.$$

Andernfalls heißt $\int_a^b f(x)\mathrm{d}x$ divergent. (Analog für f unbeschränkt in $(a, b]$)

Beispiel 3.4.

$$\begin{aligned} \int_0^1 \frac{1}{\sqrt{1-x^2}}\mathrm{d}x &= \lim_{t\to 1-} \int_0^t \frac{1}{\sqrt{1-x^2}}\mathrm{d}x = \lim_{t\to 1-} \arcsin x\Big|_0^t \\ &= \lim_{t\to 1-} \arcsin t = \arcsin 1 = \underline{\underline{\frac{\pi}{2}}} \end{aligned}$$

3.1.1.3 Absolute Integrierbarkeit

Definition 3.5. Sei $I \subseteq \mathbb{R}$ ein offenes oder halboffenes Intervall. Dann heißt f auf I *absolut integrierbar*, falls $\int_I |f(x)|\mathrm{d}x$ konvergiert oder $|f(x)|$ auf I (uneigentlich) Riemann-integrierbar ist.

Wie bei unendlichen Reihen gilt für $f \in C(I)$:
Aus absoluter Konvergenz von $\int_I f(x)\mathrm{d}x$ folgt die Konvergenz von $\int_I f(x)\mathrm{d}x$ und

$$\left| \int_I f(x)\mathrm{d}x \right| \leq \int_I |f(x)|\mathrm{d}x \,.$$

Satz 3.6. *(Hauptkriterien für absolute Integrierbarkeit).*
$f \in C(I)$ ist absolut integrierbar auf I genau dann, wenn eine der beiden folgenden Bedingungen erfüllt ist:

a. *Beschränktheitskriterium:*

$$\int_\alpha^\beta |f(x)|\mathrm{d}x \le C \quad \text{für alle } [\alpha,\beta] \subseteq I$$

b. *Majorantenkriterium:*
Es existiert eine positive Funktion $g : I \to \mathbb{R}$ mit $|f(x)| \le g(x)$ für alle $x \in I$ und $\int_I g(x)\mathrm{d}x$ konvergiert.

Beweis. zu a.:

(i). ($\Rightarrow$): $\int_\alpha^\beta |f(x)|\mathrm{d}x \le \int_I |f(x)|\mathrm{d}x < \infty$ für alle $[\alpha,\beta] \subseteq I$,
falls f absolut Riemann-integrierbar auf I ist .

(ii). ($\Leftarrow$): $F(t) := \int_a^t |f(x)\mathrm{d}x$ mit $a \in I, t > a$ ($I \supseteq [a,t]$). Dann ist $F(t)$ monoton wachsend und $F(t) \le C$ für alle $t > a$. Somit folgt die Existenz von $\lim_{t\to b} F(t)$, wenn $I = [a,b), b \in \mathbb{R} \cup \{\infty\}$, das heißt f ist absolut integrierbar auf I.
(Andere Fälle für I analog)

zu b.:

(i). ($\Rightarrow$): Falls f absolut integrierbar ist, so wähle $g(x) := |f(x)|$ für $x \in I$.

(ii). ($\Leftarrow$)

$$\int_\alpha^\beta |f(x)|\mathrm{d}x \le \int_\alpha^\beta g(x)\mathrm{d}x \le \underbrace{\int_I g(x)\mathrm{d}x}_{:=C}$$

für alle $[\alpha,\beta] \subseteq I$. Die Behauptung folgt aus a.

Beispiel 3.7. Beispiele für Majoranten:

i.

$$\begin{aligned}\int_0^1 \frac{1}{x^\alpha}\mathrm{d}x &= \lim_{a\to 0+}\int_a^1 \frac{1}{x^\alpha}\mathrm{d}x = \lim_{a\to 0+}\frac{1}{1-\alpha}x^{1-\alpha}\Big|_a^1 \\ &= \lim_{a\to 0+}\left(\frac{1}{1-\alpha} - \frac{a^{1-\alpha}}{1-\alpha}\right) = \frac{1}{1-\alpha} \quad \text{für } \alpha < 1,\end{aligned}$$

divergent für $\alpha \ge 1$.

ii.
$$\int_1^\infty \frac{1}{x^\alpha}\mathrm{d}x = \lim_{b\to\infty}\int_1^b \frac{1}{x^\alpha}\mathrm{d}x = \frac{1}{\alpha-1} \quad \text{für } \alpha > 1,$$
divergent für $\alpha \le 1$.

Beispiel 3.8. Konvergiert $\int_1^\infty \frac{\ln x}{x}\mathrm{d}x$?

$$\text{Es ist } \ln x \ge \ln e = 1 \text{ für } x \ge e \Rightarrow \frac{\ln x}{x} \ge g(x) := \begin{cases} \frac{1}{x} & \text{für } x \ge e, \\ 0 & \text{für } 1 \le x \le e. \end{cases}$$

$\int_1^\infty g(x)\mathrm{d}x = \lim_{b\to\infty}\int_e^b \frac{1}{x}\mathrm{d}x$ divergiert wegen Beispiel 3.7.ii. Aus der logischen Umkehrung des Majorantenkriteriums folgt die Divergenz von $\int_1^\infty \frac{\ln x}{x}\mathrm{d}x$.

3.1.2 Parameterintegrale

Definition 3.9. Sei $f : Q \to \mathbb{R}$, $Q = [a,b]\times[c,d]$.
Existiert $F(y) := \int_a^b f(x,y)\mathrm{d}x$ für alle $y \in [c,d]$, so heißt $F(y)$ *Parameterintegral*.

