Vorlesungsskript zu Analysis II *

Jens Kubieziel

Sommersemester 2004

*$LastChangedRevision : 137$ vom $LastChangedDate : 2005 - 09 - 1722 : 58 : 51 + 0200(Sat, 17Sep2005)$

Inhaltsverzeichnis

1 Normierte Räume 5

1.1 Definition und Beispiele für normierte Räume . . . 5
1.2 Endlichdimensionale normierte Räume . . . 8
1.3 Stetige lineare Abbildungen (Operatoren) zwischen normierten Räumen . 10

2 Differentiation 12

2.1 Differentiation - Beispiele und Rechenregeln . . . 12
2.2 Differentiation von $\mathbb{R}^n$-$\mathbb{R}^m$-Funktionen . . . 14
2.2.1 Partielle Ableitung von $\mathbb{R}^n$-$\mathbb{R}$-Funktionen . . . 14
2.2.2 Darstellung der Ableitung von $\mathbb{R}^n$-$\mathbb{R}^m$-Funktionen . . . 15
2.3 Mittelwert, Taylorscher Satz, lokale Extrema . . . 18
2.3.1 Mittelwertsätze . . . 18
2.3.2 Taylorsche Formel . . . 21
2.4 Implizite Funktionen . . . 28
2.4.1 Umkehrsatz . . . 31
2.5 Extrema mit Nebenbedingungen . . . 33
2.5.1 Lagrange Multiplikatorenmethode . . . 33
2.5.2 Praktische Lösungen von Extremalaufgaben . . . 34
2.6 Geometrische Begriffe . . . 37
2.6.1 Geometrische Interpretation . . . 38
2.6.2 Flächen und Tangentialebenen . . . 38
2.6.3 Niveauflächen . . . 39
2.6.4 Niveauflächen und Tangentialebenen . . . 40

3 Gewöhnliche Differentialgleichungen 41

3.1 Einführung . . . 41
3.2 Reduktion von Differentialgleichungssystemen höherer Ordnung auf Differentialgleichungssysteme 1. Ordnung . . . 43
3.2.1 Beispiele für Differentialgleichungen, die auf $y' = f(x)g(y)$ zurückgeführt werden können . . . 45
3.3 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten . . . 45
3.3.1 Definition . . . 45
3.3.2 Operatorfunktion (Matrixfunktionen) . . . 46
3.4 Der Existenz- und Eindeutigkeitssatz von Picard-Lindelöf . . . 54
3.4.1 Lipschitzbedingung . . . 54
3.4.2 Satz von Picard-Lindelöf . . . 56

3.4.3 Das Picard-Lindelöfsche Iterationsverfahren anhand eines Beispiels 58
3.5 Lösungstheorie linearer Differentialgleichungsysteme 1. Ordnung 58
3.5.1 Existenz- und Eindeutigkeitssatz 58
3.5.2 Lösungstheorie 59
3.5.3 Lineare Abhängigkeit und Unabhängigkeit 59
3.5.4 Die inhomogene Differentialgleichung $y' = A(x)y + b(x)$ 61
3.6 Lineare Differentialgleichungen n-ter Ordnung 63
3.6.1 Lösungstheorie 63
3.6.2 Reduktion der Ordnung einer Differentialgleichung $L(y) = 0$ bei Kenntnis einer Lösung 66
3.6.3 Praktisches Vorgehen beim Lösen von linearen DGL 2. Ordnung mit nichtkonstanten Koeffizienten 66
3.6.4 Lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten 67
3.6.5 Praktische Bestimmung von Φ 71

1 Normierte Räume

1.1 Definition und Beispiele für normierte Räume

Definition 1.1.1. *Sei E ein linearer Raum (Vektorraum) über $\mathbb{K}$*[1]*. Eine Funktion $||\cdot|| : E \to [0, \infty[$ heißt* Norm *auf E, wenn für beliebige $x, y \in E$ und $\alpha \in \mathbb{K}$ gilt:*

(N1) $||x|| \geq 0 \wedge ||x|| = 0 :\Leftrightarrow x = 0$

(N2) $||\alpha x|| = |\alpha| \cdot ||x||$ *Homogenität*

(N3) $||x + y|| \leq ||x|| + ||y||$ *Dreiecksungleichung*

Das Paar $[E, ||\cdot||]$ heißt normierter Raum*.*

Bemerkung. Durch den Ansatz $d(x, y) := ||x - y||$ wird auf E eine Metrik erklärt. Damit ist E insbesondere ein metrischer Raum. Begriffe, wie konvergente Folge, Cauchyfolge, offene/abgeschlossene Mengen etc. gelten auch für normierte Räume. z.B.: $(x_n) \subset E$ konvergiert in $E : \exists x \in E \, \forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n > n_\varepsilon : ||x_n - x|| \leq \varepsilon$

Definition 1.1.2. *Ein vollständiger normierter Raum heißt* Banachraum*.*

Beispiel.

1. $[\mathbb{R}, |\cdot|]$ ist der normierte Raum über den reellen Zahlen mit der Betragsfunktion.
2. $[\mathbb{R}^n, ||\cdot||_p], x = (\xi_1, \ldots, \xi_n), ||x||_p = \begin{cases} \left(\sum_{i=1}^n |\xi_i|^p\right)^{1/p} & 1 \leq p < \infty \\ \max_{1 \leq i \leq n} |\xi_i| & p = \infty \end{cases}$
   - $||\cdot||_1 = \sum_{i=1}^n |\xi_i|$ heißt Summennorm.
   - $||\cdot||_2 = \sqrt{\sum_{i=1}^n |\xi_i|^2}$ heißt euklidische Norm.
   - $||\cdot||_\infty$ heißt Maximumnorm.
3. (a) Normierter Raum der beschränkten Funktionen auf einer Menge Ω :
   $[B(\Omega), ||\cdot||_\infty]$, $B(\Omega) := \{f : \Omega \to \mathbb{R} | f \text{ beschränkt}\}, ||f||_\infty := \sup_{x \in \Omega} |f(x)|$
   Bemerkung: $B(\{1, \ldots, n\}) = \mathbb{R}^n, f = (f(1), \ldots, f(n)), B(\mathbb{N})$: Raum der beschränkten Folgen

   (b) Sei $[X, d]$ ein metrischer Raum. Normierter Raum der beschränkt stetigen Funktionen auf $X : [C_b(X), ||\cdot||_\infty], C_b(X) := \{f : X \to \mathbb{R} | f \text{ beschränkt stetig}\}$
4. Ist X ein kompakter metrischer Raum, so gilt: $C_b(X) = C(X), [C_b(X), ||\cdot||_\infty] = [C(X), ||\cdot||_\infty]$

[1] $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$

5. Normierter Raum der $\mathcal{R}$-integrierbaren Funktionen: $[\mathcal{R}[a,b], ||\cdot||_\infty], \mathcal{R}[a,b] = \{f : [a,b] \to \mathbb{R} | f$ Riemann-integrierbar$\}, ||f||_\infty := \sup_{x\in I} |f(x)|$

6. $[C^1[a,b], ||\cdot||_\infty], C^1[a,b] := \{f \in C[a,b] | f$ stetig differenzierbar auf $]a,b[\}$

Satz 1.1.3. *$[B(\Omega), ||\cdot||_\infty], [C_b(X), ||\cdot||_\infty], [\mathcal{R}[a,b], ||\cdot||_\infty]$ sind Banachräume. Daraus folgt, dass auch $[\mathbb{R}^n, ||\cdot||_\infty]$ ein Banachraum ist, denn $B(\{1,\ldots,n\}) = \mathbb{R}^n$.*

Proof. : Vollständigkeit von $[C_b(X), ||\cdot||_\infty]$:
Sei $(f_n) \subset C_b(X)$ eine Cauchyfolge. Dann ist zu zeigen, dass ein $f \in C_b(X)$ mit $\lim_{n\to\infty} ||f_n - f||_\infty = 0$ existiert.

1. Existenz des Limes
(f_n) ist eine Cauchyfolge, d.h. $\forall \varepsilon > 0 \, \exists n_\varepsilon \in \mathbb{N} \, \forall n, m \geq n_\varepsilon : ||f_n - f_m||_\infty \leq \varepsilon$
$||f_n - f_m|| = \sup_{x\in X} |f_n(x) - f_m(x)| \leq \varepsilon \Leftrightarrow \forall x \in X : |f_n(x) - f_m(x)| \leq \varepsilon$ Bei festem $x \in X$ gilt: $(f_n(x))$ ist eine Cauchyfolge in $\mathbb{R}$. Da $\mathbb{R}$ vollständig ist, folgt daher: $\exists \lim_{n\to\infty} f_n(x) =: f(x)$

2. Beschränkheit von f:
$\forall x \in X \, \forall m \geq n \geq n_\varepsilon : |f_n(x) - f_m(x)| \leq \varepsilon \xrightarrow{m\to\infty} \forall x \in X \, \forall n \geq n_\varepsilon : |f_n(x) - f(x)| \leq \varepsilon \Rightarrow ||f_n - f|| \leq \varepsilon$ für $n \geq n_\varepsilon$ Insbesondere gilt: $||f||_\infty = ||f_{n_\varepsilon} + f - f_{n_\varepsilon}||_\infty \leq ||f_{n_\varepsilon}||_\infty + ||f - f_{n_\varepsilon}||_\infty \leq ||f_{n_\varepsilon}||_\infty + \varepsilon < \infty$ d.h. f ist beschränkt.

3. Stetigkeit von f:
Sei $x_0 \in X$ beliebig und f_{n_ε} stetig: $\exists \delta_\varepsilon > 0 \, \forall x \in X : d(x, x_0) \leq \delta_\varepsilon \Rightarrow |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x_0)| \leq \varepsilon \Rightarrow |f(x) - f(x_0)| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(x_0)| + |f_{n_\varepsilon}(x_0) - f(x_0)| \leq \varepsilon + \varepsilon + \varepsilon = 3\varepsilon$ D.h. f ist in x_0 stetig und es folgt, dass auch f auf X stetig ist.

Die Punkte 2 und 3 implizieren, dass $f \in C_b(X)$ und $\lim_{n\to\infty} ||f_n - f|| = 0 \Rightarrow [C_b(X), ||\cdot||_\infty]$ ist ein vollständiger normierter Raum. □
Die Vollständigkeit von $[B(\Omega), ||\cdot||_\infty]$ kann analog zu 1. und 2. bewiesen werden.
Für die Vollständigkeit von $[\mathcal{R}[a,b], ||\cdot||_\infty]$ genügt es, zu zeigen, dass $[\mathcal{R}[a,b], ||\cdot||_\infty]$ in $B[a,b]$ abgeschlossen ist:
Sei $(f_n) \subset \mathcal{R}[a,b] : f_n \to f \in B[a,b] (||f_n - f||_\infty \to 0)$. zu zeigen: $f \in \mathcal{R}[a,b]$:
Sei $\varepsilon > 0$. Für $\eta := \frac{\varepsilon}{4(b-a)}$ wählen wir $n_0 \in \mathbb{N} : ||f_n - f||_\infty = \sup_{x\in[a,b]} |f_{n_0}(x) - f(x)| \leq \eta \Leftrightarrow \forall x \in [a,b] : -\eta - f_{n_0}(x) \leq f(x) \leq \eta + f_{n_0}(x) \Rightarrow \overline{S}(f;\mathfrak{Z}) \leq \overline{S}(f_{n_0} + \eta;\mathfrak{Z}) \leq \overline{S}(f_{n_0};\mathfrak{Z}) + \overline{S}(\eta;\mathfrak{Z}) = \overline{S}(f_{n_0};\mathfrak{Z}) + \eta(b-a) = \overline{S}(f_{n_0};\mathfrak{Z}) + \frac{\varepsilon}{4}$
$\underline{S}(f;\mathfrak{Z}) \geq \underline{S}(f_{n_0} - \eta;\mathfrak{Z}) \geq \underline{S}(f_{n_0};\mathfrak{Z}) + \underline{S}(-\eta;\mathfrak{Z}) = \underline{S}(f_{n_0};\mathfrak{Z}) - \eta(b-a) = \underline{S}(f_{n_0};\mathfrak{Z}) - \frac{\varepsilon}{4}$
Also:

$$\begin{aligned} \overline{S}(f;\mathfrak{Z}) - \underline{S}(f;\mathfrak{Z}) \leq & \qquad \bar{(}f_{n_0};\mathfrak{Z}) + \frac{\varepsilon}{4} - (\underline{S}(f_{n_0};\mathfrak{Z}) - \frac{\varepsilon}{4}) \\ = & \qquad \overline{S}(f_{n_0};\mathfrak{Z}) - \underline{S}(f_{n_0};\mathfrak{Z}) + \frac{\varepsilon}{2} \end{aligned}$$

$f_{n_0} \in \mathcal{R}[a,b] \Rightarrow \exists \mathfrak{Z}_\varepsilon : \overline{S}(f_{n_0};\mathfrak{Z}) - \underline{S}(f_{n_0};\mathfrak{Z} \leq \frac{\varepsilon}{2} \Rightarrow \overline{S}(f;\mathfrak{Z}_\varepsilon) - \underline{S}f;\mathfrak{Z}_\varepsilon) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, d.h. $f \in \mathcal{R}[a,b]$ ist in $B[a,b]$ abgeschlossen. Daraus folgt, dass $\mathcal{R}[a,b]$ vollständig ist. □

Bemerkung. Konvergenz bezüglich $||\cdot||_\infty$ heißt *gleichmässige Konvergenz* von Funktionen in $B(\Omega), C_b(X)$ oder $\mathcal{R}[a,b]$.
$||f_n - f||_\infty \to 0 \Leftrightarrow \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon : ||f_n - f||_\infty = \sup_x |f_n(x) - f(x)| \leq \varepsilon \Leftrightarrow$

$$\forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon\, \forall x : |f_n(x) - f(x)| \leq \varepsilon$$

Im Gegensatz dazu ist die punktweise Konvergenz einer Folge wie folgt definiert:

$$\forall x\, \forall \varepsilon > 0\, \exists n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon : |f_n(x) - f(x)| \leq \varepsilon$$

Aus gleichmässiger Konvergenz folgt punktweise Konvergenz.

Gleichmässige Konvergenz und Integration

Satz 1.1.4. *$(f_n) \subset \mathcal{R}[a,b]$ und $f_n \xrightarrow{||\cdot||_\infty} f$. Dann gilt:*

$$f \in \mathcal{R}[a,b] \text{ und } \lim_{n\to\infty} \int_a^b f_n dx = \int_a^b \lim f_n dx = \int_a^b f dx$$

Proof. $(f_n) \in \mathcal{R}[a,b], f_n \xrightarrow{||\cdot||_\infty} f \Rightarrow f \in \mathcal{R}[a,b]$
Sei $\varepsilon > 0$. Dann existiert $n_\varepsilon \in \mathbb{N}\, \forall n \geq n_\varepsilon : ||f_n - f||_\infty \leq \varepsilon$

$$\begin{aligned}
\left|\int_a^b f_n dx - \int_a^b f dx\right| &= \left|\int_a^b (f_n - f) dx\right| \leq \int_a^b |f_n - f| dx \\
&= \int_a^b |f_n(x) - f(x)| dx \leq \int_a^b ||f_n - f||_\infty dx \\
&\leq \int_a^b \varepsilon dx = \varepsilon(b-a) \quad \text{für } n \geq n_\varepsilon
\end{aligned}$$

Dies heißt, dass $\lim \int_a^b f_n(x) dx = \int_a^b \lim f_n(x) dx = \int_a^b f dx$. □

Bemerkung. (i) Der Satz 1.1.4 gilt nicht für punktweise Konvergenz.

Beispiel. $f_n : [0,1] \to \mathbb{R}, f_n(x) := \max\{n - n^2|x - \frac{1}{n}|; 0\} = \begin{cases} n^2 x & 0 \leq x \leq \frac{1}{n} \\ 2n - n^2 x & \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0 & \frac{2}{n} \leq x \leq 1 \end{cases}$

$f_n(x) \to 0$ punktweise
f_n ist nicht punktweise konvergent gegen 0. Denn $||f_n - f||_\infty = ||f_n - 0||_\infty = ||f_n||_\infty = n \to \infty$
$\int_a^b f_n(x) dx = 1\, \forall n \geq 2, \lim \int_0^1 f_n(x) dx = 1$, Aber $\int_0^1 \lim_{n\to\infty} f_n(x) dx = \int_0^1 0 dx = 0$

(ii) Satz 1.1.4 gilt nicht für uneigentliche Integrale (unendliche Intervalle).
Beispiel: $f_n : [0, \infty[\to \mathbb{R}, f_n(x) := \frac{x}{n^2} e^{-\frac{x}{n}}, ||f_n||_\infty = \frac{1}{en}, ||f_n||_\infty \to 0$.
Aber $\int_0^\infty f_n(x) dx = \frac{1}{n} + 1 \to 1$ und $\int_0^\infty \lim f_n(x) dx = \int_0^\infty 0 dx = 0$

Gleichmässige Konvergenz und Differentiation

Satz 1.1.5. *Seien $f_n : [a,b] \to \mathbb{R}$ stetig differenzierbar und $f_n \to f : [a,b] \to \mathbb{R}$ punktweise konvergent. Die Folge der Ableitungen $f_n'[a,b] \to \mathbb{R}$ konvergiere gleichmässig. Dann ist f stetig differenzierbar und es gilt*

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

Proof. $f_n' \xrightarrow{||\cdot||} g \in C[a,b]$
zu zeigen: $g = f'$
$\forall x \in [a,b]\, f_n(x) = f_n(a) + \int_a^x f_n'(t)dt$
f_n' ist stetig. Obiges kann man wegen des Hauptsatzes der Differential- und Integralrechnung schliessen. Aus dem Satz 1.1.4 folgt $\lim_{n\to\infty} \int_a^x f_n'(t)dt = \int_a^x \lim_{n\to\infty} f_n'(t)dt = \int_a^x g(t)dt$. Also: $f(x) = f(a) + \int_a^x f_n'(t)dt \Rightarrow f'(x) = g(x) \Rightarrow f' = g$ □

Bemerkung.
- Selbst wenn gilt $f_n \xrightarrow{||\cdot||} f$, impliziert f differenzierbar i.a. nicht, dass $\lim f_n'(x) = f'(x)$.
- $f_n : [-N,N] \to \mathbb{R} \quad f_n(x) = \frac{1}{n}\sin(nx)$ beliebig oft differenzierbar
  $||f_n||_\infty = \frac{1}{n} \to 0 \Rightarrow \lim f_n = 0$
  $f_n'(x) = \cos(nx) \not\to f'(x) = 0$

1.2 Endlichdimensionale normierte Räume

Definition 1.2.1. *Ein Vektorraum E (über $\mathbb{K}$) heißt genau dann* endlichdimensional*, wenn*

$$\exists m \in \mathbb{N}\, \exists x_1, \ldots, x_m \in E\, \forall x \in E\, \exists \alpha_1, \ldots, \alpha_m \in \mathbb{K} : x = \sum_{i=1}^m \alpha_i x_i$$

$$\dim E := \inf\{m | \exists x_1, \ldots, x_m \in E\, \forall x \in A\, \exists \alpha_1, \ldots, \alpha_m \in \mathbb{K} : x = \sum_{i=1}^m \alpha_i x_i\}$$

Beispiel: $\mathbb{R}^n, \mathbb{C}^n$

Satz 1.2.2 (Satz von Bolzano/Weierstrass)**.** *Sei $K \subset \mathbb{R}^n$. Dann gilt: K ist genau dann kompakt in $[\mathbb{R}^n, ||\cdot||_\infty]$, wenn K abgeschlossen und beschränkt in $[\mathbb{R}^n, ||\cdot||_\infty]$ ist.*

Proof. Der Beweis wird auf den gleichlautenden Satz in $[\mathbb{R}, |\cdot|]$ zurückgeführt:

1. Aus der Tatsache, dass K kompakt ist, folgt, dass es auch abgeschlossen und beschränkt ist.

2. Sei K abgeschlossen und beschränkt, d.h.

   a) $\exists M \geq 0\, \forall x \in \mathbb{K} : ||x||_\infty \leq M$

   b) $\forall (x_n) \subset K : x_n \xrightarrow{||\cdot||} x \in \mathbb{R}^n \Rightarrow x \in K$

Betrachten den Fall $n = 2$: Sei $(x_n) \subset K, x_n = (\xi_1^{(n)}, \xi_2^{(n)}), ||x_n||_\infty = \max\{|\xi_1^{(n)}|, |\xi_2^{(n)}|\} \leq M(\xi_1^{(n)}, \xi_2^{(n)} < \mathbb{R}$ beschränkt. Nach dem Satz von Bolzano/Weierstrass aus der Analysis I folgt, dass es eine konvergente Teilfolge $(\xi_1^{(n_k)}) \subset (\xi_1^{(n)}) : \xi_1^{(n_k)} \to \xi \in \mathbb{R}$ gibt. Die entsprechende Teilfolge $(\xi_2^{(n_k)}) \subset \xi_2^{(n)}$ ist ebenfalls beschränkt und enthält eine konvergente Teilfolge $(\xi_2^{(n_k)_l}) : (\xi_2^{(n_k)_l}) \to \xi_2 \in \mathbb{R}$. Damit hat die Folge $x_{(n_k)_l} = (\xi_1^{(n_k)_l}, \xi_2^{(n_k)_l}) \to x = (\xi_1, \xi_2)$. Denn $||x_{(n_k)_l} - x||_\infty = \max\{|\xi_1^{(n_k)_l} - \xi_1|, |\xi_2^{(n_k)_l} - \xi_2|\} \leq |\xi_1^{(n_k)_l} - \xi_1| + |\xi_2^{(n_k)_l} - \xi_2| \to 0$, d.h. $x_{(n_k)_l} \to x$. Somit enthält $(x_n) \subset K$ eine konvergente Teilfolge mit Grenzwert x. Da K abgeschlossen ist, folgt, dass x Element von K ist und dass K kompakt ist. Der Beweis für $n > 2$ ist analog. □

Definition 1.2.3. *Sei E ein Vektorraum und $||\cdot||_0, ||\cdot||$ zwei Normen auf E. $||\cdot||_0$ und $||\cdot||$ heissen genau dann auf E äquivalent, wenn*

$$\exists c_0, c > 0 \,\forall x \in E : ||x||_0 \leq c_0||x|| \qquad ||x|| \leq c||x||_0$$

Satz 1.2.4. *Auf einem endlichdimensionalen Vektorraum E über $\mathbb{K}$[2] sind sämtliche Normen äquivalent.*

Proof. Sei $\dim E = n$ und $\{x_1, \dots, x_n\}$ eine Basis von E. Für die $x \in E$ existieren $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ mit $x = \sum_{i=1}^n \alpha_i x_i$. Sei $||\cdot||$ eine beliebige Norm auf E. Des weiteren definieren wir eine Norm $||\cdot||_0$ auf E durch $||x||_0 := \sum_{i=1}^n |\alpha_i| = ||\alpha||_1$ mit $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$. Es genügt nun, zu zeigen, dass $||\cdot||_0$ und $||\cdot||$ äquivalent sind.

$$||x|| = \left|\left|\sum_{i=1}^n \alpha_i \cdot x_i\right|\right| \leq \sum_{i=1}^n |\alpha_i| \cdot ||x_i|| \leq \max_{1<i\leq n} ||x_i|| \sum_{i=1}^n |\alpha_i| = \left(\underbrace{\max_{1<i\leq n} ||x_i||}_{=c}\right) ||x||_0$$

Jetzt ist „$||\cdot||_0 \leq c_0||\cdot||$“. Dazu betrachten wir folgende Funktion: $f : [\mathbb{R}^n, ||\cdot||_\infty] \to \mathbb{R}, f(\alpha_1, \dots, \alpha_n) := ||\sum_{i=1}^n \alpha_i x_i||$. Es gilt, dass f stetig ist.
Denn $|f(\alpha_1, \dots, \alpha_n) - f(\beta_1, \dots, \beta_n)| = ||\sum_1^n (\alpha_i - \beta_i) x_i|| \leq \sum_1^n |\alpha_i - \beta_i| \cdot ||x_i|| \leq \max_{1\leq i\leq n} ||x_i|| \cdot \sum_1^n |\alpha_i - \beta_i| \leq n \cdot \underbrace{\max_{1\leq i\leq n} |\alpha_i - \beta_i|}_{=||(\alpha 1, \dots, \alpha_n) - (\beta_1 - \beta_n)||_\infty}$. Daraus folgt, dass f stetig ist.
$S = \{\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n | \sum_{i=1}^n |\alpha_i| = 1\} \subset \mathbb{R}^n$ ist abgeschlossen und beschränkt in $[\mathbb{R}^n, ||\cdot||_\infty]$. Denn für $g(\alpha_1, \dots, \alpha_n) := \sum_{i=1}^n |\alpha_i|, g : [\mathbb{R}^n, ||\cdot||_\infty] \to \mathbb{R}$ gilt, dass g stetig ist. Damit ist $S = g^{-1}(\{1\})$ (Urbild) abgeschlossen. S ist beschränkt: $||\alpha||_\infty = ||(\alpha_1, \dots, \alpha_n)||_\infty \leq \sum_{i=1}^n |\alpha_i| = 1$ für $\alpha = (\alpha_1, \dots, \alpha_n) \in S$. Somit ist S in $[\mathbb{R}^n, ||\cdot||_\infty]$ kompakt nach Satz 1.2.2. Weil f stetig ist, folgt, dass f auf S das Minimum annimmt, d.h.
$\exists (\alpha_1^0, \dots, \alpha_n^0) \in S \,\forall (\alpha_1, \dots, \alpha_n) \in S : 0 < \frac{1}{c_0} = f((\alpha_1^0, \dots, \alpha_n^0) \leq f(\alpha_1, \dots, \alpha_n) = ||\sum_1^n \alpha_i x_i||$. Sei $0 \neq x \in E$ beliebig. Dann existiert $0 \neq (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : x =$

[2] $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$

$\sum_{i=1}^n \alpha_i x_i$. Für $\left(\frac{\alpha_1}{\sum_{i=1}^n |\alpha_i|}, \frac{\alpha_2}{\sum_{i=1}^n |\alpha_i|}, \dots, \frac{\alpha_n}{\sum_{i=1}^n |\alpha_i|}\right) \in S$ gilt $\frac{1}{c_0} \leq ||\sum_{k=1}^n \frac{\alpha_k}{\sum_{i=1}|\alpha_i|} x_k|| = \frac{1}{\sum_{i=1}^n |\alpha_i|} ||\sum_1^n \alpha_n x_k|| \Rightarrow ||x_0|| = \sum_{i=1}^n |\alpha_i|| \leq c_0 ||\sum_1^n \alpha_k x_k|| = c_0||x||$. Für $x = 0$ ist die Ungleichung ebenfalls erfüllt. □

Satz 1.2.5. *(i) Ein endlichdimensionaler normierter Raum $[E, ||\cdot||]$ über $\mathbb{R}$ (bzw. $\mathbb{C}$) ist ein Banachraum.*

(ii) (Bolzano/Weierstrass-Eigenschaft) Sei K eine Menge eines endlichdimensionalen normierten Raumes E. Dann gilt: K ist genau dann kompakt, wenn K beschränkt und abgeschlossen ist.