Satz 3.10. a. Ist $f \in C(Q)$, so existiert $F(y)$ für jedes $y \in [c,d]$ und ist stetig.

b. Sind $f, \frac{\partial f}{\partial y} \in C(Q)$, so ist $F \in C^1((c,d))$ und
$$F'(y) = \frac{\mathrm{d}}{\mathrm{d}y}\int_a^b f(x,y)\mathrm{d}x = \int_a^b \frac{\partial f}{\partial y}(x,y)\mathrm{d}x\,.$$

Beweis. a. Die Existenz von $F(y)$ folgt aus der Stetigkeit von f bezüglich x und den Sätzen über Riemann-Integrale. Wir müssen zeigen:
Ist (y_n) eine beliebige Folge mit $y_n \in [c,d]$, $\lim\limits_{n\to\infty} y_n = y$, so muss $\lim\limits_{n\to\infty} F(y_n) = F(y)$ gelten.
Dazu definieren wir $h_n(x) := f(x,y_n)$, $h(x) := f(x,y)$ mit festem y. Somit folgt $\lim\limits_{n\to\infty} h_n(x) = h(x)$ für alle $x \in [a,b]$ ($f \in C(Q)$). $F(y_n) = \int_a^b h_n(x)\mathrm{d}x$ konvergiert gegen $F(y) = \int_a^b h(x)\mathrm{d}x$, falls h_n gleichmäßig gegen h konvergiert, das heißt falls h_n in $C([a,b])$ konvergiert.

$$D_n := \|h_n - h\|_{C([a,b])} = \max_{x\in[a,b]} |h_n(x) - h(x)| = |h_n(x_n) - h(x_n)| = |f(x_n,y_n) - f(x,y)|$$

für ein gewisses $x_n \in [a,b]$. (x_n) ist beschränkt. Somit existiert nach Satz von Bolzano-Weierstraß eine Teilfolge $x_{n_k} \subset (x_n)$ mit $\lim_{k\to\infty} x_{n_k} = x \in [a,b]$ und:

$$\begin{aligned} (x_{n_k}, y_{n_k}) \to (x,y) \quad &\Rightarrow \quad f(x_{n_k}, y_{n_k}) \to f(x,y) \\ (x_{n_k}, y) \to (x,y) \quad &\Rightarrow \quad f(x_{n_k}, y) \to f(x,y) \end{aligned}$$

$\Rightarrow D_{n_k} \to 0$ für $k \to \infty$. Dies gilt aber auch für jede beliebige Teilfolge von (D_n). Damit folgt $D_n \to 0$ für $n \to 0$.

b. Nach Definition von $F'(y)$ müssen wir $\lim\limits_{\eta\to 0}\frac{F(y+\eta)-F(y)}{k}$ untersuchen.

$$\frac{1}{\eta}(F(y+\eta)-F(y)) = \int_a^b \underbrace{\frac{f(x,y+\eta)-f(x,y)}{\eta}}_{=\frac{\partial f}{\partial y}(x,y+\theta\eta),\theta\in(0,1)}\mathrm{d}x = \int_a^b \frac{\partial f}{\partial y}(x,y+\theta\eta)\mathrm{d}x\,.$$

$$\text{Es ist}\quad \lim_{\eta\to 0}(y+\theta\eta) = y \;\Rightarrow\; \lim_{\eta\to 0}\frac{\partial f}{\partial y}(x,y+\theta\eta) = \frac{\partial f}{\partial y}(x,y)$$

$$\Rightarrow \lim_{\eta\to 0}\int_a^b \frac{\partial f}{\partial y}(x,y+\theta\eta)\mathrm{d}x = \int_a^b \lim_{\eta\to 0}\frac{\partial f}{\partial y}(x,y+\theta\eta)\mathrm{d}x = \int_a^b \frac{\partial f}{\partial y}(x,y)\mathrm{d}x$$

wegen a. Daraus folgt die Behauptung.

Beispiel 3.11. i. Sei $F(y) = \displaystyle\int_0^1 \frac{\mathrm{d}x}{2\mathrm{e}^x+\sin(xy)}$, gesucht: $\lim\limits_{y\to 0}F(y)$.

$f(x,y) = \dfrac{1}{2\mathrm{e}^x+\sin(xy)}$ ist stetig für alle $x\geq 0$, $y\in\mathbb{R}$, da $2\mathrm{e}^x+\sin(xy)\geq 2-1=1$.

Nach Satz 3.10.a ist

$$\lim_{y\to 0}F(y) = \lim_{y\to 0}\int_0^1 \frac{\mathrm{d}x}{2\mathrm{e}^x+\sin(xy)} = \int_0^1 \frac{\mathrm{d}x}{2\mathrm{e}^x+\sin 0} = -\frac{1}{2}\mathrm{e}^{-x}\Big|_0^1 = \underline{\underline{\frac{1}{2}-\frac{1}{2\mathrm{e}}}}$$

ii. Berechne für $y>0$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^1 \arctan\frac{x}{y}\mathrm{d}x &= \int_0^1 \frac{\partial}{\partial y}\left(\arctan\frac{x}{y}\right)\mathrm{d}x = \int_0^1 \frac{1}{1+\left(\frac{x}{y}\right)^2}\left(-\frac{x}{y^2}\right)\mathrm{d}x\\ &= -\int_0^1 \underbrace{\frac{x}{x^2+y^2}}_{=\frac{\frac{\partial}{\partial x}(x^2+y^2)}{2(x^2+y^2)}}\mathrm{d}x = -\frac{1}{2}\ln(x^2+y^2)\Big|_{x=0}^{x=1} = \underline{\underline{\frac{1}{2}\ln\frac{y^2}{1+y^2}}}\end{aligned}$$

Bemerkung 3.12. Ähnliche Resultate gelten auch für uneigentliche Riemann-Integrale über $I\subseteq\mathbb{R}$, wenn vorrausgesetzt wird
für a.: $f\in C(I\times[c,d])$ sowie zusätzlich $|f(x,y)|\leq g(x)$ für alle $x\in I$, $y\in[a,b]$ mit $\int_I g(x)\mathrm{d}y<\infty$,
bzw. für b.: $f, \frac{\partial f}{\partial x}\in C(I\times[c,d])$, sowie zusätzlich $|f(x,y)|, \left|\frac{\partial f}{\partial x}(x,y)\right|\leq g(x)$ für alle $x\in I$, $y\in[a,b]$ mit $\int_I g(x)\mathrm{d}y<\infty$.

3.2 Mehrfache Riemann-Integrale

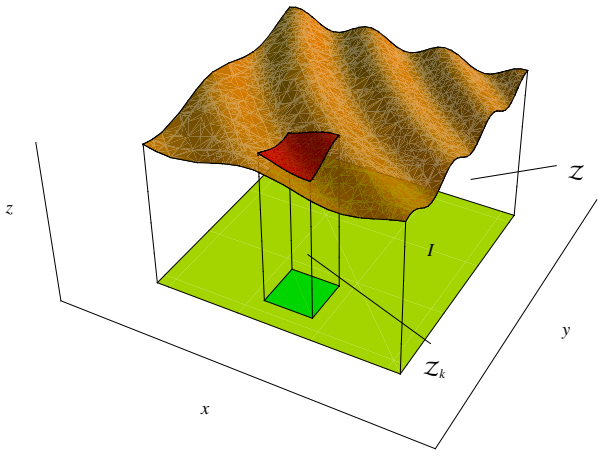
3.2.1 Einführung

3.2.1.1 Geometrische Motivation

$I = [a,b] \times [c,d]$, $f : I \to \mathbb{R}$, $f \geq 0$ auf I. Welches Volumen $|\mathcal{Z}|$ hat der „Zylinder“ $\mathcal{Z}$ mit Grundfläche I und Deckfläche

$$\{(x,y,z) | \, z = f(x,y), (x,y) \in I\} \; ?$$

Dazu wird $\mathcal{Z}$ durch genügend feine Teilzylinder $\mathcal{Z}_k$ angenähert. I wird in achsenparallele Teilrechtecke $I_1, \ldots, I_n$ zerlegt. Die Teilzylinder $\mathcal{Z}_1, \ldots, \mathcal{Z}_k$ haben eine ebene Deckfläche und eine Höhe $f(\xi_k)$ mit $\xi_k \in I_k \Rightarrow |\mathcal{Z}_k| = f(\xi_k)|I_k|$ und $\sum_{k=1}^n |\mathcal{Z}_k| = \sum_{k=1}^n f(\xi_k)|I_k|$ ist anschaulich eine Näherung zu $|\mathcal{Z}|$, falls die Zerlegung Z von I genügend fein ist.



3.2.1.2 Bezeichnungen und Definitionen

Definition 3.13. a. $I := [a_1,b_1] \times [a_2,b_2] \times \ldots \times [a_N,b_N] = \bigtimes_{i=1}^N [a_i,b_i]$ heißt *kompaktes N-dimensionales Intervall* oder *Quader*. $|I| = \prod_{i=1}^N (b_i - a_i)$.

b. Die *Zerlegung* Z ist das Kreuzprodukt $Z_1 \times \ldots \times Z_N$ von Zerlegungen $Z_i : a_i = x_{i_0} < x_{i_1} < \ldots < x_{i_{m_i}} = b_i$, $i = 1, \ldots, N$.
Die zu Z gehörigen Teilintervalle seien $I_1, \ldots, I_n$, wobei jedes Teilintervall I_k die Bauart $J_1 \times J_2 \times \ldots \times J_N$ hat mit $J_i = [x_{i_{l-1}}, x_{i_l}]$, $l \in \{1, \ldots, m_i\}$ und alle Möglichkeiten durchlaufen werden.

c. $T := \{\xi_k \in I_k | k = 1, \ldots, n\}$ ist die Menge der Zwischenpunkte und $S_f[Z,T] := \sum_{k=1}^n f(\xi_k)|I_k|$ ist die zugehörige Riemann-Summe.

d. Die *Feinheit* $\Delta(Z)$ der Zerlegung Z:

$$\Delta(Z) := \max_{i=1,\ldots,N} \Delta(Z_i) = \left(\max_{i=1,\ldots,N} \left(\max_{l=1,\ldots,m_i} (x_{i_l} - x_{i_{l-1}}) \right) \right) .$$

Definition 3.14. Sei I ein kompaktes N-dimensionales Intervall und $f : I \to \mathbb{R}$. Dann heißt f *Riemann-integrierbar auf* I, falls für jede Folge von Zerlegungen $Z^{(m)}$, $m = 1, 2, \ldots$ die Folge $\left(S_f[Z^{(m)}, T^{(m)}] \right)_{m=1}^{\infty}$ gegen ein- und denselben Grenzwert R konvergiert bei beliebiger Wahl von

$Z^{(m)}$ und $T^{(m)}$ mit $\lim_{m\to\infty} \Delta(Z^{(m)}) = 0$.
Symbolisch: $f \in R(I)$. Der Grenzwert R wird bezeichnet durch

$$R = \int_I f\mathrm{d}x = \int_I f(x)\mathrm{d}x = \int_I f(x_1, \ldots, x_N)\mathrm{d}x_1\mathrm{d}x_2 \ldots \mathrm{d}x_N$$

(*mehrfaches Riemann-Integral* oder *N-dimensionales Riemann-Integral* von f über I).

3.2.1.3 Eigenschaften

Durch wortwörtliche Übertragung aus dem $\mathbb{R}^1$ ergeben sich analoge Eigenschaften von $\int_I f\mathrm{d}x$ sowie entsprechende Integrierbarkeitskriterien außer solchen, welche auf den Monotonieeigenschaften des Integranden beruhen.

1. $\int_I f\mathrm{d}x$ ist linear im Integranden.
2. $\int_I f\mathrm{d}x$ ist additiv bezüglich I.
3. Es gilt die Dreiecksungleichung $\left|\int_I f\mathrm{d}x\right| \le \int_I |f|\mathrm{d}x$.
4. $f \in R(I) \Rightarrow f$ ist beschränkt auf I.
5. $f \in C(I) \Rightarrow f \in R(I)$.

3.2.2 Berechnung von mehrdimensionalen Riemann-Integralen

3.2.2.1 Der Satz von Fubini

Definition 3.15. Sei $I = [a,b] \times [c,d]$ und $f \in C(I)$. Dann existiert $F(y) := \int_a^b f(x,y)\mathrm{d}x$ für jedes $y \in [c,d]$. Nach Satz über die Stetigkeit von Parameterintegralen ist $F \in C([c,d])$ und es existiert

$$S := \int_c^d F(y)\mathrm{d}y = \int_c^d \left(\int_a^b f(x,y)\mathrm{d}x\right)\mathrm{d}y \qquad (\textit{iteriertes Integral}).$$

Fragestellung: Ist $S = \int_I f\mathrm{d}x\mathrm{d}y$?
Nach Definition von $S := \int_c^d F(y)\mathrm{d}y$ ist $S = \lim_{m\to\infty} S_F[Z_2^{(m)}, T_2^{(m)}]$, das heißt, für alle $\varepsilon > 0$ existiert $m_0(\varepsilon) \in \mathbb{N}$: $|S - S_F(Z_2^{(m)}, T_2^{(m)}) < \varepsilon$ für alle $m \ge m_0(\varepsilon)$. Mit $m = m_0(\varepsilon)$ folgt

$$\left|S - \sum_{k=1}^{n_m} F(\eta_k^{(m)})(y_k^{(m)} - y_{k-1}^{(m)}\right| < \varepsilon \;\Rightarrow\; \left|S - \int_a^b \underbrace{\left(\sum_{k=1}^{n_m} f(x, \eta_k^{(m)})(y_k^{(m)} - y_{k-1}^{(m)})\right)}_{:=\tilde{f}(x)} \mathrm{d}x\right| < \varepsilon$$