Proof. $x = \sum_{i=1}^n \alpha_i y_i \quad (\alpha_i \in \mathbb{R})$
$||x_0|| := \max_{1 \leq i \leq n} |\alpha_i| = ||\alpha||_\infty, \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ ist eine Norm auf E. Nach Satz 1.2.4 existieren $c, C > 0 : c||x||_0 \leq ||x|| \leq C||x||_0$.

(i) Sei $x_k = \sum_{i=1}^n \alpha_i^{(k)} y_i, \alpha_i^{(k)} = (\alpha_1^{(k)}, \dots, \alpha_n^{(k)}) \in \mathbb{R}^n$. Es gilt $c||\alpha^{(k)} - \alpha^{(l)}||_\infty \leq ||x_k - x_l|| \leq C||\alpha^{(k)} - \alpha^{(l)}||_\infty$. Sei $(x_k) \subset E$ eine Cauchyfolge. Dann folgt, dass $(\alpha^{(k)}) \subset [\mathbb{R}^n, ||\cdot||_\infty]$ eine Cauchyfolge ist. Da nun $[\mathbb{R}^n, ||\cdot||8]$ vollständig ist, gibt es ein $\alpha =)\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : ||\alpha^{(k)} - \alpha||_\infty \to 0 \Rightarrow ||x_k - x|| \to 0, x = \sum_{i=1}^n \alpha_i y_i \in E$. Also konvergiert (x_k) in E gegen x und es folgt, dass E vollständig ist.

(ii) „$\Rightarrow$“ klar, siehe Analysis I

„$\Leftarrow$“ Definieren $f : [\mathbb{R}^n, ||\cdot||_\infty] \to E, f(\alpha_1, \dots, \alpha_n) := \sum_{i=1}^n \alpha_i y_i, \{y_i\}$ ist eine Basis in E. Es gilt

a) f ist bijektiv von $\mathbb{R}^n$ auf E.

b) f, f^{-1} sind stetige Funktionen nach den Ungleichungen, die zu Beginn des Beweises genannt sind.

Sei $K \subset E$ beschränkt und abgeschlossen. $f^{-1}(K) = \{(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n : \sum_{i=1}^n \alpha_i y_i \in K\}$ ist beschränkt, da K beschränkt (wegen Ungleichung oben) und abgeschlossen (Urbilder von abg. Mengen sind wieder abg.), da f stetig ist. Nach Satz 1.2.2 folgt: $f^{-1}(K) \subset [\mathbb{R}^n, ||\cdot||_\infty]$ kompakt. Da f stetig ist, folgt $K = f(f^{-1}(K)) \subset E$ kompakt.

□

Bemerkung: Es gilt sogar: Sind in einem Banachraum E die beschränkten und abgeschlossenen Mengen kompakt, so ist E endlichdimensional.

1.3 Stetige lineare Abbildungen (Operatoren) zwischen normierten Räumen

Definition 1.3.1. *Seien E, F Vektorräume über $\mathbb{K}$. Eine Abbildung $A : E \to F$ heißt* linear*, genau dann wenn:*

$$\forall \alpha, \beta \in \mathbb{K} \, \forall x, y \in E : A(\alpha x + \beta y) = \alpha Ax + \beta Ay$$

Bemerkung: Lineare Abbildungen zwischen endlichdimensionalen Vektorräumen können durch Matrizen dargestellt werden.

Definition 1.3.2. *Seien E, F normierte Räume. Eine lineare Abbildung $A : E \to F$ heißt beschränkt, genau dann wenn:*

$$\exists M \geq 0 \, \forall x \in E : ||Ax||_F \leq M||x||_E$$

Es gilt: $\mathcal{L}(E,F) := \{A : E \to F | A \text{ beschränkt}\}$ ist ein Vektorraum. Durch $||A|| := \sup_{||x||_E \leq 1} ||Ax||_F$ wird auf $\mathcal{L}(E,F)$ eine Norm (Operatorennorm) erklärt. $[\mathcal{L}(E,F), ||\cdot||]$ ist ein normierter Raum.

Ferner gilt: $\forall x \in E : ||Ax||_F \leq ||A|| \cdot ||x||_E$. Für $x \neq 0$: $\left.\begin{array}{rcl} ||A(\frac{x}{||x||})|| & \leq & ||A|| \\ ||Ax\frac{1}{||x||}|| & = & \frac{1}{||x||}||Ax|| \end{array}\right\} \Rightarrow$

$||Ax|| \leq ||A|| \cdot ||x||$

Komposition von Operatoren: $A \in \mathcal{L}(E,F), B \in \mathcal{L}(F,G) \Rightarrow B \circ A \in \mathcal{L}(E,G)$ und es gilt $||BA|| \leq ||B|| \cdot ||A||$.

Satz 1.3.3. *Seien E, F normierte Räume und* $\dim E < \infty$*. Dann ist jede lineare Abbildung $A : E \to F$ stetig (=beschränkt).*

Proof. Sei $n := \dim E$. Wählen Basis $x_1, \ldots, x_n \in E$. Dann $\forall x \in E \, \exists \alpha_1, \ldots, \alpha_n \in \mathbb{K}$. $x = \sum_{i=1}^n \alpha_i x_i, ||x||_0 = \sum_1^n |\alpha_i|$ ist eine äquivalente Norm. Somit

$$\begin{aligned} ||Ax_i|| &= \left|\left| A\left(\sum_1^n \alpha_i x_i\right)\right|\right|_F \\ &= \left|\left|\sum_1^n \alpha_i A x_i\right|\right|_F \\ &\leq \sum_1^n |\alpha_i| \cdot ||Ax_i|| \\ &\leq \max_{1 \leq i \leq n} ||Ax_i|| \sum_1^n |\alpha_i| \\ &\leq (c \max_{1 \leq i \leq n} ||Ax_i||)||x|| \end{aligned}$$

$\Rightarrow A$ beschränkt, d.h. A ist stetig. □

Satz 1.3.4. *Sei E ein normierter Raum und F ein Banachraum. Dann ist $\mathcal{L}(E,F)$ ein Banachraum.*

2 Differentiation

2.1 Differentiation - Beispiele und Rechenregeln

Definition 2.1.1. *E und F seien Banachräume. Eine Funktion $f : D \to F, D \subset E$ offen, heißt in $x \in D$ differenzierbar, g.d.w.:*

$$\exists A \in \mathcal{L}(E,F)\, \exists r(h)\, \exists \delta > 0\, \forall ||h|| \leq \delta : x + h \in D$$

$$\textit{und } f(x+h) = f(x) + Ah + r(h) \textit{ mit } \lim_{h \to 0} \frac{r(h)}{||h||} = 0$$

Bemerkung: Die Differentiation von f im Punkt $x \in D$ bedeutet eine Approximation durch eine affive lineare Abbildung $f(x) + Ah$ mit der Eigenschaft, dass

$$\lim_{h \to 0} \frac{f(x+h) - f(x) - Ah}{||h||} = 0.$$

Setzt man $\tilde{r}(h) := \begin{cases} \frac{r(h)}{||h||} & 0 < ||h|| \leq \delta \\ 0 & h = 0 \end{cases}$, ist $\tilde{r}$ in $h = 0$ stetig. Äquivalent zur Definition der Differentiation von f im Punkt x ist die folgende:

$$f(x+h) = f(x) + Ah + ||h||\tilde{r}(h) \text{ mit } \lim_{h \to 0} \tilde{r}(h) = 0$$

Satz 2.1.2. *Es gilt: $A \in \mathcal{L}(E,F)$ ist eindeutig bestimmt.*

Definition 2.1.3. *Seien E, F Banachräume und $f : D \to F, D \subset E$ offen, differenzierbar in $x \in D$. Dann heißt der eindeutig bestimmte Operator $A \in \mathcal{L}(E,F)$ mit $f(x+h) = f(x) + Ah + r(h), \lim \frac{r(h)}{||h||} = 0$* Ableitung *(oder* Fréchet-Ableitung*) und Ah* Differential *der Funktion f im Punkt $x \in D$*

Beispiel. E, F sind Banachräume und $f : E \to F, f(x) = c + Ax, c \in F, A \in \mathcal{L}(E,F)$
Dann $f'(x) = A$
Beweis: $f(x+h) = c + A(x+h) = c + Ax + Ah = f(x) + Ah, r(h) = 0 \Rightarrow f'(x) = A$

Bemerkung. Seien E, F Banachräume und $f : D \subset E \to F$ in D differenzierbar. Dann ist f' eine Abbildung von D in $\mathcal{L}(E,F)$, $f' : D \to \mathcal{L}(E,F)$, d.h. die Bilder sind beschränkte lineare Operatoren.

Für $E = F = \mathbb{R}$ gilt $\mathcal{L}(\mathbb{R},\mathbb{R}) \approx \mathbb{R}$

Rechenregeln

(i) Seien E, F Banachräume und $f, g : D \to F, D \subset E$ offen und differenzierbar in $x \in D$. Dann $f+g, \alpha \cdot f, \alpha \in \mathbb{R}$ mind. in x differenzierbar und es gilt: $(f+g)'(x) = f'(x) + g'(x), (\alpha f)'(x) = \alpha f'(x)$

(ii) Sei E ein Banachraum, $f, g : D \to \mathbb{R}, D \subset E$ offen, in $x \in D$ differenzierbar. Dann $f \cdot g, \frac{f}{g}(g(x) \neq 0)$ sind in x differenzierbar und es gilt $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x), \left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$

(iii) Kettenregel: Seien E, F, G Banachräume, $f : D \to F$, in x differenzierbar, $g : \tilde{D} \to G, D \tilde{\subset} F$ offen und $f(D) \subset \tilde{D}$ in $f(x)$ differenzierbar. Dann ist $g \circ f$ in x differenzierbar und es gilt $(g \circ f)'(x) = g'(f(x))f'(x), E \xrightarrow{f} F \xrightarrow{g} G$

zu (iii). f, g sind differenzierbar.

$$\begin{aligned}
f(x+h) &= f(x) + f'(x)h + ||h||\tilde{r}_f(h) && \tilde{r}_f \xrightarrow{h \to 0} 0 \\
g(f(x)+k) &= g(f(x)) + g'(f(x))k + ||k||\tilde{r}_g(k) && \tilde{r}_g(k) \to 0 \\
(g \circ f)(x+h) &= g(f(x+h)) = g(f(x) + \underbrace{f'(x)h + ||h||\tilde{r}_f(h)}_{k}) \\
&= g(f(x)) + g'(f(x))(f'(x)h + ||h||\tilde{r}_f(h) + ||f'(x)||h \\
&\quad + ||h||\tilde{r}_g(f'(x)h) + ||h||\tilde{r}_f(h) \\
&= (g \circ f)(x) + g'(f(x))f'(x)h + r_{g \circ f}(h)) \\
&\text{wobei} \\
r_{g \circ f}(h) &= ||h||g'(f(x))\tilde{r}_f(h) \\
&\quad + ||f'(x+h) + ||h||\tilde{r}_f(h)|| \cdot \tilde{r}_g(f'(x)h + ||h||\tilde{r}_f(h))
\end{aligned}$$

Es gilt zu zeigen:

$$\lim_{h \to 0} \frac{r_{g \circ f}(h)}{||h||} \to 0$$

$$\begin{aligned}
||r_{g \circ f}(h)|| \leq &||h|| \cdot ||g'(f(x))|| \cdot ||\tilde{r}_f(h)|| \\
&+ (||f'(x)||||h|| + ||h||||\tilde{r}_f(h)||)||\tilde{r}_g(f'(x)h + ||h||\tilde{r}_g(h))|| \\
\frac{||r_{g \circ f}(h)||}{||h||} \leq &||g'(f(x))|| \cdot ||\tilde{r}_f(h)|| \\
&+ (||f'(x)|| + ||\tilde{r}_f(h)||)||\tilde{r}_g(f'(x)h + ||h||\tilde{r}_g(h)) \xrightarrow{h \to 0} 0 \\
\Rightarrow &\lim \frac{r_{g \circ f}(h)}{||h||} \to 0 \\
\Rightarrow &(g \circ f)'(x) = g'(f(x))f'(x)
\end{aligned}$$

□

Richtungsableitung

Definition 2.1.4. *Seien E, F Banachräume und $0 \neq h \in E$. Existiert für eine Funktion $f : D \to E, D \subset E$ offen, im Punkt $x \in E$ der Grenzwert*

$$\frac{\partial f}{\partial h}(x) := \lim_{\substack{t \to 0 \\ t \in \mathbb{R}}} \frac{f(x+th) - f(x)}{t}$$

, so heißt $\frac{\partial f}{\partial h}(x)$ Richtungsableitung *von f in x.*

Satz 2.1.5. *Seien E, F Banachräume und $f : D \to F, D \subset F$ offen, in $x \in D$ differenzierbar, so ist f in x stetig.*

Beweis: $f(x+h) = f(x) + f'(x)h + ||h||\tilde{r}_f(h) \xrightarrow{h \to 0} f(x)$. *Daher folgt, dass f in x stetig.* □

Satz 2.1.6. *Seien E, F Banachräume, $f : D \to F, D \subset E$ offen, in x differenzierbar. Dann existiert die Richtungsableitung $\frac{\partial f}{\partial h}(x)$ für jede Richtung h, und es gilt $\frac{\partial f}{\partial h}(x) = f'(x)h$.*

Beweis: $f(x+th) = f(x) + f'(x)(th) + r(th)$, *wobei* $\frac{r(th)}{||th||} \xrightarrow{t \to 0} 0$

$\frac{f(x+th)-f(x)}{t} = f'(x)h + \frac{r(th)}{t} \to f'(x)h$ □

Bemerkung. Die Umkehrung von Satz 2.1.6 gilt i.A. nicht.

Beispiel. $f : \mathbb{R}^2 \to \mathbb{R}, f(x_1, x_2) := \begin{cases} \frac{x_1 x_2^2}{x_1^2 + x_2^4} & (x_1, x_2) \neq (0,0) \\ 0 & (x_1, x_2) = (0,0) \end{cases}$ Für f gilt:

(1) f ist in (0,0) *nicht* stetig $\Rightarrow f$ in (0,0) nicht differenzierbar.
   Beweis: $(x_n) = (\frac{1}{n^2}, \frac{1}{n}) \to (0,0), f(\frac{1}{n^2}, \frac{1}{n}) = \frac{\frac{1}{n^4}}{\frac{1}{n^4} + \frac{1}{n^4}} = \frac{1}{2} \not\to 0 \Rightarrow f$ in (0,0) nicht stetig.

(2) f besitzt in (0,0) Richtungsableitungen in jede Richtung $h \neq 0$
   Beweis: $\frac{f(0+th)-f(0)}{t} = \frac{1}{t} \frac{th_1 t^2 h_2^2}{t^2 h_1^2 + t^4 h_2^4} = \frac{h_1 h_2^2}{h_1^2 + t^2 h_2^4} \xrightarrow{t \to 0} \frac{h_1 h_2^2}{h_1^2} = \frac{h_2^2}{h_1}, h = (h_1, h_2)$ für $h_1 \neq 0 \to 0$ für $h_1 = 0, h_2 \neq 0$
   $\frac{\partial f}{\partial h}(0) = \frac{h_2^2}{h_1}$ für $h_1 \neq 0$ und $\frac{\partial f}{\partial h}(0) = 0$ für $h_1 = 0$. Ferner zeigt dieser Ausdruck, dass $\frac{\partial f}{\partial h}(0)$ nicht linear in h ist.

□

2.2 Differentiation von $\mathbb{R}^n$-$\mathbb{R}^m$-Funktionen

2.2.1 Partielle Ableitung von $\mathbb{R}^n$-$\mathbb{R}$-Funktionen

Definition 2.2.1. *Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine reelle Funktion. Existiert in $x = (x_1, \ldots, x_n) \in \mathbb{R}^n \in D$ die Richtungsableitung $\frac{\partial f}{\partial e_i}(x) = \lim_{t \to 0} \frac{f(x+te_i)-f(x)}{t}, e_i = (0, \ldots, 1, \ldots, 0) \in \mathbb{R}^n$, so heißt $\frac{\partial t}{\partial x_i}(x) := \frac{\partial f}{\partial e_i}(x)$* partielle Ableitung *von f nach der i-ten Koordinate.*

Bemerkung: Die partielle Ableitung einer Funktion $f: D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, kann man als gewöhnliche Ableitung von Funktionen einer reellen Veränderlichen interpretieren, indem man nach der i-ten Koordinate differenziert und die übrigen festhält.

Ist $f(x) = f(x_1, \dots, x_n)$, so ist $\frac{\partial f}{\partial x_i} = \frac{f(x_1,\dots,x_i+t,x_1,\dots,x_n)-f(x_1,\dots,x_n)}{t}$. Also sind die partiellen Ableitungen Richtungsableitungen in die i-te Richtung.

Beispiel:

$$f(x_1, x_2) = x_1^2 + x_1 \sin x_2 \quad f: \mathbb{R}^2 \to \mathbb{R}$$

$$\frac{\partial f}{\partial x_1} = 2x + \sin x_2 \quad \frac{\partial f}{\partial x_2} = x_1 \cos x_2$$

Die Existenz der partiellen Ableitung von $f: D \to \mathbb{R}, D \subset \mathbb{R}^n$ reicht i.a. nicht für die Differenzierbarkeit von f aus!

Beispiel:

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2^2}{x_2^2 + x_2^4} & (x_1, x_2) \neq 0 \\ 0 & (x_1, x_2) = 0 \end{cases}$$

$$\frac{\partial f}{\partial h}(0,0) = \frac{h_2^2}{h_1} \text{ für } h_1 \neq 0 \wedge = 0 \text{ für } h_1 = 0$$

$$\frac{\partial f}{\partial x_1}(0,0) = 0 \quad \frac{\partial f}{\partial x_2}(0,0) = 0$$

Aber f ist in $(0,0)$ nicht differenzierbar!

2.2.2 Darstellung der Ableitung von $\mathbb{R}^n$-$\mathbb{R}^m$-Funktionen

Satz 2.2.2. *Sei $D \subset \mathbb{R}$ offen und $f: D \to \mathbb{R}^m, f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ und $f_i: D \to \mathbb{R}$ im Punkt $x \in \mathbb{R}^n$ differenzierbar. Dann gilt für die Ableitung von f die folgende Darstellung:*

$$f'(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Beweis: *$f'(x) = A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ Dann gilt: $f(x+h) = f(x) + Ah + r(h)$ mit $\frac{r(h)}{||h||} \xrightarrow{h \to 0} 0$.*

Somit $\begin{pmatrix} f_1(x+h) \\ \vdots \\ f_m(x+h) \end{pmatrix} = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} + \begin{pmatrix} \sum_{k=1}^n a_{1k} h_k \\ \vdots \\ \sum_{k=1}^n a_{mk} h_k \end{pmatrix} + \begin{pmatrix} r_1(h) \\ \vdots \\ r_m(h) \end{pmatrix}$ mit $\frac{r_i(h)}{||h||} \to 0, (i = 1, \dots, m)$

$\Rightarrow f_i(x+h) = f_i(x) + \sum_{k=1}^n a_{ik} h_k + r_i(h)$ für $i = 1, \dots, m$

$h = te_j$ mit $e_j = (0, \dots, 1, \dots, 0)$: $f_i(x + te_j) = f_i(x) + \sum_{k=1}^n a_{ik} + \delta_{jk} + r_i(te_j) = f_i(x) + ta_{ij} + r_i(te_j) \Rightarrow \frac{f_i(x+te_j)-f_i(x)}{t} = a_{ij} + \frac{r_i(te_j)}{t}$

$\Rightarrow \frac{\partial f}{\partial x_j} = \lim_{t \to 0} \frac{f(x+te_j)-f(x)}{t} = a_{ij}$ $\square$

Satz 2.2.3. *Sei $D \subset \mathbb{R}^n$ offen und $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \to \mathbb{R}^m, f_i : D \to \mathbb{R}$ eine Funktion, so dass die partielle Ableitung $\frac{\partial f_i}{\partial x_j}$ in D existieren und im Punkt $x \in D$ stetig sind, dann ist f in x differenzierbar und es gilt die Formel aus Satz 2.2.2.*

Beweis: *$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \to \mathbb{R}^m, D \subset \mathbb{R}^n$ offen, ist in $x \in D$ differenzierbar, g.d.w. $f : D \to \mathbb{R}$ in $x \in D$ differenzierbar ist.*
Daher ist o.B.d.A. für $m = 1$:
$f : D \to \mathbb{R}$ Aus der Tatsache, dass D offen ist, folgt, dass $\exists \delta > 0 : \overset{\circ}{B}(x, \delta) = \{z \in \mathbb{R}^n : ||x - z||_\infty < \delta\} \subset D$
Sei $h \in \mathbb{R}^n, ||h|| < \delta$. Setzen

$$z^{(i)} := x + \sum_{k=1}^{i} h_k e_k \quad z^{(0)} := x$$

Es gilt $z^{(n)} = x + h$. $z^{(i)}$ und $z^{(i-1)}$ unterscheiden sich in der i-ten Koordinate. Nach dem Mittwelwertsatz der Differentialrechnung für differenzierbare Funktionen in einer Veränderlichen existiert ein $\theta_i \in [0, 1] : f(z^{(i)}) - f(z^{(i-1)}) = \frac{\partial f}{\partial x_i}(y^{(i)})h_i$, wobei $y^{(i)} = z^{(i-1)} + \vartheta_i h_i e_i \quad \varphi(t) := f(z^{(i-1)} + the_j)\, \varphi(1) = f(z^{(i)})\, \varphi(0) = f(z^{(i-1)})$
Nach dem Mittelwertsatz existiert ein t_0 mit

$\frac{\varphi(1)-\varphi(0)}{1-0} = \varphi'(t_0) = \lim_{\zeta \to 0} \frac{\varphi(t_0+\zeta)+\varphi(t_0)}{\zeta} = \frac{f(z^{(i-1)}+(t_0+\zeta)h_i e_i) - \overbrace{f(z^{(i-1)} + t_0 h_i e_i}^{=:y^{(i)}}}{\zeta}$
$= \lim_{\zeta \to 0} \frac{f(y^{(i)}+\zeta h_i e_i)-f(y^{(i)})}{\zeta h_i} h_i = \frac{\partial f}{\partial x_i}(y^{(i)})h_i$ *Somit ist* $f(x+h) - f(x) = \sum_{i=1}^n (f(z^{(i)}) - f(z^{(i-1)}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(y^{(i)})h_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x)h_i + \underbrace{\sum_{i=1}^n (\frac{\partial f}{\partial x_i}(y^{(i)}) - \frac{\partial f}{\partial x_i}(x))h_i}_{=:r(h)}$

Nach der Cauchy-Schwartzschen Ungleichung gilt:

$$|r(h)| \leq \sqrt{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y^{(i)}) \right|^2} \cdot ||h||_2$$

Daher folgt nun: $\frac{|r(h)|}{||h||_2} \leq \sqrt{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i}(x) - \frac{\partial f}{\partial x_i}(y^{(i)}) \right|^2} \cdot ||h||_2 \to 0$ *für $h \to 0$, da $\frac{\partial f}{\partial x_i}$ stetig in x, geht $y^{(i)}$ gegen x.* □

Die Existenz der partiellen Ableitungen von $f : D \subset \mathbb{R}^n \to \mathbb{R}$ reicht i.a. *nicht* für die Differenzierbarkeit von f!