Hier ist $\tilde{f}(x) \in C([a,b])$ und abhängig von $Z_2^{(m)}, T_2^{(m)}$ sowie von ε. Nach Definition von $\int_a^b \tilde{f}(x)\mathrm{d}x$ ist

$$\begin{aligned}
\left|\int_a^b \tilde{f}(x)\mathrm{d}x - S_{\tilde{f}}(Z_1^{(l)}, T_1^{(l)})\right| &< \varepsilon \qquad \text{für alle } l \geq l_0(\varepsilon) \\
S_{\tilde{f}}(Z_1^{(l)}, T_1^{(l)}) &= \sum_{i=1}^{n_l}\sum_{k=1}^{n_m} f\left(\xi_i^{(l)}, \eta_k^{(m)}\right)\left(y_k^{(m)} - y_{k-1}^{(m)}\right)\left(x_i^{(l)} - x_{i-1}^{(l)}\right), \qquad \xi_i^l \in T_1^{(l)} \\
&= S_f[Z,T] \quad \text{mit} \quad Z = Z_1^{(l)} \times Z_2^{(m)}, T = T_1^{(l)} \times T_2^{(m)} \\
\Rightarrow \left|S - S_f[Z,T]\right| &\leq \underbrace{\left|S - \int_a^b \tilde{f}(x)\mathrm{d}x\right|}_{<\varepsilon} + \underbrace{\left|\int_a^b \tilde{f}(x)\mathrm{d}x - S_f[Z,T]\right|}_{<\varepsilon} < 2\varepsilon
\end{aligned}$$

Da f Riemann-integrierbar ist auf I, konvergiert $S_f[Z,T]$ für $l,m \to \infty$ gegen $\int_I f\mathrm{d}x$. Somit ist $S = \int_I f\mathrm{d}x$. Allgemein lässt sich analog zeigen:

Satz 3.16. (*Satz von Fubini*, G. Fubini (1879-1943))

a. Ist $I = [a,b] \times [c,d]$, $f \in C(I)$, so gilt

$$\int_I f\mathrm{d}x = \int_a^b \left(\int_c^d f(x,y)\mathrm{d}y\right)\mathrm{d}x = \int_c^d \left(\int_a^b f(x,y)\mathrm{d}x\right)\mathrm{d}y$$

b. Ist $I = \bigtimes_{k=1}^N [a_k, b_k]$, $f \in C(I)$, so gilt:

$$\begin{aligned}
\int_I f\mathrm{d}x &= \int_I f(x_1,\ldots,x_N)\mathrm{d}x_1,\ldots,\mathrm{d}x_N \\
&= \int_{a_N}^{b_N}\left(\ldots\left(\int_{a_2}^{b_2}\left(\int_{a_1}^{b_1} f(x_1,\ldots,x_N)\mathrm{d}x_1\right)\mathrm{d}x_2\right)\ldots\right)\mathrm{d}x_N
\end{aligned}$$

Dabei ist die Integrationsreihenfolge egal.

Beispiel 3.17. $\int_I \frac{z}{(x+y)^2}\mathrm{d}x\mathrm{d}y\mathrm{d}z$ mit $I = [1,2]\times[1,2]\times[0,1]$.

$$\begin{aligned}
\int_I \frac{z}{(x+y)^2} &= \int_1^2\left(\int_1^2\left(\int_0^1 \frac{z\mathrm{d}z}{(x+y)^2}\right)\mathrm{d}x\right)\mathrm{d}y = \int_1^2\left(\int_1^2\Bigg(\underbrace{\frac{z^2}{2(x+y)^2}\Big|_{z=0}^{z=1}}_{\frac{1}{2(x+y)^2}}\Bigg)\mathrm{d}x\right)\mathrm{d}y \\
&= \frac{1}{2}\int_1^2\Bigg(\underbrace{-\frac{1}{x+y}\Big|_{x=1}^{x=2}}_{=\frac{1}{y+1}-\frac{1}{y+2}}\Bigg)\mathrm{d}y = \frac{1}{2}(\ln|y+1| - \ln|y+2|)\Big|_{y=1}^{y=2} \\
&= \frac{1}{2}(\ln 3 - \ln 4 - \ln 2 + \ln 3) = \frac{1}{2}(\ln 9 - \ln 8) = \underline{\underline{\ln\left(\frac{3}{4}\sqrt{2}\right)}}.
\end{aligned}$$

3.2.3 Integration über beschränkte Mengen

3.2.3.1 Definition und Eigenschaften

Definition 3.18. Sei f auf $B \subset \mathbb{R}^N$ definiert, B beschränkt, so setzt man

$$f_B(x) = \begin{cases} f(x) & x \in B \\ 0 & x \in \mathbb{R}^N \setminus B \end{cases}$$

f ist *Riemann-integrierbar auf* B (symbolisch: $f \in R(B)$) genau dann, wenn $f_B \in R(I)$ für ein N-dimensionales Intervall $I \supseteq B$.

$$\int_B f \mathrm{d}x := \int_I f_B \mathrm{d}x$$

Eigenschaften und Rechenregeln übertragen sich von $\int_I f\mathrm{d}x$ auf $\int_B f\mathrm{d}x$.

3.2.3.2 Nullmengen

Für welche f, B existiert $\int_B f\mathrm{d}x$? Dazu benutzen wir

Satz 3.19. (*Lebesguesches Integrierbarkeitskriterium*, H. Lebesgue (1875 - 1941))
$f \in R(I)$ genau dann, wenn f beschränkt ist auf I und stetig auf $I \setminus \mathcal{N}$, wobei $\mathcal{N}$ eine N-dimensionale Nullmenge ist.

Definition 3.20. $\mathcal{N}$ heißt *N-dimensionale Nullmenge* oder $\mathcal{N}$ ist eine *Menge vom Maß Null* ($|\mathcal{N}| = 0$), wenn zu jedem $\varepsilon > 0$ höchstens abzählbar unendlich viele kompakte Intervalle $I_1, I_2, \ldots$ existieren mit $\bigcup_{k=1}^{\infty} I_k \supseteq \mathcal{N}$ und $\sum_{k=1}^{\infty} |I_k| \leq \varepsilon$.

Beispiel 3.21 (Beispiele für Nullmengen, vgl. auch Heuser II, Seite 463ff.)**.**

i. Endliche oder abzählbar unendliche Teilmengen des $\mathbb{R}^N$ sind Nullmengen.

ii. Jede Hyperebene $x_j = c$ mit $j \in \{1, \ldots, N\}$ ist Nullmenge, denn im Fall $N = 2$ betrachten wir die x-Achse (Hyperebene $y = 0$). Überdecke x-Achse durch Rechtecke I_k:

$$\begin{aligned} I_k &= [-k, k] \times \left[-\frac{\varepsilon}{2^{k+2}k}, \frac{\varepsilon}{2^{k+2}k} \right], k \in \mathbb{N}, \varepsilon > 0 \\ &\Rightarrow \bigcup_{k=1}^{\infty} I_k \supset \{(x, 0) | x \in \mathbb{R}\} \\ \sum_{k=1}^{\infty} (I_k) &= \sum_{k=1}^{\infty} 2k \frac{2\varepsilon}{2^{k+2}k} = \varepsilon \sum_{k=1}^{\infty} \frac{1}{2^k} = \varepsilon \left(\frac{1}{1 - \frac{1}{2}} - 1 \right) = \underline{\underline{\varepsilon}} \end{aligned}$$

iii. Jeder Graph $G_f = \{(x_1, \ldots, x_{N-1}, f(x_1, \ldots, x_{N-1})) | (x_1, \ldots, x_{N-1}) \in \Omega\}$ mit $f \in C(\Omega), \Omega \subseteq \mathbb{R}^{N-1}$ ist eine N-dimensionale Nullmenge.

iv. Sei $\Omega \subseteq \mathbb{R}^N$ ein Gebiet, $\mathcal{N} \subset \Omega$ eine Nullmenge. Dann ist $g(\mathcal{N})$ wieder eine Nullmenge, falls $g \in C^1(\Omega, \mathbb{R}^M)$, $M \geq N$.

v. Jede Umgebung $U_\delta(x_0) \subset \mathbb{R}^N$ ist keine Nullmenge, $\delta > 0$

Standardfall von Nullmengen im $\mathbb{R}^N$:
Sei $B \subset \mathbb{R}^N$, B beschränkt und bestehe ∂B aus den Graphen $G(f_1), \ldots, G(f_n)$, das heißt, $\partial B = \bigcup_{i=1}^n G(f_i)$ mit $f_i \in C(A_i)$, $A_i \subset \mathbb{R}^{N-1}$, so ist $|\partial B| = 0$.

Folgerung 3.22. Ist $f \in C\left(\overline{B}\right), |\partial B| = 0$, so ist f_B stetig auf $\mathbb{R}^N \setminus \partial B$. f_B ist beschränkt, da f beschränkt auf B ist. Somit folgt $f_B \in R(I), I \supseteq B$. Daraus ergibt sich:
Stetige Funktionen auf abgeschlossenen beschränkten Mengen des $\mathbb{R}^N$ mit einer Nullmenge als Rand sind Riemann-integrierbar.

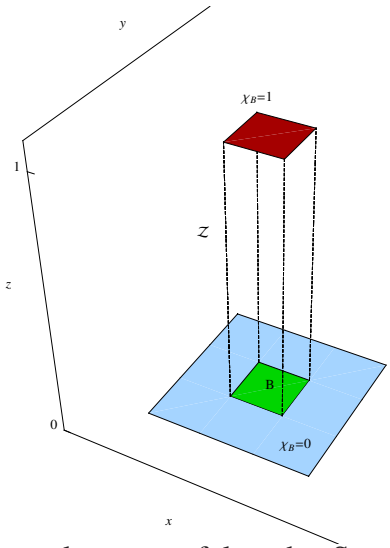
3.2.3.3 Der Satz von Cavalieri

Sei $B \subset \mathbb{R}^N$ beschränkt, $|\partial B| = 0$, $f \equiv 1$ auf B.

$$\int_B 1 \mathrm{d}x = \int_I \chi_B(x) \mathrm{d}x = \int_{\mathbb{R}^N} \chi_B(x) \mathrm{d}x =: |B| \quad (N\text{-}\textit{dimensionaler Inhalt} \text{ von } B)$$

mit der charakteristischen Funktion $\chi_B = \begin{cases} 1 & x \in B\,, \\ 0 & x \in \mathbb{R}^N \setminus B\,. \end{cases}$

Das Volumen des „Zylinders“ $\mathcal{Z}$, begrenzt durch Grund- und Deckfläche B der Höhe 1 berechnet sich elementargeometisch einerseits mit $|\mathcal{Z}| = 1 \cdot |B|$, andererseits ist $|\mathcal{Z}| = \int_I \chi_B(x,y) \mathrm{d}x\mathrm{d}y = \int_B \mathrm{d}x\mathrm{d}y$.



Zur Berechnung des Volumens beschränkter Körper kann man folgenden Satz benutzen:

Satz 3.23. (*Satz von Cavalieri*, B. Cavalieri (1598 - 1647))
Sei $B \subset \mathbb{R}^N$ beschränkt, $|\partial B| = 0$ und liege B zwischen den Ebenen $x_N = a$ und $x_N = b$. Ferner besitze der *Schnitt* $Q(\xi) = \{x \in B | x_N = \xi\}$ den $(N-1)$-dimensionalen Inhalt $q(\xi)$, $a \leq \xi \leq b$, so ist $|B| = \int_a^b q(\xi) \mathrm{d}\xi$.

Beweis. Sei $B \subseteq I = J \times [a,b]$. J sei $(N-1)$-dimensionales kompaktes Intervall.

$$|B| = \int_B \mathrm{d}x = \int_I \chi_B(x)\mathrm{d}x = \int_a^b \Bigg(\underbrace{\int_J \chi_B(x_1,\ldots,x_{N-1},\xi)\mathrm{d}x_1,\ldots,\mathrm{d}x_{N-1}}_{=\int_{Q(\xi)} 1\mathrm{d}x_1,\ldots,\mathrm{d}x_{N-1}=q(\xi)} \Bigg) \mathrm{d}\xi$$

3.2.3.4 Volumen von Rotationskörpern $\mathcal{K} \subset \mathbb{R}^3$

Wähle ein Koordinatensystem so, dass die Rotationsachse mit der z-Achse zusammenfällt. $\mathcal{K}$ entsteht somit durch Rotation der Fläche $\mathcal{F}$ um die z-Achse.