Definition 2.2.4. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, eine reellwertige Funktion, so dass die* *partiellen Ableitungen* $\frac{\partial f}{\partial x_i}$ *existieren in* $x \in D$. *Dann heißt:*

$$gradf(x) = \left(\frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x)\right)$$

Gradient *von* f *an der Stelle* x.

Bemerkung:

(i) f muss nicht notwendigerweise in x differenzierbar sein!

(ii) Ist f in x differenzierbar, so gilt: $f'(x) = \mathrm{grad} f(x)$. Für die Richtungsableitung gilt dann $\frac{\partial f}{\partial h}(x) = f'(x)h = (\mathrm{grad} f(x))h$.

Satz 2.2.5. *Sei* $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ *offen, in* $x \in D$ *differenzierbar. Ist* $gradf(x) = 0$*, so verschwinden alle Richtungsableitungen in* x. *Ist* $gradf(x) \neq 0$*, so gibt es unter allen Richtungsableitungen* $\frac{\partial f}{\partial h}(x)$ *mit* $||h||_2 = 1$ *eine grösste, nämlich die des Gradienten* $gradf(x)$ *mit* $\frac{\partial f}{\partial h}(x) = ||gradf(x)||_2$.
Beweis: $\frac{\partial f}{\partial h}(x) = (gradf(x))h$

$$\begin{aligned}
\left|\frac{\partial f}{\partial h}(x)\right| &= \left|\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x) h_i\right| \\
&\leq \sqrt{\sum_{i=1}^{n} \left|\frac{\partial f}{\partial x_i}(x)\right|^2} \sqrt{\sum_{i=1}^{n} |h_i|^2} \\
&= ||gradf(x)||_2 ||h||_2 \\
&= ||gradf(x)||_2 \text{ für } ||h||_2 = 1
\end{aligned}$$

Für $h_0 := \frac{gradf(x)}{||gradf(x)||_2}$ *gilt:*
$\frac{\partial f}{\partial h_0} = (gradf(x))h_0 = (gradf(x))\left(\frac{gradf(x)}{||gradf(x)||_2}\right) = \frac{||gradf(x)||_2^2}{||gradf(x)||_2} = ||gradf(x)||_2$ □

Sei $f : D \subset \mathbb{R}^n \to \mathbb{R}$ differenzierbar. $f(a+h) = f'(a)h + r(h)$, $\frac{r(h)}{||h||} \to 0$

(i) Für „kleine“ h, d.h. $||h||$ „klein“, ist eine Näherung für $f(a+h)$ die Formel

$$f(a+h) \approx f(a) + f'(a)h = f(a) + (\mathrm{grad} f(a))h$$

Ausführlicher: $f(a_1+h_1, \dots, a_n+h_n) \approx f(a_1, \dots, a_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a_i)h_i$. Damit wird der Unterschied $f(a+h) - f(a)$ für „kleine“ h durch den Ausdruck $\frac{\partial f}{\partial x_1}(a)h_1 + \dots + \frac{\partial f}{\partial x_n}(a)h_n = f'(a)h$ annähernd gegeben. Auf dieser Formel beruht die Fehlerrechnung.

(ii) Für kleine $t \in \mathbb{R}$ und $h \in \mathbb{R}$ mit $||h||_2 = 1$ wird der Unterschied $f(a+th) - f(a) \approx (\mathrm{grad} f(x))th$ am grössten für $h = \frac{\mathrm{grad} f(a)}{||\mathrm{grad} f(a)||_2}$, d.h. die Richtung des Gradienten.

Beispiel zur Fehlerrechnung:
Eine Pyramide mit quadratischer Grundfläche der Kantenlänge 1 m und Höhe 3 m hat ein Volumen $V(l,h) = \frac{1}{3}l^2h = 1\,\mathrm{m}^3$. Gesucht ist eine Näherung für das Volumen einer Pyramide mit einer Länge von 1,03 m und Höhe 3,06 m.
$V : \mathbb{R}^2 \to \mathbb{R}, V(l+\Delta l, h+\Delta h) \approx V(l,h) + V'(l,h)\begin{pmatrix}\Delta l\\ \Delta h\end{pmatrix} = V(l,h) + \frac{\partial V}{\partial l}(l,h)\Delta l + \frac{\partial V}{\partial h}(l,h)\Delta h = V(l,h) + \frac{2}{3}lh\Delta l + \frac{1}{3}l^2\Delta h$
$l = 1\,\mathrm{m}, h = 3\,\mathrm{m}, \Delta l = 0,03\,\mathrm{m}, \Delta h = 0,06\,\mathrm{m}$
$V(l+\Delta l, h+\Delta h) \approx \frac{2}{3}\cdot 1\cdot 3\cdot 0,03 + \frac{1}{3}1^2\cdot 0,06 = 1 + 0,06 + 0,02 = 1,08$
Der exakte Wert beträgt: $V(1,03; 3,06) = 1,082118$

Ein wichtiger Spezialfall der Kettenregel

$$u : D \subset \mathbb{R}^n \to \mathbb{R}^m \qquad g : M \subset \mathbb{R}^m \to \mathbb{R}\; u(D) \subset M$$

$$u = \begin{pmatrix}u_1\\ \vdots\\ u_m\end{pmatrix} \qquad f = g \circ u$$

$$f(x_1,\dots,x_n) = g(u_1(x_1,\dots,x_n),\dots,u_m(x_1,\dots,x_n))\frac{\partial f}{\partial x_k}(x) = \sum_{i=1}^m \frac{\partial g}{\partial u_i}(u(x))\frac{\partial u_i}{\partial x_k}(x)$$

Kettenregel: $f'(x) = g'(u(x))u'(x)$

$$\begin{aligned}
\left(\frac{\partial f}{\partial x_1}(x),\dots,\frac{\partial f}{\partial x_n}(x)\right) &= \left(\frac{\partial g}{\partial u_1}(u(x)),\dots,\frac{\partial g}{\partial u_m}(u(x))\right)\begin{pmatrix}\frac{\partial u_1}{\partial x_1}(x) & \dots & \frac{\partial u_1}{\partial x_n}(x)\\ \vdots & & \vdots\\ \frac{\partial u_m}{\partial x_1}(x) & \dots & \frac{\partial u_m}{\partial x_n}(x)\end{pmatrix}\\
&= \left(\frac{\partial g}{\partial u_1}(u(x))\frac{\partial u_1}{\partial x_1}(x) + \dots + \frac{\partial g}{\partial u_m}(u(x))\frac{\partial u_m}{\partial x_1}(x),\dots,\right.\\
&\qquad \left.\frac{\partial g}{\partial u_1}(u(x))\frac{\partial u_1}{\partial x_n}(x) + \dots + \frac{\partial g}{\partial u_m}(u(x))\frac{\partial u_m}{\partial x_n}(x)\right)\\
&\Rightarrow \quad \frac{\partial f}{\partial x_k}(x) = \frac{\partial g}{\partial u_1}(u(x))\frac{\partial u_1}{\partial x_k}(x) + \dots + \frac{\partial g}{\partial u_m}(u(x))\frac{\partial u_m}{\partial x_k}(x)
\end{aligned}$$

2.3 Mittelwert, Taylorscher Satz, lokale Extrema

2.3.1 Mittelwertsätze

Satz 2.3.1. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, auf D differenzierbar, $x, x+h \in D$ mit $x + th \in D$ für $0 \le t \le 1$. Dann*

$$\exists 0 < \vartheta < 1 : f(x+h) - f(x) = f'(x+\vartheta h)h$$

***Beweis**: (Anwendung des eindimensionalen Mittelwertsatzes) $\varphi(t) := f(x+th), \varphi : [0,1] \to \mathbb{R}$ ist stetig differenzierbar.*

Kettenregel: $\varphi'(t) = f'(x+th)h$. *Nach dem Mittelwertsatz gilt:* $\frac{\varphi(1)-\varphi(0)}{1-0} = \varphi(1) - \varphi(0) = \varphi'(\vartheta)$
$f(x+h) - f(x) = f'(x+\vartheta h)h$ □

Vorbemerkungen zum Mittelwertsatz für vektorwertige Funktionen

(i) In der Version von Satz 2.3.1 gilt der Mittelwertsatz für vektorwertige Funktionen *nicht* mehr.
Beispiel: $f : [0, 2\pi] \to \mathbb{R}^2, f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ Zu zeigen ist: $\nexists 0 < \vartheta < 1 : f(2\pi) - f(0) = f'(\vartheta 2\pi)2\pi$. Denn $f(2\pi) - f(0) = 0, f'(t) = \begin{pmatrix} (\cos t)' \\ (\sin t)' \end{pmatrix} = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}, ||f'(t)||_2 = 1 \Rightarrow f'(t) \neq 0 \, \forall t$

(ii) Ausweg: Mittelwertsatz für vektorwertige Funktionen in „Integralversionen" formulieren.

Definition 2.3.2. *Sei* $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : [a, b] \to \mathbb{R}^m$, *wobei* $f_i : [a, b] \to \mathbb{R}$ *riemannintegrierbar sind. Dann heißt*

$$\int_a^b f(t)dt := \begin{pmatrix} \int_a^b f_1(t)dt \\ \vdots \\ \int_a^b f_m(t)dt \end{pmatrix}$$

das $\mathcal{R}$*-Integral von* f *über* $[a, b]$.

f ist z.B. integrierbar, wenn f stetig ist. Ferner: $\int_a^b f(t)dt = 0, \int_a^b f(t)dt = -\int_b^a f(t)dt$. Analog definiert man für $A; [a, b] \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ das $\mathcal{R}$-Integral als:

$$\int_a^b A(t)dt := \left(\int_a^b a_{ij}(t)dt\right)_{\substack{i=1,\dots,m \\ j=1,\dots,n}}, A(t) = (a_{ij}(t))$$

Ferner: $\int_a^b A(t)hdt = \left(\int_a^b A(t)dt\right) h, h \in \mathbb{R}^n, h = \text{konstant}$

Hilfssatz: Für jede stetige Funktion $f : [a, b] \to \mathbb{R}^n$ gilt:

$$\left|\left|\int_a^b f(t)dt\right|\right|_2 \leq \int_a^b ||f(t)||_2 dt$$

Diese Ungleichung gilt auch für jede beliebige Norm $|| \cdot ||$ auf $\mathbb{R}^n$
Beweis: Sei $u := \int_a^b f(t)dt \in \mathbb{R}^n$. Dann $||u||_2^2 =< u, u >=< \int_a^b f(t)dt, u >= \int_a^b < f(t), u > dt \leq \int_a^b ||f(t)||_2 ||u||_2 dt = \left(\int_a^b ||f(t)_2 dt\right) ||u||_2 \Rightarrow \left|\left|\int_a^b f(t)dt\right|\right|_2 = ||u||_2 \leq \int_a^b ||f(t)||_2 dt$ □

Satz 2.3.3. *Sei $f : D \to \mathbb{R}^m, D \subset \mathbb{R}^n$ offen, stetig differenzierbar und $x, x+h \in D$ mit $x + th \in D, 0 \leq t \leq 1$. Dann*

$$f(x+h) - f(x) = \left(\int_0^1 f'(x+th)dt\right) h$$

Ferner gilt:

$$||f(x+h) - f(x)||_2 \leq \sup_{0 \leq t \leq 1} \underbrace{||f'(x+th)||}_{Operatorennorm} \cdot ||h||_2$$

Die Operatorennorm ist genauer $||f'(x+th) : [\mathbb{R}^n, ||\cdot||_2] \to [\mathbb{R}^m, ||\cdot||_2]||$

Beweis: *$f : \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}, f_i : D \to \mathbb{R}$ stetig differenzierbar. Wir definieren $\varphi_i : [0,1] \to \mathbb{R}, i = 1, \ldots, m, \varphi_i(t) := f_i(x+th), \varphi_i$ stetig differenzierbar. Man wendet den Hauptsatz der Differential- und Integralrechnung an: $\varphi_i(1) - \varphi_i(0) = \int_0^1 \varphi_i'(t)dt, \varphi_i'(t) = f_i'(x+th)h$. Also hat man : $f_i(x+h) - f_i(x) = \int_0^1 f_i'(x+th)hdt = \left(\int_0^1 f_i'(x+th)dt\right) h$*

$f(x+h) - f(x) = \int_0^1 \begin{pmatrix} f_1'(x+th) \\ \vdots \\ f_m'(x+th) \end{pmatrix} dth = \left(\int_0^1 f(x+th)dt\right) h$

Ferner: $||f(x+h) - f(x)||_2 = \left|\left|\int_0^1 f(x+th)dth\right|\right| \leq \left|\left|\int_0^1 f'(x+th)dt\right|\right| \cdot ||h||_2$
$\leq \int_0^1 \sup_{0 \leq t \leq 1} ||f'(x+th)||dt||h||_2 = (\sup_{0 \leq t \leq 1} ||f'(x+th)||)||h||_2$ $\square$

Definition 2.3.4. *(i) Sei E ein normierter Raum und $x_0, x_1, \ldots, x_n \in E$. Dann heißt*

$$P = \bigcup_{i=1}^{n} \{(1-\Theta)x_{i-1} + \Theta x_i : 0 \leq \Theta \leq 1\}$$

Polygonzug *durch die Punkte $x_0, \ldots, x_n$.*

(ii) $D \subset E$ heißt genau dann polygonzugzusammenhängend, *wenn $\forall x, y \in D$ existiert ein Polygonzug $P \subset D$ durch $x_0, \ldots, x_n$ mit $x_0 = x, x_n = y$.*

Satz 2.3.5. *Sei $f : D \to \mathbb{R}^m, D \subset \mathbb{R}^n$ offen und polygonzugzusammenhängend, differenzierbar mit $f' = 0$ auf D. Dann ist f konstant auf D.*

Beweis: *Es gilt: f ist stetig differenzierbar auf D, da $f' = 0$. Sei $x_0 \in D$ fest und $x \in D$ beliebig. Dann existiert ein in D verlaufender Polygonzug durch $x_0, x_1, \ldots, x_n := x$. Aus Satz 2.3.3 folgt> $f(x_0) = f(x_1) = \cdots = f(x_n) = f(x)$. Also $f(x) = f(x_0) \forall x \in D$. Daher folgt, dass f konstant ist.* $\square$

D heißt *Gebiet*, wenn D offen und polygonzugzusammenhängend.

2.3.2 Taylorsche Formel

Höhere partielle Ableitungen von $\mathbb{R}^n$-$\mathbb{R}$-Funktionen Notation: Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, eine partiell differenzierbare Funktion. Sind alle partiellen Ableitungen $\frac{\partial f}{\partial x_i} : D \to \mathbb{R}$ selbst wieder partiell differenzierbar, so heißt f zweimal differenzierbar. Man kann die partiellen Ableitungen 2. Ordnung bilden:

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

Allgemeiner definiert man durch Induktion folgendes:
$f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, heißt $(k{+}1)$-mal partiell differenzierbar, wenn sie k-mal partiell differenzierbar ist und alle partiellen Ableitungen der k-ter Ordnung $\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_2} \partial x_{i_1}} D \to \mathbb{R}$ partiell differenzierbar sind.
Eine Funktion $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, heißt k mal stetig differenzierbar, wenn sie k mal partiell differenzierbar ist und alle partiellen Ableitungen der Ordnung $\leq k$ stetig sind.

Satz 2.3.6 (Satz von Schwartz über die Vertauschbarkeit der partiellen Ableitungen)**.** *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, zweimal stetig partiell differenzierbar. Dann gilt für alle*

$$\xi \in D \text{ und } i, j = 1, \dots, n : \frac{\partial^2 f}{\partial x_j \partial x_i}(\xi) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi)$$

***Beweis**: O.B.d.A.: $n = 2, i = 1, j = 2, x_1 \to x, x_2 \to y$. Wir zeigen $\frac{\partial^2 f}{\partial x \partial y}(\xi, \eta) = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta)$ Wählen auf $\mathbb{R}^2$ die Maximumnorm. Dann existiert die ε-Kugel $B_\varepsilon(\xi, \eta) := \{(\tilde{\xi}, \tilde{\eta}) \in \mathbb{R}^2 : \max\{|\tilde{\xi}-\xi|, |\tilde{\eta}-\eta|\} < \varepsilon\} \subset D$. Damit: $(\xi{+}h, \eta{+}k) \in B_\varepsilon(\xi, \eta)$ für $|h|, |k| < \varepsilon$. Sei $0 < |h|, |k| < \varepsilon$. Wir definieren $\varphi(x) := f(x, \eta{+}k) - f(x, \eta), \varphi : [\xi, \xi{+}h] \to \mathbb{R}, h > 0$ differenzierbar. Nach dem Mittelwertsatz: $\exists x_1 \in]\xi, \xi + h[: \varphi(\xi + h) - \varphi(\xi) = h\varphi'(x_1)$.*

$$\begin{aligned}
F(h,k) := & \quad f(\xi + h, \eta + k) - f(\xi + h, \eta) - f(\xi, \eta + k) + f(\xi, \eta) \\
\Rightarrow & \quad F(h,k) = \varphi(\xi + h) - \varphi(\xi) = h\varphi'(x_1) \\
= & \quad h\left[\frac{\partial f}{\partial x}(x_1, \eta + k) - \frac{\partial f}{\partial x}(x_1, \eta)\right] \\
\Rightarrow & \quad \exists y_1 \in]\eta, \eta + k[: \frac{\partial f(x_1, \eta + k)}{\partial x} - \frac{\partial f}{\partial x}(x_1, \eta) \\
= & \quad k\frac{\partial^2 f}{\partial y \partial x}(x_1, y_1)
\end{aligned}$$

*Also $F(h,k) = hk\frac{\partial^2 f}{\partial y \partial x}(x_1, y_1), x_1 \in]\xi, \xi + h[, y_1 \in]\eta, \eta + k[$. Nun Darstellung von $F(h,k)$ mit der Funktion $\psi(y) := f(\xi + h, y) - f(\xi)$ in der Form $F(h,k) = \psi(\eta + k) - \psi(\eta)$
Analog erhält man dann wie oben $\exists x_2 \in]\xi, \xi{+}h[\quad y_2 \in]\eta, \eta{+}k[: F(h,k) = hk\frac{\partial^2 f}{\partial x \partial y}(x_2, y_2)$
$\overset{h,k \neq 0}{\Rightarrow} \frac{\partial^2 f}{\partial y \partial x}(x_1, y_1) = \frac{\partial^2 f}{\partial x \partial y}(x_2, y_2) \overset{h,k \to 0}{\Rightarrow} \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta) = \frac{\partial^2 f}{\partial x \partial y}(\xi, \eta)$* □

Bemerkung: Der Beweis hat gezeigt, dass es genügt, die Existenz der partiellen Ableitungen 2. Ordnung in einer Umgebung von (ξ, η) und die Stetigkeit in (ξ, η) vorauszusetzen.

Folgerung 2.3.7. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, k mal stetig differenzierbar. Dann gilt:*

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(\xi) = \frac{\partial^k f}{\partial x_{i_{\pi(k)}} \dots x_{i_{\pi(1)}}}$$

für jede Permutation $\pi : \{1, \dots, k\} \to \{1, \dots, k\}$
Der Beweis erfolgt induktiv.

Taylorscher Satz Notationen: $D \subset \mathbb{R}^n$ offen,
$C^m(D) := \{f : D \to \mathbb{R} | f$ besitzt stetige partielle Ableitungen der Ordnung $\leq m\}$
$m = 0 : C(D) := C^0(D)$

Satz 2.3.8. *Sei $D \subset \mathbb{R}^n$ offen, $f \in C^{m+1}(D), (m \geq 0)$ und $h := \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$. Liegen die Punkte ξ und $\xi + h$ mitsamt ihrer Verbindungsstrecke $\xi + th, 0 \leq t \leq 1$, in D, so gibt es ein Θ mit $0 < \Theta < 1$, so dass*

$$\begin{aligned} f(\xi + h) &= \sum_{k=0}^{m} \left(\frac{1}{k!} \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \dots + h_n \frac{\partial}{\partial x_n} \right)^k f(\xi) \right) \\ &\quad + \frac{1}{(m+1)!} \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^{m+1} f(\xi + \Theta h) \end{aligned}$$

wobei

$$\begin{aligned} (h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n})^k f(\xi) := & \sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}(\xi) h_{i_1} \dots h_{i_k} \\ = & \sum_{\alpha_1 + \dots + \alpha_k} \frac{k!}{\alpha_1! \dots \alpha_n!} \cdot \frac{\partial^k}{\partial x_{i_n}^{\alpha_n} \dots \partial x_{i_1}^{\alpha_1}}(\xi) h_1^{\alpha_1} \dots h_n^{\alpha_n} \end{aligned}$$

***Beweis**: Der Beweis erfolgt für den Spezialfall $n = 2$:*

$$f(\xi + h) = \sum_{k=0}^{m} \frac{1}{k!} \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} \right)^k f(\xi) + \frac{1}{(m+1)!} \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} \right)^{m+1} f(\xi + \Theta h)$$

wobei

$$\begin{aligned} \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} \right)^k f(\xi) := & \sum_{\alpha_1 + \alpha_2 = k} \frac{k!}{\alpha_1! \alpha_2!} \frac{\partial^k}{\partial x_2^{\alpha_2} \partial x_1^{\alpha_1}} h_1^{\alpha_1} h_2^{\alpha_2} \\ = & \sum_{j=0}^{k} \binom{k}{j} \frac{\partial^k f}{\partial x_2^{k-j} \partial x_1^j} h_1^j h_2^{k-j} \end{aligned}$$

$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}, h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \varphi(t) := f(\xi + th) = f(\xi_1 + th_1, \xi_2 + th_2), \varphi : [0,1] \to \mathbb{R}$ *ist* $(m+1)$-*mal stetig differenzierbar. Dann gilt* $\varphi^{(k)}(t) = \sum_{j=0}^{n} \binom{k}{j} \frac{\partial^k f}{\partial x_2^{k-j} \partial x_1^j} h_1^j h_2^{k-j} (\xi_1 + th_1, \xi_2 + th_2)$.
Durch Induktion erhält man folgendes:
Induktionsanfang:

$$\begin{aligned} \varphi'(t) = & \quad f'(\xi + th)h \\ = & \quad \left(\frac{\partial f}{\partial x_1}(\xi + th) \frac{\partial f}{\partial x_2}(\xi + th) \right) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \\ = & \quad h_1 \frac{\partial f}{\partial x_1}(\xi + th) h_2 \frac{\partial f}{\partial x_2}(\xi + th) \end{aligned}$$

für $h = 1$
$k \to k+1:$

$$\begin{aligned} \varphi^{(k+1)}(t) = & \quad (\varphi^{(k)}(t))' \\ = & \quad \left(\sum_{j=0}^{k} \binom{k}{j} h_i^j h_2^{k-j} \frac{\partial^k f}{\partial x_1^j \partial x_2^{k-j}} (\xi + th) \right)' \\ = & \quad \sum_{j=0}^{k} \binom{k}{j} h_1^{j+1} h_2^{k-j} \frac{\partial f^{k+1}}{\partial x_1^{j+1} \partial x_2^{k-j}} + \sum_{j=0}^{k} \binom{k}{j} h_1^j h_2^{k-j+1} \frac{\partial^{k+1} f}{\partial x_1^j \partial x_2^{k-j+1}} (\xi + th) \\ = & \quad \sum_{j=1}^{k+1} \binom{k}{j-1} h_1^j h_2^{k-j+1} \frac{\partial^{k+1} f}{\partial x_1^j \partial x_2^{k-j+1}} (\xi + th) \\ & \quad + \sum_{j=0}^{k} \binom{k}{j} h_1^j h_2^{k-j+1} \frac{\partial^{k+1} f}{\partial x_1^j \partial x_2^{k-j+1}} (\xi + th) \\ = & \quad \sum_{j=0}^{k+1} \binom{k+1}{j} h_1^j h_2^{k-j+1} \frac{\partial^{k+1} f}{\partial x_1^j \partial x_2^{k+1-j}} (\xi + th) \end{aligned}$$

Der taylorsche Satz für Funktionen einer Variable liefert die Behauptung:
$\varphi(t) = \sum_{k=0}^{m} \frac{\varphi^{(k)}(0)}{k!} t^k + \frac{\varphi^{(m+1)}(\Theta t)}{(m+1)!} t^{k+1}, 0 < \Theta < 1$
Für $t = 1 : f(\xi + h) = \varphi(1) = sum_{k=0}^{m} \frac{\varphi^{(k)}(0)}{k!} + \frac{\varphi^{(m+1)}(\Theta)}{(m+1)!} = \sum_{k=0}^{m} \frac{1}{k!} h_1 \frac{\partial}{\partial x_1 + h_2 \frac{\partial}{\partial x_2}})^k f(\xi) + \frac{1}{(m+1)!} (h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2})^{m+1} f(\xi + \Theta h)$ □

Für $m = 1$ gilt folgender Spezialfall:
Korollar
Sei $D \subset \mathbb{R}^n$ offen, $f \in C^2(D)$ und $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \subset \mathbb{R}^n, \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in D$. Liegen die Punkte ξ und $\xi + h$ mitsamt ihrer Verbindungsstrecke $\xi + th, 0 \le t \le 1$ in D, so existiert Θ mit

$0 < \Theta < 1$ und es gilt :

$$\begin{aligned} f(\xi+h) = & \quad f(\xi) + \left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right) f(xi) + \\ & \frac{1}{2}\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right) f(xi)^2 f(\xi+\Theta h) \\ = & \quad f(\xi) + h_1\frac{\partial h}{\partial x_1} + \cdots + h_n\frac{\partial f}{\partial x_n}(\xi) + \frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi+\Theta h)h_j h_k \\ = & \quad f(\xi) + f'(\xi)h + \frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi+\Theta h)h_j h_k \end{aligned}$$

Satz 2.3.9. *Sei $D \subset \mathbb{R}^n$ offen und $f \in C^2(D)$. Für $B_\varepsilon(\xi) \subset D$ und alle $h \in \mathbb{R}^n$ mit $\xi + h \in D_\varepsilon(\xi)$ gilt*

$$f(\xi+h) = f(\xi) + f'(\xi)h + \frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi)h_j h_k + ||h||_2^2\rho(h)$$

mit $\lim_{h\to 0}\rho(h) = 0$.