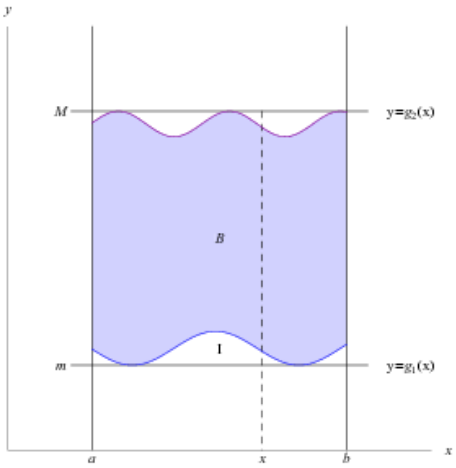
$$\begin{aligned} \mathcal{F} &:= \{(x,0,z) |\, 0 \le x \le r(z),\ a \le z \le b\}\,, \qquad r \in C([a,b]) \\ \Rightarrow\ \mathcal{K} &= \left\{ (x,y,z) \,\middle|\, \sqrt{x^2+y^2} \le r(z),\ a \le z \le b \right\}\,, \\ |\mathcal{K}| &= \int_a^b q(z)\mathrm{d}z = \pi \int_a^b r^2(z)\mathrm{d}z\,, \end{aligned}$$

$q(z) = \pi r(z)^2$, da $Q(z)$ ein Kreis ist um $(0,0,z)$ mit Radius $r(z)$ ist.

Beispiel 3.24. Volumen der Kugel um 0 mit Radius $\rho > 0$.
$\mathcal{K} = \left\{ (x,y,z) \,\middle|\, x^2+y^2+z^2 \le \rho^2 \right\}$. Hier: $r(z) = \sqrt{\rho^2 - z^2}$, $z \in [-\rho,\rho]$.

$$|\mathcal{K}| = \pi \int_{-\rho}^{\rho} \left(\sqrt{\rho^2 - z^2} \right)^2 \mathrm{d}z = \pi \left(\rho^2 z - \frac{z^3}{3} \right) \Bigg|_{-\rho}^{\rho} = 2\pi \frac{2\rho^3}{3} = \underline{\underline{\frac{4\pi}{3}\rho^3}}\,.$$

3.2.4 Integration über Normalbereiche



Definition 3.25. $B \subset \mathbb{R}^2$ heißt *Normalbereich bezüglich der x-Achse*, wenn

$$B = \{(x,y) \in \mathbb{R}^2 |\, a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$$

mit $g_1, g_2 \in C([a,b])$, $g_1 \le g_2$
($\Rightarrow\ |\partial B| = 0$, B kompakt).

Für $f \in C(B)$ gilt $f \in R(B)$ und

$$\int_B f\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\mathrm{d}y \right) \mathrm{d}x\,,$$

denn

$$\int_B f\,\mathrm{d}x\mathrm{d}y = \int_I f_B\,\mathrm{d}x\mathrm{d}y = \int_a^b \left(\int_m^M f_B(x,y)\mathrm{d}y\right)\mathrm{d}x = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\mathrm{d}y\right)\mathrm{d}x$$

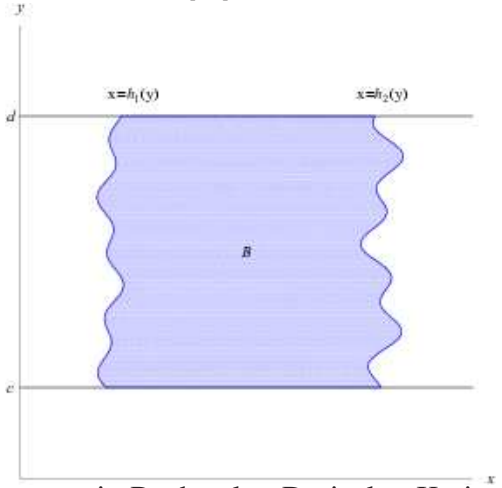
mit $I \supseteq B$, z.B. $I = [a,b]\times[m,M]$, $m = \min_{x\in[a,b]} g_1(x)$, $M = \max_{x\in[a,b]} g_2(x)$.

Definition 3.26. Hat B die Gestalt

$$B = \left\{(x,y)\in\mathbb{R}^2 \,\middle|\, c\le y\le d, h_1(y)\le x\le h_2(y)\right\},$$

so heißt *B Normalbereich bezüglich der y-Achse* und

$$\int_B f\,\mathrm{d}x\mathrm{d}y = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\mathrm{d}x\right)\mathrm{d}y\,.$$



Man beachte, dass viele elementargeometrische Figuren wie Rechtecke, Dreiecke, Kreise u.ä. Normalbereiche bezüglich beider Achsen sind.

Beispiel 3.27. $\int_B (x+y^2)\mathrm{d}x\mathrm{d}y$, wobei B ein Dreieck mit den Eckpunkten $(0,0),(1,0),(0,2)$ ist.

1. Möglichkeit: $B = \{(x,y)|\, 0\le x\le 1, 0\le y\le 2-2x\}$

$$\begin{aligned}\int_B (x+y^2)\mathrm{d}x\mathrm{d}y &= \int_0^1 \left(\int_0^{2-2x}(x+y^2)\mathrm{d}y\right)\mathrm{d}x = \int_0^1 \left(xy+\frac{y^3}{3}\right)\Big|_{y=0}^{y=2-2x}\\ &= \int_0^1 \left(2x-2x^2+\frac{8}{3}(1-x)^3\right)\mathrm{d}x\\ &= x^2-\frac{2}{3}x^3-\frac{8}{12}(1-x)^4\Big|_0^1 = \frac{1}{3}+\frac{8}{12} = \frac{1}{3}+\frac{2}{3} = \underline{\underline{1}}\,.\end{aligned}$$

2. Möglichkeit: $B = \left\{(x,y)\,\middle|\, 0\le y\le 2, 0\le x\le 1-\frac{y}{2}\right\}$

$$\begin{aligned}\int_B (x+y^2)\mathrm{d}x\mathrm{d}y &= \int_0^2 \left(\int_0^{1-\frac{y}{2}}(x+y^2)\mathrm{d}x\right)\mathrm{d}y = \int_0^2 \left(\frac{x^2}{2}+xy^2\right)\Big|_0^{1-\frac{y}{2}}\mathrm{d}y\\ &= \int_0^2 \left(\frac{1}{8}(2-y)^2+y^2-\frac{y^3}{2}\right)\mathrm{d}y = \frac{1}{24}(y-2)^3+\frac{y^3}{3}-\frac{y^4}{8}\Big|_{y=0}^{y=2}\\ &= \frac{8}{3}-\frac{16}{8}-\left(-\frac{8}{24}\right) = \frac{8}{3}+\frac{8}{24}-2 = \frac{8}{3}+\frac{1}{3}-2 = \frac{9}{3}-2 = \underline{\underline{1}}\,.\end{aligned}$$

Bemerkung 3.28. Die Verallgemeinerung auf Normalbereiche im $\mathbb{R}^N$ ist möglich, zum Beispiel im $\mathbb{R}^3$:

$$B = \left\{(x,y,z)\in\mathbb{R}^3 \,\middle|\, (x,y)\in A, g_1(x,y)\le z\le g_2(x,y)\right\}$$

ist ein *Normalbereich bezüglich der x-y-Ebene*. Dabei sei $A \subset \mathbb{R}^2$, $|\partial A|_2 = 0$, $g_1, g_2 \in C(A)$. Dann ist $|\partial B|_3 = 0$ und für $f \in C(B)$ ist

$$\int_B f \mathrm{d}x\mathrm{d}y\mathrm{d}z = \int_A \left(\int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\mathrm{d}z \right) \mathrm{d}x\mathrm{d}y \; .$$

3.3 Substitutionsregel

3.3.1 Der Transformationssatz

Im $\mathbb{R}^1$ erfolgt die Substitution $x = g(t)$ nach der Formel $\int_a^b f(x)\mathrm{d}x = \int_\alpha^\beta f(g(t)) \underbrace{\frac{\mathrm{d}x}{\mathrm{d}t}}_{=g'(t)} \mathrm{d}t$,

wobei $g(\alpha) = a, g(\beta) = b, g'(t) \neq 0, g \in C^1([\alpha,\beta]), f \in C(\mathbb{R})$. Mit $B = [\alpha,\beta]$ gilt:

$$g' > 0: \quad \Rightarrow \quad a < b: \int_{g(B)} f\mathrm{d}x = \int_a^b f\mathrm{d}x = \int_B f(g(t))|g'(t)|\mathrm{d}t \; ,$$

$$g' < 0: \quad \Rightarrow \quad a > b: \int_{g(B)} f\mathrm{d}x = \int_b^a f\mathrm{d}x = -\int_a^b f\mathrm{d}x = -\int_\alpha^\beta f(g(t))g'(t)\mathrm{d}t = \int_B f(g(t))|g'(t)|\mathrm{d}t \; .$$

Das N-dimensionale Analogon ist folgender Satz:

Satz 3.29. *(Transformationssatz)*
Seien $\Omega, \tilde{\Omega} \subseteq \mathbb{R}^N$ Gebiete, $g : \Omega \to \tilde{\Omega}$ bijektiv und g, g^{-1} C^1-Abbildungen. Dann gilt für jedes beschränkte $B \subset \Omega$ mit $|\partial B| = 0$ die Formel

$$\int_{g(B)} f(x)\mathrm{d}x = \int_B f(g(t)) \, |\det g'(t)| \, \mathrm{d}t$$

für $f \in C(\tilde{\Omega})$.

Bemerkung 3.30. Benutzt man bei der Variablensubstitution $x = g(t)$ die Bezeichnungen:

$$g'(t) = \left(\frac{\partial x_j}{\partial t_j} \right)_{i,j=1,\ldots,N} = \frac{\partial(x_1,\ldots,x_N)}{\partial(t_1,\ldots,t_N)} \qquad \textit{(Funktionalmatrix)} \text{ und}$$

$$\det g'(t) = \det\left(\frac{\partial x_i}{\partial t_j} \right) = \frac{D(x_1,\ldots,x_N)}{D(t_1,\ldots,t_N)} \qquad \textit{(Funktionaldeterminante)},$$

so ergibt sich formal

$$\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_N = \left| \frac{D(x_1,\ldots,x_N)}{D(t_1,\ldots,t_N)} \right| \mathrm{d}t_1,\ldots,\mathrm{d}t_N \; .$$

Wegen $g^{-1}(g(t)) = t$ für alle $t \in B$ folgt nach Kettenregel

$$g^{-1\prime}(g(t)) \cdot g'(t) = E_N \;\Rightarrow\; \text{Matrix } g'(t) \text{ regulär} \;\Leftrightarrow\; \det g'(t) \neq 0 \; .$$

Eine Abbildung g mit obigen Eigenschaften heißt auch C^1-*Diffeomorphismus*.

Der Beweis des Transformationssatzes ist sehr aufwendig (siehe Heuser Band II, S. 475 ff.).

Beweis (Skizzierung der Beweisidee im $\mathbb{R}^2$). Sei $\Omega, \tilde{\Omega} \subseteq \mathbb{R}^2$ und $B \subset \Omega, \tilde{B} = g(B), g : \Omega \to \tilde{\Omega}$ ein C^1-Diffeomorphismus, weiterhin

$$(u, v) \in B,\ g(u, v) = (F(u, v), G(u, v)) = (x, y) \in \tilde{B}$$

und $\tilde{I}_k = g(I_k)$ (krummliniges Parallelogramm). I_k habe Eckpunkte $\alpha_k, \beta_k, \gamma_k, \delta_k$, dabei

$$\alpha_k = (u_k, v_k),\ \beta_k = (u_k + \Delta u_k, v_k),\ \gamma_k = (u_k + \Delta u_k, v_k + \Delta v_k),\ \delta_k = (u_k, v_k + \Delta v_k).$$