Beweis: *Nach dem Korollar oben* $f(\xi+h) = f(\xi)+f'(\xi)h+\frac{1}{2}\sum_{j,k=1}^{n}\frac{\partial^2 f}{\partial x_j \partial x_k}(\xi)h_j h_k + r(h)$ *mit* $r(h) = \frac{1}{2}\sum_{j,k=1}^{n}\left[\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi+\Theta h) - \frac{\partial^2 f}{\partial x_k \partial x_j}(\xi)\right] h_j h_k$. *Somit*

$$|r(h)| \leq \frac{1}{2}\left(\sum_{j,k=1}^{n}\sqrt{\left|\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi+\Theta h) - \frac{\partial^2 f}{\partial x_k \partial x_j}\right|^2}\right)\underbrace{\sqrt{\sum_{j,k=1}^{n}|h_j|^2|h_k|^2}}_{\sum_{j,k=1}^{n}|h_j|^2 = ||h||_2^2}$$

$$\begin{aligned} \sum_{j,k=1}^{n}|h_j|^2|h_k|^2 = & \quad \sum_{k=1}^{n}\sum_{j=1}^{n}|h_j|^2|h_k|^2 \\ = & \quad \sum_{k=1}^{n}(|h_k|^2\sum_{j=1}^{n}|h_j|^2) \\ = & \quad \left(\sum_{j=1}^{n}|h_j|^2)(\sum_{k=1}^{n}|h_k|^2\right) \\ = & \quad \left(\sum|h_j|^2\right) \end{aligned}$$

Also: $\frac{|r(h)|}{||h||_2^2} \leq \frac{1}{2}\sqrt{\sum_{j,k=1}^{n}\left|\frac{\partial^2 f}{\partial x_k \partial x_j}(\xi+\Theta h) - \frac{\partial^2 f}{\partial x_k \partial x_j}(\xi)\right|} \to 0$

Also $\rho(h) := \begin{cases} \frac{r(h)}{||h||_2^2} & h \neq 0 \\ 0 & h = 0 \end{cases} \to 0$ □

Lokale Extrema

Definition 2.3.10. *Eine Funktion $f : D \subset \mathbb{R}^n \to \mathbb{R}$ besitzt in $x \in D$ ein* lokales Maximum *bzw.* lokales Minimum, *falls eine Kugel $B_\varepsilon(x) \subset D, \varepsilon > 0$ existiert, so dass $f(x) \geq f(y)$ bzw. $f(x) \leq f(y)$ für alle $y \in B_\varepsilon(x)$.*

Tritt in dieser Definition der Fall $f(x) = f(y)$ nur für $x = y$ ein, so spricht man von einem *isoliertem lokalem Maximum bzw. Minimum.* Ein lokales Extremum ist ein lokales Maximum bzw. Minimum.

Satz 2.3.11. *Sei $D \subset \mathbb{R}^n$ offen und $f : D \to \mathbb{R}$ eine partiell differenzierbare Funktion. Dann gilt f besitzt in $x \in D$ ein lokales Extremum und es folgt, dass der Gradient von f an der Stelle $x = 0$. Dies ist gleichbedeutend mit $\frac{\partial f}{\partial x_1}(x) = \cdots = \frac{\partial f}{\partial x_1}(x) = 0$*

Proof. Wir definieren $\varphi_i(t) := f(x+te_i), i = 1, \ldots, n, e_i = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$. f besitzt in x ein lokales Extremum. Das impliziert φ_i haben in 0 ein lokales Extremum. Somit gilt $\varphi_i'(0) = 0 = \lim_{t\to 0} \frac{f(x+te_i)-f(x)}{t} = \frac{\partial f}{\partial x_i}(x)$. Also ist der Gradient in $f(x) = \left(\frac{\partial f}{\partial x_i}(x), \ldots, \frac{\partial f}{\partial x_n}(x)\right) = 0$ □

Definition 2.3.12. *Sei $A = (a_{ij})_{i,j=1}^n$ eine symmetrische reelle $n \times n$ Matrix A heißt* positiv definit, *genau dann wenn*

$$\forall x \in \mathbb{R}^n \backslash \{0\} : \langle Ax, x\rangle > 0$$

A heißt positiv semidefinit, *genau dann wenn*

$$\forall x \in \mathbb{R}^n : \langle Ax, x\rangle \geq 0$$

A heißt negativ definit, genau dann wenn $-A$ positiv definit.

$$\Leftrightarrow \forall x \in \mathbb{R}^n \backslash \{0\} : \langle Ax, x\rangle < 0$$

A heißt indefinit, *genau dann wenn*

$$\exists x, y \in \mathbb{R}^n : \langle Ax, x\rangle > 0 \text{ und } \langle Ay, y\rangle < 0$$

Kriterien für positiv definite Matrizen Eine symmetrische relle $n \times n$ Matrix $A = (a_{ij})$ ist genau dann positiv, wenn $\exists \alpha > 0 \, \forall x \in \mathbb{R}^n : \langle Ax, x\rangle \geq \alpha ||x||_2^2$ oder alle Eigenwerte von A sind positiv, d.h. $\lambda > 0$ oder $\forall k = 1, \ldots, n : \det \begin{pmatrix} a_{11} & \ldots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \ldots & a_{kk} \end{pmatrix} > 0$

Beweis siehe Vorlesung Algebra I+II

Definition 2.3.13. *Sei $D \subset \mathbb{R}^n$ offen und $f \in C^2(D)$. Die* Hessesche Matrix *von f im Punkt $x \in D$ ist definiert durch:*

$$(Hessf)(x) := \left(\frac{\partial^2 f}{\partial x_j \partial x_i} \right)_{i,j=1}^{n}$$

Wegen $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}$ ist $(Hessf)(x)$ symmetrisch. Nach dem Satz 2.2.3 hat man $f''(x) = (Hessf)(x)$.

Mit $f'(x)h = \langle \text{grad } f(x), h\rangle$ und $\langle f''(x)h, h\rangle = \langle \text{Hess } f(x)h, h\rangle = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i} h_i h_j$ kann Satz 2.3.9 wie folgt beschrieben werden:
Sei $D \subset \mathbb{R}^n$ offen und $f \in C^2(D)$. Dann gilt für $B_\varepsilon(x) \in D$ und alle $h \in \mathbb{R}^n$ mit $x + h \in B_\varepsilon(x) : f(x+h) = f(x) + \langle f'(x), h\rangle + \frac{1}{2}\langle f''(x)h, h\rangle + ||h||^2 \rho(h)$ wobei $\lim_{h\to 0} \rho(h) = 0$.

Satz 2.3.14. *Sei $D \subset \mathbb{R}^n$ offen und $f \in C^2(D)$ mit $(\text{grad } f)(x) = f'(x) = 0$. Dann gilt:*

(i) Ist $f''(x) = (\text{Hess } f)(x)$ positiv definit, so hat f in x ein isoliertes lokales Minimum.
Beweis: (Die Beweise zu (i) und (ii) können analog geführt werden.) $f(x+h) = f(x) + \langle f'(x), h\rangle + \frac{1}{2}\langle f''(x)h, h\rangle + ||h||^2\rho(h) = f(x) + \frac{1}{2}\langle f''(x)h, h\rangle + ||h||^2\rho(h)$
Wählen $\delta > 0 : \forall h : ||h||_2 \leq \delta \Rightarrow |\rho(h)| \leq \frac{\alpha}{4}$. Damit erhält man für $0 < ||h||_2 \leq \delta$: $f(x+h) = f(x) + \frac{1}{2}\langle f''(x)h, h\rangle + ||h||_2^2\rho(h) \geq f(x) + \frac{1}{2}\langle f''(x)h, h\rangle - ||h||_2^2|\rho(h)| \geq f(x) + \frac{\alpha}{2}||h||_2^2 - \frac{\alpha}{4}||h||_2^2 = f(x) + \frac{\alpha}{4}||h||_2^2 > f(x)$

(ii) Ist $f''(x)$ negativ definit, so hat die Funktion f in x ein isoliertes lokales Maximum.
Beweis erfolgt analog zu (i).

(iii) Ist $f''(x)$ indefinit, so besitzt f in x kein lokales Extremum.
Beweis: Zeigen: In jeder ε-Kugel $B_\varepsilon(x)$ um x existieren $y', y'' : f(y'') < f(x) < f(y')$. $f''(x)$ ist indefinit: $\exists \xi \in \mathbb{R}^n \backslash \{0\} : \langle f''(x)\xi, \xi\rangle =: \alpha > 0$. Für kleine $|t|$ gilt: $f(x+t\xi) = f(x) + \frac{1}{2}\langle f''(x)t\xi, t\xi\rangle + ||t\xi||_2^2\rho(t\xi)$ und für hinreichend kleine $|t|$ gilt $|\rho(t\xi)| \leq \frac{\alpha}{4}\frac{1}{||\xi||_2^2}$. Also hat man $f(x+t\xi) = f(x) + \frac{t^2}{2}\langle f''(x)\xi, \xi\rangle + t^2||\xi||_2^2\rho(t\xi) = f(x + \frac{\alpha}{2}t^2 + t^2||\xi||_2^2\rho(t\xi) \geq f(x) + \frac{\alpha}{2}t^2 - t^2||\xi||_2^2|\rho(t\xi)| \geq f(x) + \frac{\alpha}{2}t^2 - \frac{\alpha}{4}t^2 = f(x) + \frac{\alpha}{4}t^2 > f(x)$ für $0 < |t| \leq \delta_0$
Ebenso zeigt man: Ist $\eta \in \mathbb{R}^n$ ein Vektor mit $\langle f''(x)\eta, \eta\rangle < 0$, so gilt für genügend kleine $|t| > 0 : f(x+t\eta) < f(x)$

□

Spezialfall $n = 2 \quad x := x_1, y := x_2$
Sei $D \subset \mathbb{R}^2$ offen, $f \in C^2(D)$ und grad $f(\xi, \eta) = 0$, d.h. $\frac{\partial f}{\partial x}(\xi, \eta) = \frac{\partial f}{\partial y}(\xi, \eta) = 0$

$$f'' = \text{Hess } f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{cases} \text{positiv definit } \Leftrightarrow a_{11} > 0 & \det A = a_{11}a_{22} - a_{12}^2 > 0 \\ \text{negativ definit } \Leftrightarrow a_{11} < 0 & \det A = a_{11}a_{22} - a_{12}^2 > 0 \end{cases}$$

Setzen $\Delta := \det f'' = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial y \partial x}\right)^2$, so hat

(i) f ein isoliertes lokales Minimum in (ξ, η), wenn $\frac{\partial^2 f}{\partial x^2} > 0, \Delta > 0$

(ii) f ein isoliertes lokales Maximum in (ξ, η), wenn $\frac{\partial^2 f}{\partial x^2} < 0, \Delta > 0$

(iii) f kein lokales Extremum, wenn $\Delta < 0$

Beispiele $f : \mathbb{R}^2 \to \mathbb{R}$

1. $f(x,y) = c + x^2 + y^2, \text{grad } f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2x, 2y)$
   $f''(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, f''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ positiv definit. Somit besitzt f in $(0,0)$ ein isoliertes lokales Minimum (sogar globales Minimum).

2. $g(x,y) = c - x^2 - y^2, \text{grad } g = (-2x, -2y), (\text{grad } g)(0,0) = (0,0), g''(x,y) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, -2 < 0, \det g''(0,0) = 4 \Rightarrow g$ besitzt in $(0,0)$ ein isoliertes lokales Maximum.

3. $h(x,y) = c + x^2 - y^2, \text{grad } h = (2x, -2y), \text{grad } (0,0) = (0,0), h''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$
   $\det h''(0,0) = -4 < 0 \Rightarrow h$ hat in $(0,0)$ kein lokales Extremum.

4. Ist die Hessesche Matrix $f''(x,y)$ in einer Nullstelle des Gradienten semidefinit, so lassen sich keine allgemeinen Aussagen machen.

   a) $f_1(x,y) = x^2 + y^4, \text{grad } f_1 = (2x, 4y^3), f_1''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix}$

   b) $f_2(x,y) = x^2, \text{grad } f_2 = (2x, 0), f_2''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$

   c) $f_3(x,y) = x^2 + y^3, \text{grad } f_3 = (2x, 3y^2), f_3''(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 6y \end{pmatrix}$

   $f_i''(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, i = 1, 2, \ldots, h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \langle f_i''(0,0)h, h\rangle = \left\langle \begin{pmatrix} 2h_1 \\ 0 \end{pmatrix}, \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\rangle = 2h_1^2 + 0 = 2h_1^2 \geq 0$. Somit ist $f_i''(0,0)$ semidefinit. Es gilt:

   - f_1 hat in $(0,0)$ ein isoliertes lokales Minimum.
   - f_2 hat in $(0,0)$ lokales Minimum (kein isoliertes)
   - f_3 hat in $(0,0)$ kein lokales Extremum.

Beispiel: Methode der kleinsten Quadrate Sei $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} : \mathbb{R}^n \to \mathbb{R}^m, n \leq m, b = (b_i) = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m, f(x_1, \dots, x_n) = \sum_{k=1}^m \left(\sum_{j=1}^n a_{kj} x_j - b_k \right)^2, f \to$ Minimum, $f(x) = ||Ax - b||_2^2$

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "f"] \arrow[d, "g"] & \mathbb{R} \\ \mathbb{R}^m \end{tikzcd}$$

$g(x) := Ax - b, h(y) := ||y||_2^2, g'(x) = A, h'(y) = \left(\frac{\partial h}{\partial y_1}, \dots, \frac{\partial h}{\partial y_n} \right) = 2(y_1, \dots, y_n) = 2y^T, f'(x) = g'(g(x))g'(x) = 2g(x)^T g'(x) = 2(Ax - b)^T A \ (f'(x)^T)' = (2A^T(Ax - b))' = 2A^T A, (f'(x)^T)' = f''(x)^T = f''(x)$
Notwendige Bedingung für Extremum: $f'(x) = 2(Ax - b)^T A = 0 \Leftrightarrow f'(x)^T = 2A^T(Ax - b) = 0$

$$A^T Ax = A^T b$$

Man setzt voraus, dass A injektiv ist. Dies bedeutet, dass $A^T A$ invertierbar ist. Denn es gilt: $A^T Ax = 0 \Rightarrow 0 = \langle A^T Ax, x \rangle = \langle Ax, Ax \rangle = ||Ax||_2^2 \Rightarrow Ax = 0 \Rightarrow x = 0 \Rightarrow A^T A$ ist invertierbar. Für eine injektive Matrix A gilt, dass f in $x = (A^T A)^{-1} A^T b$ ein isoliertes lokales Minimum besitzt. Denn $\langle f''(x)h, h \rangle = \langle 2A^T Ah, h \rangle = 2\langle Ah, Ah \rangle = 2||Ah||^2 > 0$ für $h \neq 0$

2. Lösungsweg: $f(x) = ||Ax - b|_2^2$ und A injektiv:
$f(x + h) = \langle A(x + h) - b, A(x + h) - b \rangle = \langle Ax - b + Ah, Ax - b + Ah \rangle = \langle Ax - b, Ax - b \rangle + \langle Ax - b, Ah \rangle + \langle Ah, Ax - b \rangle + \langle Ah, Ah \rangle = ||Ax - b||_2^2 + 2\langle Ax - b, Ah \rangle + ||Ah||_2^2 = f(x) + 2\langle A^T(Ax - b), h \rangle + ||Ah||_2^2 = f(x) + ||Ah||_2^2$ für $x = (A^T A)^{-1} A^T b \Rightarrow \forall h \neq 0 : f(x+h) = f(x) + ||Ah||_2^2 > f(x) \Rightarrow f$ besitzt in $x = (A^T A)^{-1} A^T b$ ein absolutes isoliertes Minimum.
Verallgemeinerung: $A : \mathbb{R}^n \to \mathbb{R}^m, n \leq m, b_k \in \mathbb{R}^m, 1 \leq k \leq N$
$f(x) = \sum_{k-1}^N ||Ax - b_k||_2^2 \to$ Minimum.
A ist injektiv $\Rightarrow x = \frac{1}{N} \sum_{k=1}^N (A^T A)^{-1} A^T b_k$ liegt ein absolutes isoliertes Minimum vor.

2.4 Implizite Funktionen

Eine Funktion $f : D \to \mathbb{R}^m, D \subset \mathbb{R}^n$ wird *implizit* genannt, wenn sie durch eine Gleichung der Form

$$F(x, y) = 0$$

gegeben ist, d.h. $F(x, f(x)) = 0 \, \forall x \in D$.

Problem: Unter welchen Bedingungen an F wird durch $F(x, y) = 0$ eine Funktion $y = f(x)$ definiert? Oder unter welchen Bedingungen an F ist die Gleichung $F(x, y) = 0$ eindeutig nach y auflösbar?

In Komponenten zerlegt bedeutet dies, das Gleichungssystem

$$\begin{matrix} F_1(x_1,\ldots,x_n,y_1,\ldots,y_m)=0 \\ \vdots \\ F_m(x_1,\ldots,x_n,y_1,\ldots,y_m)=0, \end{matrix} \quad \text{wobei } x=\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}, y=\begin{pmatrix} y_1\\ \vdots \\ y_m\end{pmatrix}, F=\begin{pmatrix} F_1\\ \vdots \\ F_m\end{pmatrix} \text{ auf } D \text{ nach}$$

$y_1,\ldots,y_m$ aufzulösen, also reellwertige Funktionen $f_1,\ldots,f_M : D\to\mathbb{R}$ zu bestimmen, dass gilt $F_j(x_1,\ldots,x_n,f_1(x_1,\ldots,x_n),\ldots,f(x_1,\ldots,x_n))=0, 1\le j\le m, \begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix}\in D$

Beispiele: $n=m, F(x,y)=Ay-x, A:\mathbb{R}^n\to\mathbb{R}^n, F:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$
$F(x,y)=0\Leftrightarrow Ay-x=0\Leftrightarrow Ay=x$. Ist A invertierbar, so gilt $y=f(x)=A^{-1}x$

(a) $m=n=1: F:\mathbb{R}\times\mathbb{R}\to\mathbb{R}, F(x,y)=y^3-x^2, F(x,y)=0\Leftrightarrow y^3-x^2=0\Leftrightarrow y^3=x^2\Leftrightarrow y=\sqrt[3]{x^2}$. Durch $F(x,y)=y^3-x^2=0$ wird eindeutig eine Funktion $y=f(x)$ definiert. Die Auflösung nach y ist eindeutig

(b) $F:\mathbb{R}\times\mathbb{R}\to\mathbb{R}, F(x,y):=x^2+y^2-1=0$
Funktionen sind z.B. $y_1=f_1(x)=\sqrt{1-x^2}, y_2=f_2(x)=-\sqrt{1-x^2}$ Durch die Gleichung $F(x,y)=x^2y^2-1=0$ wird nicht eindeutig eine Funktion $y=f(x)$ definiert. Auflösung nach y ist nicht eindeutig möglich. In den meisten Fällen interessiert man sich dafür, ob durch die Gleichung $F(x,y)=0$ in einer Umgebung (i.d.R. ε- Kugeln) $B_\varepsilon(\xi,\eta)$ mit $F(\xi,\eta)=0$ eine Funktion $y=f(x)$ mit $\eta=f(\xi)$ definiert wird (lokale Auflösbarkeit von $F(x,y)=0$ nach y).

Notation:

$$x=\begin{pmatrix} x_1\\ \vdots \\ x_n\end{pmatrix} \qquad y=\begin{pmatrix} y_1\\ \vdots \\ y_m\end{pmatrix}$$

$$\begin{pmatrix} x\\ y\end{pmatrix}=\begin{pmatrix} x_1\\ \vdots \\ x_n\\ y_1\\ \vdots\\ y_m\end{pmatrix} \qquad F:K\to\mathbb{R}^m$$

$$K\subset\mathbb{R}^n\times\mathbb{R}^m \qquad F=\begin{pmatrix} F_1\\ \vdots \\ F_m\end{pmatrix}$$

$$F_i:K\to\mathbb{R}$$

$$F(x,y)=F(x_1,\ldots,x_n,y_1,\ldots,y_m)=\begin{pmatrix} F_1((x_1,\ldots,x_n,y_1,\ldots,y_m)\\ \vdots \\ F_m(x_1,\ldots,x_n,y_1,\ldots,y_m)\end{pmatrix}$$

Falls die Punkte existieren, gilt:

$$\frac{\partial F}{\partial x} := \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}$$

ist eine $m \times n$-Matrix und

$$\frac{\partial F}{\partial y} := \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_n} \end{pmatrix}$$

ist eine $m \times m$-Matrix.