$\tilde{I}_k$ hat dann Eckpunkte $g(\alpha_k), g(\beta_k), g(\gamma_k), g(\delta_k)$, $|I_k| = \Delta u_k \Delta v_k$. $|\tilde{I}_k|$ hat annähernd den gleichen Flächeninhalt des geradlinigen Parallelogramms mit gleichen Eckpunkten wie $\tilde{I}_k$.

$$\Rightarrow\ |\tilde{I}_k| \approx \left| \det \begin{pmatrix} g(\beta_k) - g(\alpha_k) \\ g(\delta_k) - g(\alpha_k) \end{pmatrix} \right| .$$

Nach Taylorschem Lehrsatz ist

$$\begin{aligned} & g_1(\beta_k) - g_1(\alpha_k) = F(u_k + \Delta u_k, v_k) - F(u_k, v_k) = F_u(u_k, v_k)\Delta u_k + R \\ \Rightarrow\ & g_1(\beta_k) - g_1(\alpha_k) \approx F_u(u_k, v_k)\ \Rightarrow\ |\tilde{I}_k| \approx \left| \begin{vmatrix} \frac{\partial g_1}{\partial u}(\alpha_k) & \frac{\partial g_2}{\partial u}(\alpha_k) \\ \frac{\partial g_1}{\partial v}(\alpha_k) & \frac{\partial g_2}{\partial v}(\alpha_k) \end{vmatrix} \Delta u_k \Delta v_k \right| = |\det g'(\alpha_k)||I_k| \\ \Rightarrow\ & \int_{\tilde{B}} f \mathrm{d}x\mathrm{d}y = \int_{\mathbb{R}^2} f_{\tilde{B}} \mathrm{d}x\mathrm{d}y \approx \sum_k f(g(\alpha_k))|\tilde{I}_k| \approx \sum_k f(g(a_k))|I_k||\det g'(a_k)| \end{aligned}$$

(Riemannsche Summe zu $\int_B f(g(u, v))|\det g'|\mathrm{d}u\mathrm{d}v$).

3.3.2 Beispiele

Beispiel 3.31. i. Volumen eines N-dimensionalen *Parallelepipeds* (oder *Parallelflachs*)

$$\mathcal{P} = \left\{ x \in \mathbb{R}^N \,\middle|\, x = a + \sum_{i=1}^N t_i a_i, t_i \in [0, 1] \right\} ,$$

$\mathcal{P} = g\left([0, 1]^N\right)$ mit $g(t) = a + \sum_{i=1}^N t_i a_i$, $t = (t_1, \ldots, t_N)$, $t \in [0, 1]^N$, wobei $a, a_1, \ldots, a_N$ gegebene Vektoren des $\mathbb{R}^N$ sind.

$$\begin{aligned} \left|\mathcal{P}\right| &= \int_{\mathcal{P}} 1 \mathrm{d}x = \int_{\mathbb{R}^N} \chi_{\mathcal{P}}(x) \mathrm{d}x = \int_{[0,1]^N} |\det g'(t)| \mathrm{d}t \\ &= \int_{[0,1]^N} |\det A| \mathrm{d}t = |\det A| \left|[0, 1]^N\right| = \underline{\underline{|\det A|}} , \end{aligned}$$

weil $g'(t) = \left(\frac{\partial g_i}{\partial t_j} \right)_{i,j=1,\ldots,N} = (a_1, \ldots, a_N) := A.$

ii. Integrale über kugelsymmetrische Potentiale f.

$f = f(x)$ heißt *kugelsymmetrisches Potential*, wenn es $V \in C((0,\infty))$ mit $f(x) := V(\|x\|)$ für $x \in \mathbb{R}^3 \setminus \{0\}$ gibt. Berechnung von

$$\int_{\rho\le\|x\|\le R} f\,\mathrm{d}x = \int_{U_R(0)\setminus U_\rho(0)} f\,\mathrm{d}x\,.$$

$\overline{U_R(0)} \setminus U_\rho(0) = g(B)$ mit $B = [\rho, R]\times[0,2\pi)\times[-\frac{\pi}{2},\frac{\pi}{2}]$ und

$$x = g(r,\phi,\theta) \;\Leftrightarrow\; \begin{cases} x_1 &= r\cos\phi\cos\theta,\ 0\le r\le R\\ x_2 &= r\sin\phi\cos\theta,\ \phi\in[0,2\pi]\\ x_3 &= r\sin\theta,\ \theta\in[-\frac{\pi}{2},\frac{\pi}{2}]\end{cases}$$

$$\begin{aligned}\int_{\rho\le\|x\|\le R} f\,\mathrm{d}x &= \int_B f(g(r,\phi,\theta))\ \underbrace{|\det g'(r,\phi,\theta)|}_{=|r^2\cos\theta|\text{ (s. Übungsaufgabe 5.a)}}\ \mathrm{d}r\mathrm{d}\phi\mathrm{d}\theta\\ &= \int_0^{2\pi}\left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\left(\int_\rho^R V(r)r^2\cos\theta\mathrm{d}r\right)\mathrm{d}\theta\right)\mathrm{d}\phi\\ &= \left(\int_\rho^R V(r)r^2\mathrm{d}r\right)\underbrace{\left(\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\cos\theta\mathrm{d}\theta\right)}_{=2}\underbrace{\left(\int_0^{2\pi}\mathrm{d}\phi\right)}_{=2\pi} = 4\pi\int_\rho^R V(r)r^2\mathrm{d}r\end{aligned}$$

Ist $V(r)r^2$ zusätzlich beschränkt auf $(0,R]$, so existiert

$$\int_0^R V(r)r^2\mathrm{d}r = \lim_{\rho\to 0+}\int_\rho^R V(r)r^2\mathrm{d}r \quad\text{und}\quad \int_{\|x\|\le R} f\,\mathrm{d}x = 4\pi\int_0^R V(r)r^2\mathrm{d}r\,.$$

$$\text{Speziell ist}\quad \int_{\|x\|\le R}\frac{\mathrm{d}x}{\|x\|^2} = 4\pi\int_0^R\frac{1}{r^2}r^2\mathrm{d}r = \underline{\underline{4\pi R}}\,.$$

Bemerkung 3.32. Ebenso wie $\int_{\|x\|\le R} f\mathrm{d}x$ als uneigentliches mehrfaches Riemann-Integral aufgefasst wird, erklärt man $\int_B f\mathrm{d}x$ durch Grenzübergang, wenn B eine unbeschränkte Menge des $\mathbb{R}^N$ ist.