Satz 2.4.1. *Sei $D \subset \mathbb{R}^n, G \subset \mathbb{R}^m$ offen und die Funktion $F : D \times G \to \mathbb{R}^m$ stetig differenzierbar. Ferner seien $\xi \in D$ und $\eta \in G$ Punkte, für die $F(\xi, \eta) = 0$ und $\frac{\partial F}{\partial y}(\xi, \eta)$ invertierbar ist. Dann gilt:*

(i) Es existiert eine offene δ-Kugel $B_\delta(\xi) \subset D$ und eine offene ε-Kugel $B_\varepsilon(\eta) \subset G$ und genau eine stetige Funktion $f :_\delta (\xi) \to B_\varepsilon(\eta)$ mit $f(\xi) = \eta$ und $F(x, f(x)) = 0 \,\forall x \in B_\delta(\xi)$. Für jedes feste $x \in B_\delta(\xi)$ ist $f(x)$ die einzige Lösung in $B_\varepsilon(\eta)$ mit $F(x, f(x)) = 0$.

(ii) Es existiert eine weitere offene δ_1-Kugel $B_{\delta_1}(\xi), \delta_1 \le \delta$ und die Funktion $f : B_{\delta_1}(\xi) \to B_\varepsilon(\eta)$ ist stetig differenzierbar auf $B_{\delta_1}(\xi)$ und es gilt die Formel:

$$f'(x) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1} \frac{\partial F}{\partial x}(x, f(x))$$

Beweis in Heuser (292-299)

Eine Bemerkung zur Formel in (ii):

$$\begin{tikzcd} B_{\delta_1}(\xi) \arrow[r, "h"'] \arrow[d, "g"] & \mathbb{R}^m \\ D \times G \end{tikzcd}$$

$$h(x) = F(x, f(x))$$

$$g(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix} \qquad F = F(x, y)$$

$$h(x) = F(x, f(x)) \qquad h'(x) = F'(x, f(x))g(x)$$

$$F'(x, y) = \begin{pmatrix} \frac{\partial F}{\partial x} \frac{\partial F}{\partial y} \end{pmatrix} \qquad g'(x) = \begin{pmatrix} I \\ f'(x) \end{pmatrix}$$

$$F'(x, f(x))g'(x) = \frac{\partial F}{\partial x}(x, f(x))I + \frac{\partial F}{\partial y}(x, f(x))f'(x) = h'(x) \qquad h(x) = 0$$

$\Rightarrow h'(x) = 0 \Rightarrow \frac{\partial F}{\partial x}(x, f(x)) + \frac{\partial F}{\partial y}(x, f(x))f'(x) = 0$
$f'(x) = -\left(\frac{\partial F}{\partial y}(x, f(x))\right)^{-1} \frac{\partial F}{\partial x}(x, f(x))$.

Beispiel:

$$\begin{aligned} F(x,y) &= x^2 + y^2 - 1 && F : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ \frac{\partial F}{\partial y} &= 2y \neq 0 && \text{für } y \neq 0 \\ \left(\frac{\partial F}{\partial y}\right)^{-1} &= \frac{1}{2y} \end{aligned}$$

Damit ist die Gleichung $F(x, y) = x^2 + y^2 - 1 = 0$ nach y eindeutig in einer offenen Umgebung von (ξ, η) mit $F(\xi, \eta) = 0$ und $\eta \neq 0$ auflösbar. In $(-1, 0)$ und $(1, 0)$ ist $F(x, y) = 0$ nicht lokal nach y auflösbar.
$F(x, f(x)) = x^2 + f(x)^2 - 1 = 0, \frac{\partial F}{\partial x} = 2x$
$f'(x) = -\left(\frac{\partial F}{\partial y}\right)^{-1}(x, f(x)) \cdot \left(\frac{\partial F}{\partial x}\right)(x, f(x)) = -\frac{1}{2f(x)} \cdot 2x = -\frac{x}{f(x)}, y > 0, f(x) = \sqrt{1-x^2}, f'(x) = -\frac{x}{\sqrt{1-x^2}}$

2.4.1 Umkehrsatz

Problem: Unter welchen Bedingungen besitzt eine Funktion $f : D \to \mathbb{R}^n, D \subset \mathbb{R}^n$, eine Umkehrung? Welche analytischen Eigenschaften besitzt die Funktion f^{-1}?

$f(x) = Ax, A : \mathbb{R}^n \to \mathbb{R}^n, f(x) = Ax = y \to x = A^{-1}y, f^{-1}(y) = A^{-1}y, f'(x) = A$

Satz 2.4.2. *Sei $f : D \to \mathbb{R}^n, D \subset \mathbb{R}^n$ offen, stetig differenzierbar und in $\xi \in D$ sei $f'(\xi)$ invertierbar. Dann existiert eine offene Umgebung $U \in D$ von ξ und eine ε-Kugel von $\eta = f(\xi)$, so dass $f|_U : U \to B_\varepsilon(\eta)$ bijektiv die Umgebung U auf $B_\varepsilon(\eta)$ abbildet. Die Umkehrung f_U^{-1} von $f|_U$ ist stetig differenzierbar und für die Ableitung gilt:*

$$(f|_U^{-1})'(y) = f'(x)^{-1} \text{ mit } y = f(x) \quad x \in U$$

$$(f|_U^{-1})'(f(x)) = f'(x)^{-1} \quad x \in U$$

***Beweis**: $F : D \times \mathbb{R}^n \to \mathbb{R}^n, F(x, y) := f(x) - y = 0$*
F ist stetig differenzierbar, weil f stetig differenzierbar ist, mit $\frac{\partial F}{\partial x}(\xi, f(\xi)) = f'(\xi)$ und $F(\xi, \eta) = 0$ invertierbar. Satz 2.4.1 impliziert, dass es eine offene ε-Kugel $B_\varepsilon(\eta)$, eine offene δ-Kugel $B_\delta(\xi)$ und genau eine stetige Funktion $\varphi : B_\varepsilon(\eta) \to B_\delta(\xi), \varphi(\eta) = \xi, F(\varphi(y), y) = f(\varphi(y)) - y = 0$ gibt. Wegen Punkt (ii) aus Satz 2.4.1 können wir $B_\varepsilon(\eta)$ so klein gewählt denken, dass φ auf $B_\varepsilon(\eta)$ sogar stetig differenzierbar ist. Ferner gilt, da $f : D \to \mathbb{R}^n$ stetig ist, dass $f^{-1}(B_\varepsilon(\eta))$ offen in $\mathbb{R}^n$. Denn $x \in f^{-1}(B_\varepsilon(\eta)) \overset{f \text{ stetig}}{\Rightarrow} \exists 0 < r < \infty : B_r(x) \cap D \subset f^{-1}(B_\varepsilon(\eta)) \overset{D \text{ offen}}{\Rightarrow} \exists 0 < r_0 \leq r : B_{r_0}(x) \subset f^{-1}(B_\varepsilon(\eta)) \Rightarrow f^{-1}(B_\varepsilon(\eta))$ ist offen im $\mathbb{R}^n$.
Für $U := B_\delta(\xi) \cap f^{-1}(B_\varepsilon(\eta)) = \{x \in B_\delta(\xi) : f(x) \in B_\varepsilon(\eta)\}$ (offen) gilt $f(U) = B_\varepsilon(\eta)$.
Klar ist, dass $f(U) \subset B_\varepsilon(\eta)$. Wichtig zu zeigen, ist die Gleichheit:
Sei $y \in B_\varepsilon(\eta)$. Dann gilt: $\varphi(y) \in B_\delta(\xi)$ und $f(\varphi(y)) = y \in B_\varepsilon(\eta)$. Also $\varphi(y) \in$

$B_\delta(\xi) \cap f^{-1}(B_\varepsilon(\eta)) = U$, d.h. $f(U) = B_\varepsilon(\eta)$. f ist also surjektiv.
Injektivität: Seien $x_1, x_2 \in U : f(x_1) = f(x_2) = y \Rightarrow F(x_1, y) = F(x_2, y) = 0 \overset{\text{Satz 2.4.1}}{\Rightarrow}$ $x_1 = x_2$. Also ist $f|_U : u \to B_\varepsilon(\eta)$ von U auf $B_\varepsilon(\eta)$ bijektiv und $\varphi = f|_u^{-1}$ ist die stetig differenzierbare Umkehrabbildung von $f|_U$. Die Formel $(f|_U^{-1})'(f(x)) = f'(x)^{-1}$ folgt mit der Kettenregel:
$\varphi = f|_U^{-1}, \varphi(f(x)) = x, \varphi'(f(x)) = I \Rightarrow (f|_U^{-1})'(f(x)) = \varphi'(f(x))f'(x) = If'(x)^{-1}$ □

Bemerkungen:

1. Sei $f : U \to V$ eine bijektive Abbildung der offenen Mengen $U, V \subset \mathbb{R}^n$. Sind f, f^{-1} stetig differenzierbar, so heißt f ein *Diffeomorphismus*. In dieser Terminologie besagt der Umkehrsatz, dass die Funktion $f|_U$ bei hinreichend kleiner Umgebung U ein Diffeomorphismus ist.

2. Der Umkehrsatz ist eine „lokale" Aussage, d.h. existiert für eine stetig differenzierbare Funktion $f : D \to \mathbb{R}^n, f(x)^{-1}$, so existiert i.a. nicht die „globale" Inverse auf ganz D.
   Beispiel: $f : \mathbb{R}^2 \to \mathbb{R}^2, f(x,y) = \begin{pmatrix} e^x \cos y \\ e^x \sin y \end{pmatrix}$. f ist auf $\mathbb{R}^2$ stetig differenzierbar.
   $f'(x,y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}, f'(x,y)^{-1} = \begin{pmatrix} e^{-x} \cos y & e^{-x} \sin y \\ -e^{-x} \sin y & e^{-x} \cos y \end{pmatrix}$
   Aber $f : \mathbb{R}^2 \to \mathbb{R}^2$ ist nicht global invertierbar auf $\mathbb{R}^2$: $f(x,y) = f(x, y + 2k\pi)$ mit $k \in \mathbb{Z}$

Satz 2.4.3. *Sei $f : D \to \mathbb{R}^n, D \subset \mathbb{R}^n$ offen, stetig differenzierbar und existiert $f'(x)^{-1}$ für jedes $x \in D$. Dann gilt:*

(i) f ist eine offene Abbildung, *d.h. das Bild jeder offenen Teilmenge von D ist offen im $\mathbb{R}^n$.*
Beweis: $A \subset D$ offen in D. Da D offen ist, folgt, dass A in $\mathbb{R}^n$ offen ist.

1. Fall $A = \emptyset \Rightarrow f(A) = \emptyset$

2. Fall $a \neq \emptyset, y \in f(A)$ und $x \in A : f(x) = y$.
Da $f'(x)$ invertierbar ist, folgt es existiert eine offene Umgebung $U \subset A$ mit $x \in U$ und $f(U) \subset f(A)$ offen (nach dem Umkehrsatz), d.h. $f(A)$ ist offen.

(ii) Die Funktion $\varphi : D \to \mathbb{R}, \varphi(x) := ||f(x)||$, besitzt kein Maximum und, falls $f(x) \neq 0$ in D, auch kein Minimum on D.
Beweis: Annahme: $x_0 \in D$ sei eine Maximalstelle von φ, d.h. $\forall x \in D : ||f(x)|| \leq ||f(x_0)||$. Dann ist $f(x_0) > 0$. Denn wäre $f(x_0) = 0$, dann wäre $f = 0$ auf D und f' wäre auf D nirgends invertierbar.↯
Der Umkehrsatz impliziert: $\exists B_\varepsilon(f(x_0)) \subset f(D)$. Insbesondere gilt für $\rho = \frac{\varepsilon}{2||f(x_0)||}$ $\exists x_1 \in D\, y_1 = f(x_1) = f(x_0) + \rho f(x_0) = (1+\rho)f(x_0)$. Denn $||y_1 - f(x_0)|| = \rho||f(x_0)|| = \frac{\varepsilon}{2} < \varepsilon$. Somit hat man $||f(x_1)|| = (1+\rho)||f(x_0)|| > ||f(x_0)||$↯

(iii) Ist f injektiv, so ist $f^{-1}: f(D) \to \mathbb{R}$ stetig differenzierbar.
Beweis: $f(D)$ ist nach (i) offen. Nach dem Umkehrsatz und da f injektiv ist, folgt $f^{-1}: f(D) \to \mathbb{R}^n$ ist stetig differenzierbar.

$\square$

2.5 Extrema mit Nebenbedingungen

Seien $f : X \to \mathbb{R}, X \subset \mathbb{R}^n$ offen und $g : X \to \mathbb{R}^m, m < n$.
f besitzt in $\xi \in X$ ein lokales Maximum bzw. lokales Minimum unter der Nebenbedingung, $g(x) = 0 :\Leftrightarrow \xi \in N := \{x \in X : g(x) = 0\}$ und es existiert ein $B_\delta(\xi) \subset X \, \forall x \in B_\delta(\xi) \cap N : f(x) \leq f(\xi)$ (lokales Maximum) bzw.$f(x) \geq f(\xi)$ (lokales Minimum)

Problem: Gesucht sind die also die Stellen lokaler Extrema von f unter der angegebenen Nebenbedigung und die Werte, die f in ihnen annimmt.

Setzen $x = \begin{pmatrix} y \\ z \end{pmatrix}, y = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, z = \begin{pmatrix} x_{m+1} \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n-m}$. Dann $g(x) = g(y, z)$. Falls $g(y, z) = 0$ (m Komponentengleichungen) nach y auflösbar ist, d.h. $y = h(z)$ mit $g(h(z), z) = 0$, dann läuft das Problem der Extremwertbestimmung darauf hinaus, die „freien“ lokalen Extrema (d.h. lokale Extrema ohne Nebenbedingungen) der Funktion $\varphi(z) = f(h(z), z)$ zu bestimmen. Ist eine solche „explizite“ Auflösung von $g(y, z) = 0$ nach y nicht möglich, dann hilft uns die sog. Multiplikatorenmethode.

2.5.1 Lagrange Multiplikatorenmethode

Satz 2.5.1. *Sei $X \in \mathbb{R}^n$ eine offene Menge, $f : X \to \mathbb{R}, g := \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : X \to \mathbb{R}^m$ seien stetig differenzierbar und f besitzt in $\xi \in X$ ein lokales Extremum unter der Nebenbedingung $g(x) = 0$. Ferner existieren in $g'(\xi) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\xi) & \cdots & \frac{\partial g_1}{\partial x_n}(\xi) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(\xi) & \cdots & \frac{\partial x_m}{\partial x_n}(\xi) \end{pmatrix}$ eine m-reihige Unterdeterminante, die nicht verschwindet. Dann existieren m Zahlen $\lambda_1, \ldots, \lambda_m$ (*Lagrange Multiplikatoren*), so dass die Gleichung*

$$f'(\xi) + \sum_{i=1}^{m} \lambda_i g_i'(\xi) = 0$$

besteht.

Proof. O.B.d.A. sei die Determinante der Matrix

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\xi) & \cdots & \frac{\partial g_1}{\partial x_n}(\xi) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(\xi) & \cdots & \frac{\partial x_m}{\partial x_n}(\xi) \end{pmatrix} \neq 0 \tag{2.1}$$

Seien $x = \begin{pmatrix} y \\ z \end{pmatrix}, y = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, z = \begin{pmatrix} x_{m+1} \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n-m}, \xi = \begin{pmatrix} \eta \\ \zeta \end{pmatrix}, \eta = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_m \end{pmatrix}, \zeta = \begin{pmatrix} \xi_{m+1} \\ \vdots \\ \xi_n \end{pmatrix}$. Dann ist $f(x) = f(y,z), g(x) = g(y,z)$. Die Aussage, f hat in ξ ein lokales Extremum unter der Nebenbedingung $g(x) = 0$ impliziert:
$g(\xi) = g(\eta,\zeta) = 0$ und wegen Gleichung 2.1 ist $\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\eta,\zeta) & \dots & \frac{\partial g_1}{\partial x_n}(\eta,\zeta) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(\eta,\zeta) & \dots & \frac{\partial x_m}{\partial x_n}(\eta,\zeta) \end{pmatrix}$. Nach Satz 2.4.1 folgt $g(y,z) = 0$ ist stetig differenzierbar nach y auflösbar, d.h.

$$\exists B_\delta(\zeta) \subset \mathbb{R}^{n-m}\, h : B_\delta(\zeta) \to \mathbb{R}^m : h(\zeta) = \eta \wedge \forall z \in_\delta (\zeta) : g(h(z), z) = 0 \tag{2.2}$$

Ferner sei δ so klein, dass $\varphi(z) := f(h(z), z)$ differenzierbar ist. Es gilt:

$$\varphi'(\zeta) = \frac{\partial f}{\partial y}(\xi)h'(\zeta) + \frac{\partial f}{\partial z}(\xi)$$

f lokales Extremum in ξ unter der Nebenbedingung $g(x) = 0$ impliziert, dass $\varphi'(\zeta) = 0$. Also hat man

$$\frac{\partial f}{\partial y}(\xi)h'(\zeta) + \frac{\partial f}{\partial z}(\xi) = 0 \tag{2.3}$$

Ausserdem hat man wegen der Gleichung 2.2 mit Satz 2.4.1, dass

$$h'(\zeta) = -\left(\frac{\partial g}{\partial y}(\xi)\right)^{-1} \frac{\partial g}{\partial z}(\xi)$$

Einsetzen in die Gleichung 2.3 liefert: $\underbrace{-\frac{\partial f}{\partial y}(\xi)\left(\frac{\partial g}{\partial y}(\xi)\right)^{-1}}_{\lambda^T = (\lambda_1,\dots,\lambda_m)} \frac{\partial g}{\partial z}(\xi) + \frac{\partial f}{\partial z}(\xi) = 0$.

Somit ist $\frac{\partial f}{\partial y}(\xi) + \lambda^T \frac{\partial g}{\partial y}(\xi) = 0$ und $\frac{\partial f}{\partial z}(\xi) + \lambda^T \frac{\partial g}{\partial z}(\xi) = 0$. Beide Gleichungen liefern: $f'(\xi) + \lambda^T g'(\xi) = 0 \Leftrightarrow f'(\xi) + \sum_{i=1}^m \lambda_i g_i'(\xi) = 0$ $\square$

2.5.2 Praktische Lösungen von Extremalaufgaben

$f : X \subset \mathbb{R}^n \to \mathbb{R}, g : X \to \mathbb{R}^m, m < n$
Man betrachtet das System der $n + m$-Gleichungen

$$\frac{\partial f}{\partial x_k}(x) + \lambda_1 \frac{\partial g_1}{\partial x_k}(x) + \dots + \lambda_m \frac{\partial g_m}{\partial x_k}(x) = 0$$

und

$$g_j(x) = 0 \quad j = 1, \dots, m$$

für die $n+m$ Unbekannten $x_1, \dots, x_n, \lambda_1, \dots, \lambda_m$, die sich aus Satz 2.5.1 ergeben.

Aufgabe: Seien $f : X \subset \mathbb{R}^n \to \mathbb{R}, g = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix} : X \to \mathbb{R}^m$. f soll unter der Nebenbedingung $g(x) = 0$ extremal werden.

Merkregel:

1. Man bildet die Funktion $F : X \times \mathbb{R}^m \to \mathbb{R}, F(x, \lambda) := f(x) + \lambda^T g(x), x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$

2. Differentiation nach (x, λ) und Null setzen. $F'(x, \lambda) = 0 = \left(\frac{\partial F}{\partial x}, \frac{\partial F}{\partial \lambda}\right) = (f'(x) + \lambda^T g'(x), g^T(x)) = 0 \Leftrightarrow f'(x) + \lambda^T g'(x) = 0$

Ob in den Punkten $x = \xi$, die das obige Gleichungssystem erfüllen, lokale Extrema vorliegen, muss tatsächlich noch überprüft werden. Hinreichende Bedingungen lassen sich schwer angeben. In diesem Zusammenhang sind folgende Bemerkungen dabei nützlich:

(i) $N = \{x \in X : g(x) = 0\}$. Ist N kompakt, so besitzt die Funktion $f|_N$ ein Maximum und ein Minimum, d.h. f hat sogar ein globales und damit ein lokales Maximum und Minimum.

(ii) Seien $A, B \subset \mathbb{R}^n$ und $A \neq \emptyset$ ist kompakt und B ist abgeschlossen. Dann existiert $\xi \in A, \eta \in B : ||\xi - \eta|| = \inf_{\substack{x \in A \\ y \in B}} ||x - y||$
Die Kompaktheit der Menge ist notwendig, denn:
$A = \{(t, \sqrt{t^2 - 1} : 1 \leq t \leq \infty\}, \nexists \xi \in A, \eta \in B : ||\xi - \eta|| = \inf_{\substack{x \in A \\ y \in B}} ||x - y||$

Beispiel. 1. Bestimme Minima und Maxima der Funktion $f(x, y, z) = 5x + y - 3z$ unter den Nebenbedingungen: $x + y + z = 0$ und $x^2 + y^2 + z^2 = 1$:

$$N = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : g(x, y, z) = 0 \right\} \qquad g(x, y, z) = \begin{pmatrix} x + y + z \\ x^2 + y^2 + z^2 - 1 \end{pmatrix}$$

$$g : \mathbb{R}^3 \to \mathbb{R}^2 \qquad g'(x, y, z) = \begin{pmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \end{pmatrix}$$

Eine zweireihige Unterdeterminante von g' auf N ist $\neq 0$.
$F(x, y, z, \lambda, \nu) = f(x, y, z) + \lambda(x + y + z) + \nu(x^2 + y^2 + z^2 - 1) = 5x + y - 3z +$

$\lambda(x+y+z)+\nu(x^2+y^2+z^2-1), F'(x,y,z,\lambda,\nu)=0 \Leftrightarrow$

$$\frac{\partial F}{\partial x} = 5+\lambda+2\nu x = 0 \tag{2.4}$$
$$\frac{\partial F}{\partial y} = 1+\lambda+2\nu y = 0 \tag{2.5}$$
$$\frac{\partial F}{\partial z} = -3+\lambda+2\nu z = 0 \tag{2.6}$$
$$\frac{\partial F}{\partial \lambda} = x+y+z = 0 \tag{2.7}$$
$$\frac{\partial F}{\partial \nu} = x^2+y^2+z^2-1 = 0 \tag{2.8}$$

2.4, 2.5 und 2.6: $3+3\lambda+2\nu(\underbrace{x+y+z}_{=0}) = 0 \Rightarrow 3+3\lambda = 0 \Rightarrow \lambda = -1 \Rightarrow$

$\left.\begin{array}{rcl} 4+2\nu x &=& 0 \\ 2\nu y &=& 0 \end{array}\right\} \Rightarrow \nu \neq 0, y = 0 \overset{2.7,2.8}{\Rightarrow} \left.\begin{array}{rcl} x+z &=& 0 \\ x^2+y^2 &=& 1 \end{array}\right\} \Rightarrow 2x^2 = 1 \Rightarrow x = \pm\frac{1}{\sqrt{2}}, z = \mp\frac{1}{\sqrt{2}}$

2.4: $4+2\nu x = 0, x = \pm\frac{1}{\sqrt{2}}, 4+\frac{2}{\sqrt{2}}\nu = 0, \nu = \pm 2\sqrt{2}$

$(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}, -1, -2\sqrt{2})$

Die Punkte $(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}), (-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$ erfüllen das Gleichungssystem 2.4 bis 2.8 in Verbindung mit $\lambda = -1, \nu = -2\sqrt{2}$ bzw. $\nu = 2\sqrt{2}$.

Einsetzen in f: $f(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}) = 5\frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} = \frac{8}{\sqrt{2}} = 4\sqrt{2}$ und $f(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}) = -5\frac{1}{\sqrt{2}} - \frac{3}{\sqrt{2}} = -\frac{8}{\sqrt{2}} = -4\sqrt{2}$

Die Menge N ist kompakt. Da f stetig ist, nimmt f auf N sowohl ihr Maximum als auch ihr Minimum an. Daher ist $4\sqrt{2}$ das Maximum und $-4\sqrt{2}$ das Minimum von f auf N.

Globale Extrema direkt: Sei

$$N = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 : \begin{array}{c} x+y+z=0 \\ x^2+y^2+z^2=1 \end{array} \right\}$$
$$f(x,y,z) = 5x+y-3z = \underbrace{x+y+z}_{=0} + 4(x-z) \text{ auf } N$$

Es gilt folgende Abschätzung:

$$\begin{aligned} |f(x,y,z)| &= 4|x-z| \leq 4(|x|+|z|) \\ &\leq 4(\sqrt{1^2+1^2})\sqrt{|x|^2+|z|^2} = 4\sqrt{2}(\sqrt{|x|^2+|z|^2}) \\ &\leq 4\sqrt{2} \end{aligned}$$

Also ist $-4\sqrt{2} \leq f(x,y,z) \leq 4\sqrt{2}$ mit $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in N$.

2. Sei $A = (a_{ij})_{i,j=1}^n$ eine reelle symmetrische $n \times n$-Matrix ($A = A^T$). Man bestimme die Extrema $f(x) = \langle Ax, x\rangle$ unter der Nebenbedingung $||x||_2 = 1, x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$.

   (Uebung)

3. Man bestimme die Extrema der Funktion $f(x) = \sum_{i=1}^n a_i x_i$ mit $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, a_i > 0, x_i > 0$ mit der Nebenbedingung $\sum_{i=1}^n \frac{a_i b_1}{x_1} = 1$ mit $b_1 > 0$.

   $N = \{x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n : x_1, \ldots, x_n > 0 \wedge \sum_{i=1}^n \frac{a_i b_i}{x_i} = 1\}$

   $g(x) = g(x_1, \ldots, x_n) = \sum_{i=1}^n \frac{a_i b_i}{x_i} - 1$

   $F(x, \lambda) = f(x) + \lambda g(x) = \sum_{i=1}^n a_i x_i + \lambda(\sum_{i=1}^n \frac{a_i b_i}{x_i} - 1), F'(x, \lambda) = 0 \Leftrightarrow \frac{\partial F}{\partial x_i} = a_i + \lambda(-\frac{a_i b_i}{x_i^2} = 0 \Leftrightarrow \begin{matrix} a_1 - \lambda \frac{a_1 b_1}{x_1^2} = 0 \\ \vdots \\ a_n - \lambda \frac{a_n b_n}{x_n^2} = 0 \end{matrix}$ und $\frac{\partial F}{\partial \lambda} = \sum_{i=1}^n \frac{a_i b_i}{x_i} - 1 = 0 \Leftrightarrow \sum_{i=1}^n \frac{a_i b_i}{x_i} = 1$

   Somit folgt $\lambda = \frac{x_1^2}{b_2} = \cdots = \frac{x_n^2}{b_n} \Leftrightarrow x_i = \sqrt{\lambda b_i} \Rightarrow 1 = \sum_{i=1}^n \frac{a_i b_i}{x_i} = \sum_{i=1}^n \frac{a_i b_i}{\sqrt{\lambda b_i}} = \sum_{i=1}^n \frac{a_i \sqrt{b_1}}{\sqrt{\lambda}} \Rightarrow \lambda = \left(\sum_{i=1}^n a_i \sqrt{b_i}\right)^2 \Rightarrow x_k = \sqrt{b_k}\sqrt{\lambda} = \sqrt{b_k}\left(\sum_{i=1}^n a_i \sqrt{b_i}\right)$

   Für f ergibt sich folgendes:

   $f(x_1, \ldots, x_n) = \sum_{k=1}^n a_k x_k = \sum_{k=1}^n a_k \sqrt{b_k} (\sum_{i=1}^n a_i \sqrt{b_1}) = \left(\sum_{i=1}^n a_i \sqrt{b_i}\right)^2$

   N ist *nicht* kompakt, da N nicht beschränkt ist.

   Behauptung: f besitzt in $\xi = \left(\sqrt{b_k} \sum_{i=1}^n a_i \sqrt{b_i}\right)_{k=1}^n$ ein Minimum unter der Nebenbedingung N:

$$
\begin{aligned}
f(\xi) = \left(\sum_{i=1}^n a_i \sqrt{b_i}\right)^2 &= \sum_{i=1}^n \left(\sqrt{a_i x_i}\sqrt{\frac{a_i b_i}{x_i}}\right)^2 \\
&\leq \left(\sqrt{\sum_{i=1}^n (\sqrt{a_i x_i})^2}\sqrt{\sum_{i=1}^n (\sqrt{\frac{a_i b_i}{x_i}})^2}\right)^2 \\
&= \left(\sum_{i=1}^n a_i x_i\right)\left(\sum_{i=1}^n \frac{a_i b_i}{x_i}\right) = f(x)
\end{aligned}
$$

   für $x \in N$. Daher folgt, dass f in ξ ein globales Minimum besitzt.

2.6 Geometrische Begriffe

Definition 2.6.1. *Unter einer* Kurve *im $\mathbb{R}^n$ versteht man eine stetige Abbildung $f : I \to \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein Intervall ist. Die Kurve f heißt differenzierbar (bzw. stetig* *differenzierbar), wenn f differenzierbar (bzw. stetig differenzierbar) ist.*

Definition 2.6.2. *Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}^n$ eine differenzierbare Kurve. Für $t \in I$ heißt $f'(t)$ der* Tangentialvektor *der Kurve f zum Parameter t. Falls $f'(t) \neq 0$, heißt $\frac{f'(t)}{||f'(t)||_2}$ Tangentialeinheitsvektor.*

2.6.1 Geometrische Interpretation

Der Tangentialvektor $f'(t)$ lässt sich als Limes von Sekanten auffassen. Denn $f'(t) = \lim_{h\to 0} \frac{f(t+h)-f(t)}{h}$

Bemerkung: Eine Kurve $f : I \to \mathbb{R}^n$ braucht nicht injektiv zu sein. Gilt $f(t_1) = f(t_2) = x$ für $t_1 \neq t_2$, so heißt x *Doppelpunkt* der Kurve f. In x hat dann f zwei i.allg. verschiedene Tangentialvektoren.

Beispiel:

$$f : \mathbb{R} \to \mathbb{R}^2 \qquad f(t) = \begin{pmatrix} t^2 - 1 \\ t^3 - t \end{pmatrix}$$

$$x = t^2 - 1 \qquad y = t^3 - t = t(t^2 - 1)$$

$$y^2 = t^2(t^2 - 1)^2 = x^2 + x^3 f(\mathbb{R}) = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : y^2 = x^2 + x^3 \right\}$$

$$f(-1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = f(1) \qquad f'(t) = \begin{pmatrix} 2t \\ 3t^2 - 1 \end{pmatrix}$$

$$f'(-1) = \begin{pmatrix} -2 \\ 2 \end{pmatrix} \qquad f'(1) = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

Definition 2.6.3. *Sei $f : I \to \mathbb{R}^n$ eine stetig differenzierbare Kurve. Die Kurve heißt* regulär, *falls $f'(t) \neq 0$ für alle $t \in I$. Ein Parameter $t \in I$ mit $f'(t) = 0$ heißt* singulär.

Beispiel: $f : \mathbb{R} \to \mathbb{R}^2, f(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}, x = t^2, y = t^3, y^2 = x^3, y = \pm\sqrt{x^3}$

$f'(t) = \begin{pmatrix} 2t \\ 3t \end{pmatrix}, f'(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. In $t = 0$ liegt ein singulärer Punkt vor.

Definition 2.6.4. ***Schnittwinkel zweier Kurven*** *Seien $f : I_1 \to \mathbb{R}^n, g : I_2 \to \mathbb{R}^n$ zwei reguläre Kurven. Für $t_1 \in I_1, t_2 \in I_2$ gelte $f(t_1) = g(t_2)$. Unter dem* Schnittwinkel *ϑ der Kurven f und g bei den Parametern t_1 bzw. t_2 versteht man den Winkel zwischen den Tangentialvektoren $f'(t_1)$ und $g'(t_2)$. Der Winkel ϑ wird also bestimmt durch*

$$\cos\vartheta = \frac{\langle f'(t_1), g'(t_2)\rangle}{||f'(t_1)||_2||g'(t_2)||_2} \quad 0 \leq \vartheta \leq \pi$$

2.6.2 Flächen und Tangentialebenen

Definition 2.6.5. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$. Die Menge*

$$graph\ (f) = \left\{ \begin{pmatrix} x \\ f(x) \end{pmatrix} \in \mathbb{R}^{n+1} : x \in D \right\}$$

heißt Fläche *im* $\mathbb{R}^{n+1}$.

Definition 2.6.6. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen und differenzierbar. Die Menge* $T := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^{n+1} : x \in \mathbb{R}^n, y = f(\xi) + f'(\xi)(x - \xi) \right\}$ *heißt* Tangentialebene *von f im Punkt* $\begin{pmatrix} \xi \\ f(\xi) \end{pmatrix}$ *mit $\xi \in D$.*

Bemerkung:

(1) Im Fall $n = 1$ benutzt man den Begriff Kurve statt Fläche und Tangente statt Tangentialebene.

(2) Die Tangentialebene T kann man sich im Punkt $(\xi, f(\xi))^T$ wie folgt erzeugt denken: Man betrachtet folgende Kurven in f:

$$g_i(t) = (\xi_1, \dots, \xi_{i-1}, t, \xi_{i+1}, \dots, \xi_n, f(\xi_1, \dots, \xi_{i-1}, t, \xi_{i+1}, \dots, \xi_n))^T$$
$$g_i : T \to \mathbb{R}^{n+1} \qquad g_i(\xi_i) = (f(\xi))$$

Tangentialvektoren: $g_i(\xi_i) = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \frac{\partial f}{\partial x_i}(\xi) \end{pmatrix}$

Dann:

$$\begin{aligned} T &= \begin{pmatrix} \xi \\ f(\xi) \end{pmatrix} + \operatorname{span} \{g_1'(\xi), \dots, g_n'(\xi)\} \\ &= \left\{ \begin{pmatrix} \xi \\ f(\xi) \end{pmatrix} + \sum_{i=1}^{n} h_i g_i'(\xi) : h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n \right\} \\ &= \left\{ \begin{pmatrix} \xi + h \\ f(\xi) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_1}(\xi) \end{pmatrix} : h \in \mathbb{R}^n \right\} \qquad h_i = x_i - \xi_i \end{aligned}$$

2.6.3 Niveauflächen

Definition 2.6.7. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$. Dann heißt die Menge $\{x \in D : f(x) = c\}$* Niveaufläche *der Funktion f zum Niveau c.*

Ein Beispiel hierfür sind Wetterkarten (Isobaren).

Bemerkung: Man kommt von einer Niveaufläche $f(x) = c_1$ zu einer „unmittelbar benachbarten" Niveaufläche $f(x) = c_2$ mit $c_2 - c_1$ „klein" am schnellsten in Richtung des Gradienten:

$$c_2 - c_1 = f(x + h)f(x) \approx f'(x)h$$

Es gilt:

$$\max\{|f'(x)h| : ||h||_2 \leq 1\} = ||f'(x)||_2$$

für $h = \frac{(f'(x)^T}{||f'(x)||_2}$ Richtung des Gradienten.

2.6.4 Niveauflächen und Tangentialebenen

Satz 2.6.8. *Sei $f : D \to \mathbb{R}, D \subset \mathbb{R}^n$ offen, stetig differenzierbar und $f'(\xi) \neq 0$ in $\xi \in D$. Dann gilt: Die Tangentialebene an die Niveaufläche $f(x) = c$ mit $c = f(\xi)$ ist charakterisiert durch:*

$$\langle f'(\xi), x - \xi \rangle = 0$$

In Komponenten:

$$\sum_{i=1}^{n} (x - \xi) \frac{\partial f}{\partial x_i}(\xi_1, \ldots, \xi_n) = 0$$

***Beweis**: Sei o.B.d.A. $\frac{\partial f}{\partial x_n}(\xi) \neq 0$. Dann existieren nach dem Auflösungssatz der impliziten Funktion eine Umgebung U von $\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_{n-1} \end{pmatrix}$ und eine Umgebung V von ξ_n und $\varphi : U \to V$ mit*

1. $\varphi(\xi_1, \ldots, \xi_{n-1}) = \xi_n$
2. $f(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_n)) = 0 \quad \forall \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \in U$

T ist die Tangentialebene von φ im Punkt $\begin{pmatrix} \tilde{\xi} \\ \varphi(\tilde{\xi} \end{pmatrix} = \xi$ mit $\tilde{\xi} = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_{n-1} \end{pmatrix}$

$$T = \left\{ \begin{pmatrix} \tilde{x} \\ \varphi(\tilde{\xi} + \varphi'(\tilde{\xi})(\tilde{x} - \tilde{\xi}) \end{pmatrix} : \tilde{x} \in U \right\}$$

Sei $y \in T$. Dann $y_n = \varphi(\tilde{\xi} + \varphi'(\tilde{\xi})(\tilde{y} - \tilde{\xi}) = \xi_n - \frac{\partial f}{\partial x_n}(\xi)^{-1} \frac{\partial f}{\partial \tilde{x}}(\xi)(\tilde{y} - \tilde{\xi}) \Leftrightarrow \frac{\partial f}{\partial x_n}(\xi)(y_n - \xi_n) = \underbrace{-\frac{\partial f}{\partial x_n}(\xi)(\tilde{y} - \tilde{\xi})}_{=\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi)(y - \xi_i)} \Rightarrow \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi)(y_i - \xi_i) = 0$ $\square$

3 Gewöhnliche Differentialgleichungen

3.1 Einführung

Definition 3.1.1. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$, $f : G \to \mathbb{R}^n$ stetig. Dann heißt*

$$y' = f(x, y) \tag{3.1}$$

ein System von n Differentialgleichungen 1. Ordnung.

Für $n = 1$ heißt 3.1 eine einfache Differentialgleichung *1. Ordnung.*

Unter einer Lösung von 3.1 versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$\varphi : I \to \mathbb{R}^n$$

mit folgenden Eigenschaften:

a) graph $\varphi \subset G$, graph $\varphi = \{(x, \varphi(x)) | x \in I\}$

b) Es gilt: $\varphi'(x) = f(x, \varphi(x))$

In Koordinatenschreibweise:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

$$y' = f(x, y) \Leftrightarrow \begin{matrix} y_1' = f_1(x, y_1, \ldots, y_n) \\ \vdots \\ y_n' = f_n(x, y_1, \ldots, y_n) \end{matrix}$$

Die Lösung $\varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix}$ erfüllt $\varphi_i' = (f_i(x, \varphi_1(x), \ldots, \varphi_n(x)))$.

Bemerkungen:

1. Nach dem Hauptsatz der Differential- und Integralrechnung ist b) äquivalent zu

$$\varphi(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))dt \text{ mit } \varphi(x_0) = y_0$$

Für $y' = f(x)$ erhält man als Lösung $\varphi(x) = c + \int_{x_0}^{x} f(t)dt, \varphi(x_0) = c$.

2. Implizite Gestalt eines DGL-Systems 1 Ordnung:

$$F(x, y, y') = 0$$

$$F : G \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \text{ stetig}$$

3. Treten in 3.1 neben x weitere unabhängige Variable auf sowie partielle Ableitungen, so spricht man von *partiellen Differentialgleichungen*.
   Beispiel: Wärmeleitungsgleichungen:
   $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, u = u(x,t)$
   Korteweg-deVries-Gleichungen:
   $\frac{\partial u}{\partial t} = \frac{\partial^3 u}{\partial x^3} + 6u\frac{\partial u}{\partial x}$

Anfangswertproblem Folgende Aufgabe heißt Anfangswertproblem für ein DGL-System 1. Ordnung:
Sei $(x_0, y_0) \in G$.
Gesucht ist eine Lösung $y = \varphi(x)$ von $y' = f(x, y)$ mit $\varphi(x_0) = y_0$.

Beispiel: $y' = f(x), f : I \to \mathbb{R}^n$ stetig, $x_0 \in I, \varphi(x_0) = y_0$
Integration: $\varphi(x) = y_0 + \int_{x_0}^x f(t)dt$
$\int_{x_0}^x f(t)dt := \begin{pmatrix} \int_{x_0}^x f_1(t)dt \\ \vdots \\ \int_{x_0}^x f_n(t)dt \end{pmatrix}$

Geometrische Intepretation:
Eine DGL $y' = f(x, y)$ in einem Gebiet G, $G \subset \mathbb{R} \times \mathbb{R}^n$, bestimmt ein *Richtungsfeld*:

$$\{(x, y, y') : (x, y) \in G, y' = f(x, y)\}$$

(x, y, y') heißt *Linienelement*. In jedem Punkt $(x, y) \in G$ wird durch $y' = f(x, y)$ eine Steigung vorgegeben.

Problem: Gesucht wird eine Funktion $y = \varphi(x)$, deren Tangente im Punkt $x, \varphi(x))$ die Richtung besitzt, die durch das Richtungsfeld vorgeschrieben ist, d.h. die auf das Richtungsfeld passt.

Kurven konstanten Anstiegs heissen *Isoklinen*.

Beispiel: $y' = f(x, y) = x, G \subset \mathbb{R} \times \mathbb{R}$
Die Lösungen sind Parabeln: $\varphi(x) = \frac{x^2}{2} + c$

Die allgemeine Theorie der gewöhnlichen Differentialgleichungen beschäftigt sich mit folgenden Problemstellungen für das AWP $y' = f(x, y), \varphi(x_0) = y_0$.

1. Existenz einer Lösung
2. Eindeutigkeit der Lösung
3. Berechnung von Lösungen
4. Abhängigkeit der Lösung von Anfangsbedingungen (z.B. ob stetige Abhängigkeit vorliegt)
5. Stabilitätsverhalten von Lösungen

Lösungsvielfalt Die Gesamtheit aller Lösungen

$$L = \{y = \varphi(x) | \varphi'(x) = f(x, \varphi(x))\}$$

heißt *allgemeine Lösung* von $y' = f(x, y)$.

Ein $\varphi \in L$ heißt *spezielle Lösung* von $y' = f(x, y)$.

Die allgemeine Lösung kann oft durch Parameter beschrieben werden.

3.2 Reduktion von Differentialgleichungssystemen höherer Ordnung auf Differentialgleichungssysteme 1. Ordnung

Definition 3.2.1. *$G \subset \mathbb{R} \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n$, $f : G \to \mathbb{R}^n$ stetig*
Dann heißt

$$y^{(k)} = f(x, y, y', \ldots, y^{(k)}) \tag{3.2}$$

ein System von n Differentialgleichungen k-ter Ordnung.

Für $k = 1$ erhält man Gleichung 3.1 und für $n = 1$ heißt die Gleichung 3.2 einfach Differentialgleichung k-ter Ordnung.

Unter einer Lösung von 3.2 versteht man eine Funktion φ, die auf einem Intervall $I \subset \mathbb{R}$ definiert ist und k-mal differenzierbar ist, mit folgenden Eigenschaften:

a) $\{(x, \varphi(x), \varphi'(x), \ldots, \varphi^{(k-1)}(x)) | x \in I\} \subset G$

b) $\varphi^{(k)} = f(x, \varphi(x), \ldots, \varphi^{(k-1)}(x))$

Koordinatenschreibweise: $f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, d.h. wegen Gleichung 3.2 $\Leftrightarrow y_i^{(k)} = f_i(x, y_1, \ldots, y_n, y_1', \ldots, y_n', \ldots, y_1^{(k-1)}, \ldots, y_n^{(k-1)})$

Reduktion von 3.2 auf DGL-Systeme 1. Ordnung Man setzt $Y := \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{k-1} \end{pmatrix} \in (\mathbb{R}^n)^k$

$F(x, Y) := \begin{pmatrix} y_1 \\ \vdots \\ y_{k-1} \\ f(x, y_0, \ldots, y_{k-1}) \end{pmatrix}$. Dann gilt:

a) φ ist eine Lösung von 3.2 $\Rightarrow \Phi := \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(k-1)} \end{pmatrix}$ ist eine Lösung von

$$Y' = F(x, Y) \tag{3.3}$$

Denn: Sei $\varphi : I \to \mathbb{R}^n$ eine Lösung von 3.2, d.h. $\varphi^{(k)}(x) = f(x, \varphi(x), \ldots, \varphi^{(k-1)}(x)) \Rightarrow$

$$\begin{aligned}
\Phi' &= \begin{pmatrix} \varphi'(x) \\ \varphi''(x) \\ \vdots \\ \varphi^{(k)}(x) \end{pmatrix} \\
&= \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_{k-1} \\ f(x, \varphi_0(x), \ldots, \varphi_{k-1}(x)) \end{pmatrix} \\
&= F(x, \varphi_0(x), \ldots, \varphi_{k-1}(x)) \\
&= F(x, \Phi(x))
\end{aligned}$$

b) Φ ist eine Lösung von 3.3. Dann ist $\varphi := \varphi_0$ eine Lösung von 3.3.
Denn: Sei Φ eine Lösung von $Y' = F(x, Y) \Rightarrow \varphi_0' = \varphi_1 \ldots$. Setze $\varphi := \varphi_0 \Rightarrow \varphi Y' = F(x, Y)Y' = F(x, Y)i_1 = \varphi', \varphi_2 = \varphi'', \ldots, \varphi_{k-1} = \varphi^{(k-1)}$ und $\varphi^{(k)}(x) = f(x, \varphi(x), \ldots, \varphi^{(k-1)}(x))$. Daraus folgt die Aussage.

Fazit: Die Lösungen 3.2 und 3.3 sind in eineindeutiger Beziehung zueinander.

Festlegung von C durch Anfangsbedingung $\varphi(x_0) = y_0$ Die Struktur der allgemeinen Lösung $\varphi(x)$ der inhomogenen DGL $y' = a(x)y + b(x)$ kann auch wie folgt charakterisiert werden:
$\varphi(x) = C\varphi_0(x) + \varphi_{\text{inh}}(x), \varphi_0(x) = e^{\int_{x_0}^{x} a(t)dt}$
Beweis: Sei $\psi(x)$ eine Lösung des inhomogenen DGL $y' = a(x)y + b(x)$. Wir zeigen, $\exists C \in \mathbb{R} : \psi(x) = C\varphi_0(x) + \varphi_{\text{inh}}(x)$ und setzen $h(x) := \psi(x) - \varphi_{\text{inh}}(x)$. Dann gilt $h'(x) = \psi'(x) - \varphi_{\text{inh}}'(x) = a(x)\psi(x) + b(x) - (a(x)\varphi_{\text{inh}}(x)b(x)) = a(x)(\psi(x) - \varphi_{\text{inh}}(x)) = a(x)h(x) \Rightarrow h(x)$ ist Lösung des homogenen DGL $y' = a(x)y$. Somit existiert ein $C \in \mathbb{R} : h(x) = C\varphi_0(x)$ oder $\psi(x) - \varphi_{\text{inh}}(x) = C\varphi_0(x) \Rightarrow \psi(x) = C\varphi_0(x) + \varphi_{\text{inh}}(x)$

Beispiel: Gesucht ist die allgemeine Lösung von $y' = ay + 1$.

1. Schritt Berechnung der allgemeinen Lösung des homogenen DGL $y' = ay$
$\varphi(x) = C\varphi_0(x) = Ce^{\int_0^x a dt} = Ce^{ax}$

2. Schritt Finden einer speziellen Lösung der inhomogenen DGL $y' = ay + 1$
Die erste Möglichkeit ist die Variation der Konstanten, die zweite ist Raten.
$\varphi_{\text{inh}}(x) = -\frac{1}{a}$ $0 = \varphi_{\text{inh}}(x) = aa\varphi_{\text{inh}}(x) + 1 = a(-\frac{1}{a}) + 1 = 0$
Allgemeine Lösung der inhomogenen DGL = allgemeine Lösung des homogenen DGL + spezielle Lösung des inhomogenen DGL

$$\varphi(x) = Ce^{ax} - \frac{1}{a} \quad \varphi(x) = Ce^{ax} + e^{ax} - \frac{1}{a}$$

Zur ersten Möglichkeit (Variation der Konstanten):
Ansatz: $\varphi(x) = e^{ax}C$ Festlegung von C durch Anfangsbedingung (x)

Einsetzen in inhomogene DGL: $y' = ay + 1$
$\varphi'(x) = ae^{ax}C(x) + e^{ax}C'(x) = a\varphi(x) + 1 = ae^{ax}C(x) + 1 \Rightarrow e^{ax}C'(x) = 1 \Rightarrow C'(x) = e^{-ax} \Rightarrow C(x) - \underbrace{C(0)}_{=C} int_0^x a^{-at}dt = -\frac{1}{a}e^{-ax} + \frac{1}{a} \Rightarrow C(x) = C - \frac{1}{a}e^{-ax} + \frac{1}{a}$
$\varphi(x) = e^{ax}C(x) = e^{ax}(C - \frac{1}{a}e^{-ax} + \frac{1}{a}) = e^{ax}C - \frac{1}{a} + \frac{1}{a}e^{ax}$

3.2.1 Beispiele für Differentialgleichungen, die auf $y' = f(x)g(y)$ zurückgeführt werden können

(a) Die inhomogene DGL $y' = f(\frac{y}{x})$
Sei $I \subset \mathbb{R}$ ein Intervall, $f : I \to \mathbb{R}$ stetig, $G := \{(x,y) \in \mathbb{R}\backslash\{0\}\times\mathbb{R} | \frac{y}{x} \in I\}$. Dann heißt $y' = f(\frac{y}{x})$ homogene DGL. Es gilt, $\varphi(x)$ ist eine Lösung von $y' = f(\frac{y}{x}) \Leftrightarrow \psi(x) = \frac{\varphi(x)}{x}$ ist Lösung von $z' = \frac{1}{x}(f(z) - z)$.
Beweis:

„⇒" Sei φ Lösung von $y' = f(\frac{y}{x})$. Dann $\varphi'(x) = \frac{\varphi'(x)x = \varphi(x)}{x^2} = \frac{1}{x}(\varphi'(x) - \frac{\varphi(x)}{x}) = \frac{1}{x}(f(\frac{\varphi(x)}{x}) - \frac{\varphi(x)}{x}) = \frac{1}{x}(f(\psi(x)) - \psi(x)) \Rightarrow \psi(x) = \frac{\varphi(x)}{x}$ ist Lösung von $z' = \frac{1}{x}(f(z) - z)$

„⇐" Sei ψ Lösung von $z' = \frac{1}{x}(f(x) - x)$. Dann gilt für $\varphi(x) = \psi(x) \quad x : \varphi'(x) = \psi'(x)x + \psi(x) = \frac{1}{x}(f(\psi(x)) - \psi(x))x + \psi(x) = f(\psi(x)) - \psi(x) + \psi(x) = f(\psi(x)) = f(\frac{\varphi(x)}{x})$
φ ist Lösung des DGL $y' = f(\frac{y}{x})$

(b) Die Differentialgleichung $y'' = f(y), f : J \to \mathbb{R}$ stetig
Die Lösung erfolgt mit Hilfe der sog. Energiemethode: Sei $\varphi(x)$ Lösung von $y'' = f(y)$. Dann Multiplikation mit $\varphi'(x)$: $\varphi''(x)\varphi'(x) = f(\varphi(x)\varphi'(x) \Leftrightarrow \frac{1}{2}\frac{d}{dx}((\varphi'(x)^2)) = \frac{d}{dx}\int_{x_0}^{x} f(\varphi(z))\varphi'(z)dz = \frac{d}{dx}\int_{\varphi(x_0)}^{\varphi(x)} f(t)dt$
Physiker setzen dann: $U(y) := -\int_{y_0}^{y} f(t)dt \Leftrightarrow \frac{d}{dx}(\frac{1}{2}(\varphi'(x))^2 + U(\varphi(x))) = 0 \Rightarrow \frac{1}{2}(\varphi'(x))^2 + U(\varphi(x)) = E =$konstant
$y' = \pm\sqrt{2(E - U(y))}$ - Differentialgleichung mit getrennten Variablen

3.3 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten

3.3.1 Definition

Sei $\mathfrak{M}_n(\mathbb{K}) := \{A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} : a_{ij} \in \mathbb{K}\}$ die Menge aller $n \times n$-Matrizen. Dann heißt

$$y' = Ay \; y \in \mathbb{R}^n \; A \in \mathfrak{M}_n(\mathbb{K}) \tag{3.4}$$

homogenes lineares Differentialgleichungssystem 1. Ordnung mit konstanten Koeffizienten. Weiter sei $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} : I \to \mathbb{R}^n$ eine stetige Funktion. Dann heißt

$$y' = Ay + b(x) \tag{3.5}$$

ein inhomogenes lineares Differentialgleichungssystem 1. Ordnung mit konstanten Koeffizienten. In Komponenten lautet die Gleichung 3.5 wie folgt:

$$\begin{aligned} y_1' &= a_{11}y_1 + \cdots + a_{1n}y_n + b_1(x) \\ &\vdots \\ y_n' &= a_{n1}y_1 + \cdots + a_{nn}y_n + b_n(x) \end{aligned}$$

Bemerkungen: Die Funktion $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n, f(x, y) := Ay$ ist linear in y. Daher stammt der Name lineare Differentialgleichung. Im eindimensionalen Fall $n = 1$ hat man die allgemeine Lösung der homogenen Differentialgleichung $y' = ay, \varphi(x) = e^{ax}c, c \in \mathbb{R}$. Mehrdimensional $n > 1$ hätte man folgende Analogie:

$$y' = Ax, \varphi(x) = e^{Ax}c, c \in \mathbb{R}^n$$

Die entstehen folgende Probleme:

(a) Was versteht man unter e^{Ax}?

(b) Wie berechnet man e^{Ax}?

3.3.2 Operatorfunktion (Matrixfunktionen)

Definition 3.3.1. *Sei $(A_k) \subset \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$ eine Folge. Die Reihe $\sum_{k=0}^{\infty} A_k$ heißt genau dann konvergent, wenn die Folge der Partialsummen $(\sum_{k=0}^{n} A_k)_{n=1}^{\infty}$ in $\mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$ konvergent ist.*

Beispiel: Sei $A \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n), e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$ ist für jede Matrix $A \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n)$ konvergent. Denn $a_n = \frac{A^n}{n!}$. Die Norm ist $||a_n|| = \frac{||A^n||}{n!} \leq \frac{||A||^n}{n!} = b_n$.

$$\lim \frac{b_{n+1}}{b_n} = \lim \frac{||A||^{n+1}}{(n+1)!} \cdot \frac{n!}{||A||^n} = \lim \frac{||A||}{n+1} = 0$$

Satz 3.3.2. *Sei $A = (a_{ij})_{i,j=1}^n$ konstant. Dann gilt*

(i) Die allgemeine Lösung des homogenen DGL $y' = Ay$ ist

$$\varphi(x) = e^{Ax}c \quad c \in \mathbb{R}^n$$

Beweis:

(a) $\varphi(x) = e^{Ax}c$ ist Lösung von $y' = Ay : \varphi'(x) = Ae^{Ax}c = A\varphi(x)$

(b) *Jede Lösung $\varphi(x)$ von $y' = Ay$ lässt sich in der Form $\varphi(x) = e^{Ax}c$ darstellen. Sei φ Lösung von $y' = e^{Ax}y$. Zeigen, dass $c(x) := e^{-Ax}\varphi(x) = c$ konstant ist. $c'(x) = -Ae^{-Ax}\varphi(x) + e^{-Ax}\varphi'(x) = -Ae^{-Ax}\varphi(x) + e^{-Ax}A\varphi(x) = -Ae^{-Ax}\varphi(x) + Ae^{-Ax}\varphi(x) = 0 \Rightarrow c$ ist konstant. Daher folgt weiter $e^{Ax}c = e^{Ax}e^{-Ax}\varphi(x) = I\varphi(x) = \varphi(x)$*

(ii) *Die allgemeine Lösung der inhomogenen DGL $y' = Ay + b(x)$ ist*

$$\varphi(x) = e^{Ax}c + e^{Ax}\int_{x_0}^{x} e^{-At}b(t)dt$$

Beweis: Variation der Konstanten: Sei $\varphi(x)$ Lösung des inhomogenen DGL $y' = Ay + b(x)$. Man setzt $\varphi(x) = e^{Ax}c(x)$. Einsetzen in DGL: $\varphi'(x0 = Ae^{Ax}c(x) + e^{Ax}c'(x) = A\varphi(x)b(x) = Ae^{Ax}c(x) + b(x) \Rightarrow e^{Ax}c'(x) = b(x)$. Man multipliziert von links mit e^{-Ax} : $e^{-Ax}e^{Ax}c'(x) = e^{-Ax}b(x) \Rightarrow c'(x) = e^{-Ax}b(x)$.
Integration: $c(x) - \underbrace{c(x_0)}_{=c} = \int_{x_0}^{x} e^{-At}b(t)dt \Rightarrow c(x) = c + \int_{x_0}^{x} e^{-At}b(t)dt \Rightarrow \varphi(x) = e^{Ax}c(x) = e^{Ax}(c + \int_{x_0}^{x} e^{-At}b(t)dt)$
$e^{Ax}c$ ist die allgemeine Lösung der homogenen DGL $y' = Ay$ und $e^{Ax}\int_{x_0}^{x} e^{-At}b(t)dt$ ist die spezielle Lösung der inhomogenen DGL $y' = Ay + b(x)$.

$\square$

Bemerkung: $\Phi(x) = e^{Ax}, \Phi : \mathbb{R} \to \mathcal{L}(\mathbb{K}^n, \mathbb{K}^n), \Phi'(x) = Ae^{Ax} = e^{Ax}A$
Man benutzt die Reihe

$$\Phi(x) = e^{Ax} = I + Ax + \frac{A^2x^2}{2!} + \frac{A^3x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{A^kx^k}{k!}$$

$$\begin{aligned}
\frac{\Phi(x+h) - \Phi(x)}{h} &= \frac{e^{A(x+h)} - e^{Ax}}{h} \\
&= \frac{e^{Ax+Ah} - e^{Ax}}{h} \\
&= \frac{e^{Ax}e^{Ah} - e^{Ax}}{h} \\
&= e^{Ax}\frac{e^{Ah} - I}{h} \\
&= e^{Ax}\frac{1}{h}\sum_{k=1}^{\infty} \frac{A^kx^k}{k!} \\
&= e^{Ax}A\sum_{k=1}^{\infty} \frac{A^{k-1}h^{k-1}}{k!} \\
&= e^{Ax}A(I + \frac{Ah}{2!} + \frac{A^2h^2}{3!} + \cdots) \xrightarrow{h\to 0} e^{Ax}A \\
&= Ae^{Ax}
\end{aligned}$$

Beispiel:

$$\begin{aligned} y_1' &= y_1 + y_2 + \dots + y_n \\ y_2' &= y_1 + y_2 + \dots + y_n \\ &\vdots \\ y_n' &= y_1 + y_2 + \dots + y_n \end{aligned} \qquad \text{Gesucht ist } y_i = \varphi_i x$$

Matrixschreibweise:

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \qquad \varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_n(x) \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

Die allgemeine Lösung ist $\varphi(x) = e^{Ax}c, c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$

Berechnung von e^{Ax}: $e^{Ax} = \sum_{k=0}^{\infty} \frac{A^k x^k}{k!}$. Seien $e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, Ae_i = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = e_1 + e_2 + \dots + e_n$

$$\begin{aligned} A^2 e_i &= Ae_1 + Ae_2 + \dots + Ae_n = n(e_1 + \dots + e_n) = nAe_i \Rightarrow A^2 = nA \\ A^3 &= AA^2 = AnA = nA^2 = n^2 A \\ A^k &= n^{k-1} A \\ e^{Ax} &= I + \sum_{k=1}^{\infty} \frac{A^k x^k}{k!} = I + \sum_{k=1}^{\infty} \frac{n^{k-1} x^k}{k!} A \\ &= I + \frac{A}{n} \sum_{k=1}^{\infty} \frac{(nx)^k}{k!} = I + \frac{A}{n}(e^{nx} - 1) \\ &\boxed{e^{Ax} = I + \frac{e^{nx} - 1}{n} A} \end{aligned}$$

Allgemeine Lösung von $y' = Ay$ ist $\varphi(x) = e^{Ax}c = (I + \frac{e^{nx}-1}{n}A)c$
Probe: $\varphi'(x) = e^{nx}Ac, A\varphi(x) = A(I + \frac{E^{nx}-1}{n}A)c = (A + \frac{e^{nx}-1}{n}A^2)c = (A + \frac{e^{nx}-1}{n}nA)c = (A + (e^{nx} - 1)A)c = e^{nx}Ac \Rightarrow \varphi'(x) = A\varphi(x) \Rightarrow \varphi(x) = (I + \frac{e^{nx}-1}{n}A)c$ ist die allgemeine Lösung von $y' = Ay$.

Man berechne die allgemeine Lösung des inhomogenen DGL $y' = Ay + b(x), b(x) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = e_1$. In Komponenten lautet die Gleichung:

$$\begin{aligned} y_1' &= & y_1 + \cdots + y_n + 1 \\ \vdots &= & \vdots \\ y_n' &= & y_1 + \cdots + y_n \end{aligned}$$

Ansatz (Variation der Konstanten): $\varphi(x) = e^{Ax}c(x)$
$\varphi'(x) = Ae^{Ax}c(x) + e^{Ax}c'(x) = A\varphi(x) + b(x) = Ae^{Ax}c(x) + b(x), e^{Ax}c'(x) = b(x), c'(x) = e^{-Ax}b(x) = e^{A(-x)}b(x) = (I + \frac{e^{-nx}-1}{n}A)e_1$
$c(x) = c + \left(Ix - \frac{1}{n}\left(\frac{e^{-nx}}{n} + x\right)A\right)e_1$
Damit ergibt sich die DGL wie folgt:
$\varphi(x) = e^{Ax}c(x) = e^{Ax}\left(c + \left(Ix - \frac{1}{n}\left(\frac{e^{-nx}}{n} + x\right)A\right)e_1\right) = e^{Ax}c + e^{Ax}(Ix - \frac{1}{n}(\frac{e^{-nx}}{n} + x)A)e_1 = (I + \frac{e^{-nx}-1}{n}A)c + (I + \frac{e^{nx}-1}{n}A)(Ix - \frac{1}{n}(\frac{e^{-nx}+x}{n} + x)A)e_1 =$

$$\boxed{\left(I + \frac{e^{nx}-1}{n}A\right)c + \left(Ix - \left(\frac{x}{n} + \frac{1}{n^2}\right)A\right)e_1}$$

Spezialfall $n = 2$: $y = (y_1, y_2)^T, \varphi = (\varphi_1, \varphi_2)^T, A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, b(x) = e_1 = (1, 0)^T$
$\varphi(x) = (I + \frac{e^{2x}-1}{2}A)c + Ix - (\frac{x}{2} + \frac{1}{4})e_1, c = (c_1, c_2)^T$
$\varphi_1(x) = c_1 + \frac{e^{2x}-1}{2}(c_1 + c_2) + (x - (\frac{x}{2} + \frac{1}{4})) = \frac{e^{2x}+1}{2}c_1 + \frac{e^{2x}-1}{2}c_2 + \frac{x}{2} - \frac{1}{4}$
$\varphi_2(x) = c_2 + \frac{e^{2x}-1}{2}(c_1 + c_2) - \frac{x}{2} - \frac{1}{4} = \frac{e^{2x}-1}{2}c_1 + \frac{e^{2x}+1}{2}c_2 - \frac{x}{2} - \frac{1}{4}$
Probe: $\varphi_1'(x) = e^{2x}c_1 + e^{2x}c_2 + \frac{1}{2}, \varphi_1(x) + \varphi_2(x) + 1 = e^{2x}c_1 + e^{2x}c_2 - \frac{1}{2} + 1 \Rightarrow \varphi_1'(x) = \varphi_1(x) + \varphi_2(x) + 1$
$\varphi_2'(x) = e^{2x}c_1 + e^{2x}c_2 - \frac{1}{2}, \varphi_1(x) + \varphi_2(x) = e^{2x}c_1 + e^{2x}c_2 - \frac{1}{2} \Rightarrow \varphi_2'(x) = \varphi_1(x) + \varphi_2(x)$

Definition 3.3.3. *Sei $A \in \mathfrak{M}_n(\mathbb{K})$. Eine matrixwertige Funktion $\Phi = (\varphi_1, \ldots, \varphi_n) : \mathbb{R} \to \mathfrak{M}_n(\mathbb{K})$ heißt ein* Fundamentalsystem *von Lösungen von $y' = Ay$, falls $\Phi(x)$ für alle x invertierbar ist und die Matrixdifferentialgleichung $Y' = AY$ erfüllt, d.h. $\Phi'(x) = A\Phi(x)$. Es gilt:*

(1) Jedes Fundamentalsystem Φ hat die Gestalt $\varphi(x) = e^{Ax}S$, wobei $S \in \mathfrak{M}_n(\mathbb{K})$ invertierbar.
Beweis: $\Phi(x) = e^{Ax}S, \Phi'(x) = Ae^{Ax}S = A\Phi(x)$
Zeigen $S(x) = e^{-Ax}\Phi(x)$, wobei die invertierbare Matrix $\Phi(x)$ die DGL $Y' = AY$ erfüllt. $S'(x) = -Ae^{-Ax}\Phi(x) + e^{-Ax}\Phi'(x) = -Ae^{-Ax}\Phi(x) + e^{-Ax}A\Phi(x) = 0 \Rightarrow S'(x) = S =$konstant. $e^{-Ax}\Phi(x) = S \Rightarrow \Phi(x) = e^{Ax}S$. Da Φ, e^{Ax} invertierbar, folgt S invertierbar.

(2) *Die allgemeine Lösung von $y' = Ay$ ist $\varphi(x) = \Phi(x)c, c \in \mathbb{K}^n$*
Beweis: $\varphi(x) = \Phi(x)c = e^{Ax}Sc$ ist Lösung von $y' = Ay$. Zu zeigen ist $\varphi(x)$ Lösung von $y' = Ay$, so exitiert $c \in \mathbb{K}^n : \varphi(x) = e^{Ax}Sc$. Setzen $c(x) = S^{-1}e^{-Ax}\varphi(x)$. Dann gilt $c'(x) = S^{-1}Ae^{-Ax}\varphi + S^{-1}e^{-Ax}\varphi(x) = S^{-1}Ae^{-Ax}\varphi(x) - S^{-1}e^{-Ax}A\varphi(x) = 0 \Rightarrow$ $c(x) = x =$konstant. $c = S^{-1}e^{-Ax}\varphi(x) \Rightarrow \varphi(x) = e^{Ax}Sc$ □

Berechnung von $\Phi(x) = e^{Ax}S$

$$\begin{aligned}\Phi(x) &= (\varphi_1(x), \dots, \varphi_n(x))\\ \varphi_i(x) &= e^{Ax}Se_i\\ e_i &= (0, \dots, 0, 1, 0, \dots, 0)^T\end{aligned}$$

Es gilt:

Satz 3.3.4. *Die Matrix $\mathfrak{M}_n(\mathbb{C})$ besitzt die Eigenwerte $\mu_1, \dots, \mu_p, \lambda_1, \dots, \lambda_q$ und das System $\{v_{10}, v_{20}, \dots, v_{p0}, v_{11}, \dots, v_{r_1 1}, \dots, v_{12}, \dots, v_{r_2 2}, \dots, v_{1q}, \dots, v_{r_q q}\}, p + r_1 + \dots + r_q = n$ von Eigenvektoren und Hauptvektoren (Wurzelvektoren), d.h.*

$$\begin{aligned}(A - \mu_1 I)v_{10} &= 0, \dots, (A - \mu_p I)v_{p0} = 0\\ (A - \lambda_1 I)v_{11} &= 0, (A - \lambda_1 I)v_{21} = v_{11}, \dots, (A - \lambda_1 I)v_{r_1 1} = v_{(r_1 - 1)1}\\ &\vdots\\ (A - \lambda_q I)v_{1q} &= 0, (A - \lambda_q I)v_{2q} = v_{1q}, \dots, (A - \lambda_q I)v_{r_q q} = v_{(r_q - 1)q}\end{aligned}$$

Dann wird mit $\varphi_{10}(x) = e^{\mu_1 x}v_{10}, \dots, \varphi_{p0}(x) = e^{\mu_r x}v_{p0}$

$$\varphi_{ei}(x) = e^{\lambda_i x}\left(v_{ei} + xv_{(e-1)i} + \frac{x^2}{2!}v_{(e-2)i} + \dots + \frac{x^{l_i - 1}}{(l_i - 1)!}v_{1i}\right) \quad l = 1, \dots, r_1, i = 1, .., q$$

ein Fundamentalsystem von Lösungen von $y' = Ay$ gegeben, d.h. die Matrix $\Phi(x) = (\varphi_{10}(x), \varphi_{20}(x), \dots, \varphi_{p0}(x), \varphi_{11}(x), \dots, \varphi_{r_1 1}(x), \dots, \varphi_{1q}(x), \dots, \varphi_{r_q q}(x))$ erfüllt die DGL $Y' = AY$ und ist für jedes x invertierbar. Die allgemeine Lösung von $y' = Ay$ ist $\varphi(x) = \Phi(x)c, c \in \mathbb{K}^n$ oder $\varphi(x) = c_{10}\varphi_{10}(x) + \dots + c_{p0}\varphi_{p0}(x) + c_{11}\varphi_{11}(x) + \dots + c_{r_1 1}\varphi_{r_1 1}(x) + \dots + c_{1q}\varphi_{1q}(x) + \dots + c_{r_q q}\varphi_{r_q q}(x)$
Beweis:

(1) *Betrachten für J eine Diagonalmatrix $J = D = \begin{pmatrix}\mu_1 & & 0\\ & \ddots & \\ 0 & & \mu_n\end{pmatrix}$. Dann gilt: Die Spalten von S sind dann Eigenvektoren von A. Denn für e_1 hat man $ASe_i = SJe_i = S\mu_i e_i = \mu_i Se_i, Av_i = \mu_i v_i, v_i := Se_i$.*
$\Phi(x)e_i = e^{Ax}Se_i = Se^{Jx}e_i = Se^{\mu_i x}e_i = e^{\mu_i x}Se_i = e^{\mu_i x}v_i$
$\Phi(x) = (e^{\mu_1 x}v_1, \dots, e^{\mu_n x}v_n)$

$$\varphi(x) = \Phi(x)(c_1 e_1 + \dots + c_n e_n) = c_1\Phi(x)e_1 + \dots + c_n\Phi(x)e_n = c_1 e^{\mu_1 x}v_1 + \dots + c_n e^{\mu_n x}v_n$$

ist allgemeine Lösung von $y' = Ay$.

(2) *Jede Matrix* $A \in \mathfrak{M}_n(\mathbb{K})$ *kann nach einem Satz der linearen Algebra in eine Jordanmatrix* $J = S^{-1}AS$ *transformiert werden. Es ist* $J = \begin{pmatrix} J_0 & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_q \end{pmatrix}$ *blockdiagonal mit einem diagonalen Block* $J_0 = \begin{pmatrix} \mu_1 & & 0 \\ & \ddots & \\ 0 & & \mu_p \end{pmatrix}$ *und Jordanblöcken* $J_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda_i & 1 & \dots & 0 \\ \vdots & & & & & \\ 0 & \dots & & & \lambda_1 & 1 \\ 0 & \dots & & & 0 & \lambda_i \end{pmatrix}$ *der Größe* $r_i \geq 2$. *Die Zahlen* $\mu_1, \dots, \mu_p, \lambda_1, \dots, \lambda_q$ *sind dabei die nicht notwendigerweise verschiedenen Eigenwerte von* A. *Ferner hat man* $p + r_1 + \dots + r_q = n$ *und* $e^{Jx} = \begin{pmatrix} e^{J_0 x} & & 0 \\ & \ddots & \\ 0 & & e^{J_q x} \end{pmatrix}$

Wie ist die Lösung von $e^{J_i x}$*? Man setzt* $J_i = J = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & \dots & & & & \lambda \end{pmatrix}$, $J = \lambda I + N, N = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & & & & 1 \\ 0 & \dots & & & 1 \end{pmatrix}$

$e^{Jx} = e^{\lambda Ix + Nx} = e^{\lambda Ix} + e^{Nx} = e^{\lambda x} e^{Nx}$

$\varphi_l(x) = e^{Ax} Se_l = Se^{Ix} e_l = Se^{\lambda x}(e_l + e_{l-1} + \dots + \frac{x^{l-1}}{(l-1)!} e_1) = e^{\lambda x}(Se_l + xSe_{l-1} + \dots + \frac{x^{l-1}}{(l-1)!} Se_1)$

$v_i = Se_i : \varphi_l(x) = e^{\lambda x}(v_l + xv_{l-1} + \dots + \frac{x^{l-1}}{(l-1)!} v_1)$ *Es gilt*

$$\begin{aligned} Ne_1 = 0, Ne_2 = e_1, Ne_3 = e_2, \dots \quad & , Ne_l = e_{l-1}, 2 \leq l \leq r \\ N^2 e_1 = 0, N2e_2 = Ne_1 = 0, N2e_3 = Ne_2 = e_1, \dots \quad & , N^2 e_l = Ne_{l-1} = e_{l-2} \\ N^k e_1 = N^k e_2 = \dots = N^k e_k = 0 \quad & N^k e_l = e_{l-k} \\ N^{r-1} e_1 = \dots = N^{r-1} e_{r-1} = 0 \quad & N^{r-1} e_r = e_1, N^r = 0 \end{aligned}$$

$e^{Nx} e_l = \sum_{k=0}^{r-1} \frac{x^k N^k e_l}{k!} = Ie_l + xNe_l + \frac{x^2}{2!} N^2 e_l + \dots + \frac{x^{r-1}}{(r-1)!} N^{r-1} e_l = e_l + xe_{l-1} + \frac{x^2}{2!} e_{l-2} + \dots + \frac{x^{l-1}}{(l-1)!} e_1$ *Dabei ist* λ *der Eigenwert von* A *und die Eigenvektoren und Wurzelvektoren erfüllen die Gleichungen* $(A - \lambda I)v_1 = 0, (A - \lambda I)v_2 = v_1, \dots, (A - \lambda I)v_l = v_{l-1}, \dots, (A - \lambda I)v_r = v_{r-1}$. *Denn* $AS = SJ$ *impliziert:* $ASe_1 = SJe_1 =$

$S\lambda e_1 = \lambda S e_1 \Leftrightarrow (A-\lambda I)v_1 = 0$ *und* $ASe_l = SJe_l = S(\lambda e_l + e_{l-1}) = \lambda S e_l + S e_{l-1} \Leftrightarrow (A-\lambda I)v_l = v_{l-1}, 2 \leq l \leq r$

□

Bemerkung: Sei $A \in \mathfrak{M}_n(\mathbb{K})$ und sei $S \in \mathfrak{M}_n(\mathbb{K})$ invertierbar. Dann gilt für $J := S^{-1}AS$ die Formel $e^{Ax}S = Se^{Jx}$. Denn $J^2 = (S^{-1}AS)^2 = S^{-1}ASS^{-1}AS = S^{-1}A^2S$ und somit ist $e^{Jx} = \sum_{k=0}^{\infty} \frac{J^k x^k}{k!} = \sum_{k=0}^{\infty} \frac{S^{-1}ASx^k}{k!} = S^{-1}(\sum_{k=0}^{\infty} \frac{A^k x^k}{k!})S = S^{-1}e^{Ax}S \Leftrightarrow e^{Ax}S = Se^{Jx}$.

Beispiel:

$$\begin{aligned}
y_1' &= 3y_1 \\
y_2' &= y_2 - y_3 \\
y_3' &= y_2 + 3y_3 \\
y &= \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, y' = Ay \\
A &= \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{pmatrix} \\
y &= \varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \varphi_3(x) \end{pmatrix}
\end{aligned}$$

1. Schritt Bestimmung der Eigenwerte von A

$\det(\lambda I - A) = \det \begin{pmatrix} \lambda-3 & 0 & 0 \\ 0 & \lambda-1 & 1 \\ 0 & -1 & \lambda-3 \end{pmatrix} = (\lambda-3)\det\begin{pmatrix} \lambda-1 & 1 \\ -1 & \lambda-3 \end{pmatrix} =$
$(\lambda-3)((\lambda-1)(\lambda-3)+1) = (\lambda-3)(\lambda^2-4\lambda+4) = (\lambda-3)(\lambda-2)^2$

λ_1 ist drei und λ_2 ist ein zweifacher Eigenwert.

2. Schritt Bestimmung von Eigenvektoren und Wurzelvektoren

Zu $\lambda = 3$ existiert nur ein Eigenvektor:

$$\begin{aligned}
\psi_1(x) &= e^{\lambda_1 x}w_1 = e^{3x}w_1, (A-\lambda_1 I)w_1 = (A-3I)w_1 = 0 \\
\psi_2(x) &= e^{\lambda_2 x}w_2 = e^{2x}w_2, (A-\lambda_2 I)w_2 = (A-2I)w_2 = 0 \\
\psi_3(x) &= e^{\lambda_2 x}(w_3 + xw_2) = e^{2x}(w_3 + xw_2) \\
& (A-\lambda_2 I)w_3 = (A-2I)w_3 = w_2 \\
(A-3I)w_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} w_{11} \\ w_{21} \\ w_{31} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}$$

Daher folgt: w_{11} ist beliebig, $-2w_{21} - w_{31} = 0$ und $w_{21} = 0$. Setzen $w_{11} = 1$ und errechnen $w_{21} = w_{31} = 0$. Damit ist $w_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ein Eigenvektor zum Eigenwert $\lambda_1 = 3$.

$$\boxed{\psi_1(x) = e^{3x} w_1 = e^{3x} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} e^{3x} \\ 0 \\ 0 \end{pmatrix}}$$

ist Lösung von $y' = Ay$.

$$(A - 2I)w_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} w_{12} \\ w_{22} \\ w_{32} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Aus der Gleichung $w_{12} = 0, -w_{22} - w_{32} = 0 = w_{22} + w_{32}$ folgt, dass $w_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ ein Eigenvektor ist.

$$\boxed{\psi_2(x) = e^{2x} w_2 = e^{2x} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ e^{2x} \\ -e^{2x} \end{pmatrix}}$$

$$(A - 2I)w_3 = w_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} w_{13} \\ w_{23} \\ w_{33} \end{pmatrix} = w_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$w_{13} = 0, -w_{23} - w_{33} = 1, w_{23} + w_{33} = -1 \Rightarrow w_{13} = 0, w_{23} = -1, w_{33} = 0$

Somit ist $w_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$ ein Eigenvektor.

$$\boxed{\begin{aligned} \psi_3(x) = e^{2x}(w_3 + x w_2) &= e^{2x} \left(\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} + x \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right) \\ &= e^{2x} \begin{pmatrix} 0 \\ x - 1 \\ -x \end{pmatrix} = \begin{pmatrix} 0 \\ (x-1)e^{2x} \\ -xe^{2x} \end{pmatrix} \end{aligned}}$$

Fundamentalmatrix:

$$\Phi(x) = (\psi_1(x), \psi_2(x), \psi_3(x) = \begin{pmatrix} e^{3x} & 0 & 0 \\ 0 & e^{2x} & (x-1)e^{2x} \\ 0 & -e^{2x} & -xe^{2x} \end{pmatrix}$$

Probe: Zeigen $Y = \Phi(x)$ erfüllt Matrixdifferentialgleichung $Y' = AY$:

$$\Phi'(x) = \begin{pmatrix} 3e^{3x} & 0 & 0 \\ 0 & 2e^{2x} & (2x-1)e^{2x} \\ 0 & -2e^{2x} & -(2x+1)e^{2x} \end{pmatrix},$$

$$A\Phi(x) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 3e^{3x} & 0 & 0 \\ 0 & 2e^{2x} & (2x-1)e^{2x} \\ 0 & -2e^{2x} & -(2x+1)e^{2x} \end{pmatrix} \Rightarrow$$

$\Phi'(x) = A\Phi(x)$

Allgemeine Lösung von $y' = Ay$: $\varphi(x) = \Phi(x)c, c \in \mathbb{R}^3$

$\varphi(x) = \Phi(x) \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = c_1\varphi(x)e_1 + c_2\varphi(x)e_2 + c_3\varphi(x)e_3 = c_1\psi(x)e_1 + c_2\psi(x)e_2 + c_3\psi(x)e_3$

In Komponenten:

$$\begin{aligned} \varphi(x) &= \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \varphi_3(x) \end{pmatrix} \\ &= c_1 \begin{pmatrix} e^{3x} \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ e^{2x} \\ -e^{2x} \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ (x-1)e^{2x} \\ -xe^{2x} \end{pmatrix} \\ &= \begin{pmatrix} c_1e^{3x} \\ c_2e^{2x} + c_3(x-1)e^{2x} \\ -c_2e^{2x} - c_3xe^{2x} \end{pmatrix} \end{aligned}$$

$y_1 = \varphi_1(x) = c_1e^{3x}, y_2 = \varphi_2(x) = c_2e^{2x}c_3(x-1)e^{2x}, y_3 = \varphi_3(x) = -c_2e^{2x} - c_3xe^{2x}$

3.4 Der Existenz- und Eindeutigkeitssatz von Picard-Lindelöf

Man betrachtet das DGL $y' = f(x,y), y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, f = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$. In Komponenten:

$$\begin{aligned} y_1' &= f_1(x, y_1, \dots, y_n) \\ &\vdots \\ y_n' &= f_n(x, y_1, \dots, y_n \end{aligned}$$

3.4.1 Lipschitzbedingung

Definition 3.4.1. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n, f : G \to \mathbb{R}^n$ stetig.*

(a) f genüge in G einer Lipschitzbedingung

$$\exists L \geq 0 \, \forall (x,y),(x,\tilde{y}) \in G : ||f(x,y) - f(x,\tilde{y})|| \leq L||y-\tilde{y}||$$

(Genauer: $f = f(x,y)$ genüge in G bzgl. y einer Lipschitzbedingung)

(b) f genüge in G lokal *einer Lipschitzbedingung*

$$\forall (a,b) \in G \, \exists \textit{Umgebung } U \textit{ von } (a,b) \textit{und } f \textit{ genügt in } G \cap U \textit{ einer Lipschitzbedingung}$$

Hinreichendes Kriterium für Lipschitzbedingung Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n, f = f(x,y), y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ *stetig* partiell differenzierbar. Dann genügt f in G lokal einer Lipschitzbedingung.

Beweis: $(a,b) \in G : \exists r > 0$, so dass $U := \{(x,y) \in \mathbb{R} \times \mathbb{R}^n : |x - a \leq r, ||y-b|| \leq r\} \subset G$. U ist in $\mathbb{R} \times \mathbb{R}^n$ abgeschlossen und beschränkt und somit also kompakt. Somit ist wegen Stetigkeit von $\frac{\partial f}{\partial y} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix}$ auf U: $L := \sup\{\left|\left|\frac{\partial f}{\partial y}(x,y)\right|\right| : (x,y) \in U\} < \infty$.

Nach dem Mitelwertsatz gilt:

$$f(x,y) - f(x,\tilde{y}) = \left(\int_0^1 \frac{\partial f}{\partial y}(x, y + t(\tilde{y} - y))dt \right)(\tilde{y} - y)$$

Daher folgt:

$$\begin{aligned}
||f(x,y) - f(x,\tilde{y})||_2 = \quad & ||\int_0^1 \frac{\partial f}{\partial y}(x, y + t(\tilde{y} - y))dt)(\tilde{y} - y) \\
\leq \quad & \int_0^1 ||\frac{\partial f}{\partial y}(x, y + t(\tilde{y} - y)))(\tilde{y} - y)||_2 dt \\
\leq \quad & \int_0^1 \left|\left|\frac{\partial f}{\partial y}(x, y + t(\tilde{y} - y))\right|\right| ||\tilde{y} - y||_2 dt \\
\leq \quad & \int_0^1 \underbrace{\left|\left|\frac{\partial f}{\partial y}(x, y + t(\tilde{y} - y))\right|\right|}_{\leq L \, \forall (x,z) \in U} dt ||\tilde{y} - y||_2 \\
\leq \quad & L \int_0^1 1dt ||\tilde{y} - y||_2 = L||\tilde{y} - y||_2
\end{aligned}$$

Da auf endlichdimensionalen Räumen die Normen äquivalent sind, erhalten wir unter der Ausgangsnorm $||\cdot||$: $||f(x,y) - f(x,\tilde{y}|| \leq \tilde{L}||\tilde{y} - y||$. □

3.4.2 Satz von Picard-Lindelöf

Satz 3.4.2. *Sei $G \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : G \to \mathbb{R}^n$ stetig, die lokal einer Lipschitzbedingung genügt. Dann gilt, dass für alle $(a, c) \in G$ ein $\varepsilon > 0$ existiert und genau eine Lösung $\varphi : [a - \varepsilon, a + \varepsilon] \to \mathbb{R}^n$ vom Anfangswertproblem $y' = f(x, y)$ mit $\varphi(a) = c$.*
Beweis:

1. Schritt *Raum der stetigen Funktionen*
Sei I ein abgeschlossenes und beschränktes Intervall (=kompakt). $\mathcal{C}(I, \mathbb{R}^n) := \{\varphi : I \to \mathbb{R}^n : \varphi \text{ stetig}\}, ||\varphi|| := \sup_{x \in I} ||\varphi(x)||_2$. Der Raum $[\mathcal{C}(I, \mathbb{R}^n), ||\cdot||]$ ist ein vollständiger normierter Raum. Sei $I_\varepsilon := [a - \varepsilon, a + \varepsilon], \varepsilon > 0$. Wählen $\varepsilon, r > 0$ so, dass $U = \{(x, y) \in I_\varepsilon \times \mathbb{R}^n : ||y - c||_2 \leq r\} \subset G$ und f auf U einer Lipschitzbedingung genügt, d.h. $\exists L \geq 0 \forall (x, y), (x, \tilde{y}) \in U : ||f(x, y) - f(x, \tilde{y})||_2 \leq ||y - \tilde{y}||_2$. U ist kompakt und f stetig auf U. Man setzt $B_r^\varepsilon(c) := \{\varphi \in \mathcal{C}(I_\varepsilon, \mathbb{R}^n) : ||\varphi - c|| \leq r\}$ (ist i.a. nicht kompakt). B_r^ε ist als abgeschlossene Kugel $\mathcal{C}(I_\varepsilon, \mathbb{R}^n)$ ein vollständiger metrischer Raum mit der Metrik $(d(\varphi, \psi) = ||\varphi - \psi||)$.

2. Schritt *Kontrahierende Abbildung*
Wir definieren die Abbildung

$$T : B_r^\varepsilon(c) \to \mathcal{C}(I_\varepsilon, \mathbb{R}^n)$$

durch den Ansatz $(T\varphi)(x) := c + \int_a^x f(t, \varphi(t))dt$.

a) Es gilt die Abschätzung

$$\begin{aligned} ||(T\varphi)(x) - c||_2 = & \left|\left|\int_a^x f(t, \varphi(t))dt\right|\right|_2 \\ \leq & \left|\int_a^x ||f(t, \varphi(t))||_2 dt\right| \\ \leq & |x - a|M \leq \varepsilon M \quad \forall x \in I_\varepsilon \end{aligned}$$

Hieraus folgt:

$$\boxed{\forall \varphi \in B_r^\varepsilon(c) : ||T\varphi - c|| = \sup_{x \in I_\varepsilon} ||(T\varphi)(x) - c||_2 \leq \varepsilon M}$$

b)

$$\begin{aligned} ||(T\varphi)(x) - (T\tilde{\varphi})(x)||_2 = & \left|\left|\int_a^x (f(t, \varphi(t)) - f(t, \tilde{\varphi}(t)))dt\right|\right|_2 \\ \leq & \left|\int_a^x \underbrace{||f(t, \varphi(t)) - f(t, \tilde{\varphi}(t)))||_2}_{L||\varphi - \tilde{\varphi}||} dt\right| \\ \leq & L|x - a|||\varphi - \tilde{\varphi}|| \\ \leq & \varepsilon L||\varphi - \tilde{\varphi}|| \quad x \in I_\varepsilon \end{aligned}$$

$$\Rightarrow \boxed{\forall \varphi, \tilde{\varphi} \in B_r^\varepsilon(c) : ||T\varphi - T\tilde{\varphi}|| \leq \varepsilon L ||\varphi - \tilde{\varphi}||}$$

Man wählt $\varepsilon > 0$ so, dass $\varepsilon \leq \min\{\frac{r}{M}, \frac{1}{2L}\}$. Dann gilt:

a) $T(B_r^\varepsilon(c)) \subset B_r^\varepsilon(c)$

b) $\forall \varphi, \tilde{\varphi} \in B_r^\varepsilon(c) : ||T\varphi - T\tilde{\varphi}|| \leq \frac{1}{2}||\varphi - \tilde{\varphi}||$

3. Schritt Anwendung des Banachschen Fixpunktsatzes
Man setzt $X := B_r^\varepsilon(c)$ (vollständiger metrischer Raum) und $T : X \to X$ ist kontrahierend. Nach dem Banachschen Fixpunktsatzes existiert genau ein $\varphi \in X = B_r^\varepsilon(c) : T\varphi = \varphi$.

Der Hauptsatz der Differential- und Integralrechnung impliziert $\varphi'(x) = f(x, \varphi(x))$ und $\varphi(a) = c$. □

Beispiele für Anfangswertprobleme:

(a) $y' = \sqrt[3]{y^2}, y = \varphi(a) = c, f = f(x,y) = \sqrt[3]{y^2}, f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

1. Fall $a \in \mathbb{R}, c \neq 0 : \varphi(a) = c$
$\frac{\partial f}{\partial y}(x,y) = \frac{2}{3}y^{\frac{2}{3}-1} = \frac{2}{3}y^{-\frac{1}{3}} \Rightarrow f$ erfüllt lokal eine Lipschitzbedingung in $\mathbb{R} \times (\mathbb{R}\backslash\{0\})$. Nach dem Satz von Picard-Lindelöf folgt, dass das Anfangswertproblem eindeutig lösbar ist. Berechnung der Lösung:

$$\varphi'(x) = \sqrt[3]{(\varphi(x))^2} = \frac{\varphi'(x)}{\sqrt[3]{(\varphi(x))^2}} = 1$$

Integration:

$$\int_a^x \frac{\varphi'(t)}{\sqrt[3]{\varphi(t)^2}} dt = \int_a^x dt = x - a$$

$$\int_a^x \frac{\varphi'(x)dt}{\sqrt[3]{\varphi(t)^2}} \begin{bmatrix} u = \varphi(t) \\ du = \varphi'(t)dt \end{bmatrix} = \int_{c=\varphi(a)}^{\varphi(x)} \frac{du}{\sqrt[3]{u^2}} = \int_u^{\varphi(x)} u^{-\frac{2}{3}} du$$

$$= 3u^{\frac{1}{3}}|_c^{\varphi(x)} = 3\left((\varphi(x))^{\frac{1}{3}} - c^{\frac{1}{3}}\right) = x - a$$

$$\boxed{\varphi(x) = \frac{1}{27}(x + 3\sqrt[3]{c} - a)^3}$$

Bemerkung: Die Lösung ist sogar auf $I = \mathbb{R}$ definiert.

2. Fall $y = \varphi(a) = 0, c = 0$
$\varphi_0(x) = 0$ ist eine Lösung des Anfangswertproblems.
$\varphi_a(x) = \frac{1}{27}(x-a)^3$ ist ebenfalls eine Lösung des Anfangswertproblems. Ferner ist auch für $a_0 < a < a_1, \varphi(x) = \begin{cases} \varphi_{a_0}(x) = \frac{1}{27}(x-a_0)^3 & x \leq a_0 \\ 0 & a_0 < x < a_1 \\ \varphi_{a_1}(x) = \frac{1}{27}(x-a_1)^3 & x \geq a_1 \end{cases}$ eine Lösung des Anfangswertproblems. Damit ist das Anfangswertproblem für kein $a \in \mathbb{R}$ mit $\varphi(a) = 0$ eindeutig lösbar.

3.4.3 Das Picard-Lindelöfsche Iterationsverfahren anhand eines Beispiels

$y' = f(x, y), y = \varphi(x), y = \varphi(a) = c, \varphi'(x) = f(x, \varphi(x)) = \int_a^x \varphi'(t)dt = \int_a^x f(t, \varphi(t))dt$
Iterationsprozedur: $\varphi_{n+1} = (T\varphi_n)(x) = c + \int_a^x f(t, \varphi_n(t))dt$
Der Startwert ist $\varphi_0 = c$.
Beispiel: $y' = y, y = \varphi(0) = c$ Die Lösung ist $\varphi(x) = ce^x$. Mit Hilfe des Iterationsverfahrens wird ebenfalls die Lösung ermittelt: $f(x, y) = y, |f(x, y) - f(x, \tilde{y})| = |y - \tilde{y}|$
f genügt einer Lipschitzbedingung mit $L = 1$. Damit führt das Iterationsverfahren zu genau einer Lösung mit $\varphi(0) = c$.

$$\begin{aligned}
\varphi_{n+1} &= & c + \int_a^x f(t, \varphi_n(t))dt \\
\varphi_1(x) &= & c + \int_0^x f(t, \varphi_0(t))dt = c + \int_0^x \varphi_0(t)dt = c + \int_0^x cdt = c + cx \\
\varphi_2(x) &= & c + \int_0^x \varphi_1(t)dt = c + \int_0^x c(1+t)dt = c(1 + x + \frac{x^2}{2}) \\
&\vdots & \\
\varphi_k &= & c(1 + x + \frac{x^2}{2!} + \dots + \frac{x^k}{k!})
\end{aligned}$$

Induktion: $k \to k + 1$:
$\varphi_{k+1}(x) = c + \int_0^x \varphi_k(t)dt = c + c\int_0^x (1 + t + \frac{t^2}{2!} + \dots + \frac{t^k}{k!})dt = c + c(t + \frac{t^2}{2!} + \dots + \frac{t^{k+1}}{(k+1)!})|_0^x = c + c(x + \frac{x^2}{2} + \dots + \frac{x^{k+1}}{(k+1)!}) = c(1 + x + \frac{x^2}{2} + \dots + \frac{x^{k+1}}{(k+1)!})$
$\varphi(x) = \lim_{n\to\infty} \varphi_n(x) = c \lim \sum_{k=0}^n \frac{x^k}{k!} = c\sum_{k=0}^\infty \frac{x^k}{k!} = ce^x$

3.5 Lösungstheorie linearer Differentialgleichungssysteme 1. Ordnung

$$y' = A(x)y + b(x), A(x) = (a_{ij}(x))_{i,j=1}^n, b(x) = \begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix}$$

3.5.1 Existenz- und Eindeutigkeitssatz

Satz 3.5.1. *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \to \mathfrak{M}_n(\mathbb{K}), b : I \to \mathbb{K}^n$ stetige Abbildungen. Dann gilt:*

$$\forall x_n \in I\, \forall c \in \mathbb{K}^n\, \exists^1 \textit{Lösung}\ \varphi : I \to \mathbb{K}^n : y' = A(x)y + b(x), \varphi(x_0) = c$$

***Beweis**: Setzen $f(x, y) = A(x)y + b(x)$. Man zeigt, dass für jedes kompakte Intervall $J \subset I$ ein $L \geq 0$ existiert mit $\forall x \in J\, \forall y, \tilde{y} \in \mathbb{K}^n : ||f(x, y) - f(x, \tilde{y})|| \leq L||y - \tilde{y}||$*
$A(x) = (a_{ij}(x))$ ist stetig. Daher ist auch $a_{ij}(x)$ stetig. Aus der Tatsache, dass J kompakt

ist, folgt $M_J := \max_{i,j=1,\dots,n} \sup_{x\in J} |a_{ij}(x)| < \infty$. *Für* $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \tilde{y} = \begin{pmatrix} \tilde{y_1} \\ \vdots \\ \tilde{y_n} \end{pmatrix}$ *gilt:*

$$\begin{aligned} ||f(x,y) - f(x,\tilde{y})||_1 = \quad & ||A(x)y - A(x)\tilde{y}||_1 = \sum_{i=1}^n |\sum_{j=1}^n a_{ij}(x)(y_j - \tilde{y_j})| \\ \leq \quad & \sum_{i=1}^n \sum_{j=1}^n |a_{ij}(x)||y_j - \tilde{y_j}| \leq nM_J \sum_{i=1}^n |y_j - \tilde{y_j}| \\ = \quad & nM_j||y-\tilde{y}||_1 \end{aligned}$$

Da die $||\cdot||_1$ *äquivalent zur Ausgangsnorm* $||\cdot||$ *ist, erhält man mit einer weiteren Konstanten* $\alpha \geq 0$ *die Abschätzung* $||f(x,y) - f(x,\tilde{y})|| \leq \alpha n M_J ||y - \tilde{y}||$. *Nach Existenz und Eindeutigkeitssatz von Picard-Lindelöf existiert ein offenes Intervall* $\tilde{J} \subset J$ *und* $x_0 \in \tilde{J}$ *und das AWP* $y' = A(x)y + b(x), y = \varphi(x_0) = c$ *ist eindeutig lösbar. Ferner kann man zeigen, dass die Lösung des AWP eindeutig auf J und weiter auf das gesamte Intervall I fortgesetzt werden kann (siehe Forster Analysis II).* □

3.5.2 Lösungstheorie

Dazu wird $\mathcal{L}_h$ für homogene Lösungen eingeführt: $\mathcal{L}_h := \{\varphi : I \to \mathbb{K}^n : \varphi \text{Lösung von } y' = A(x)y\}$
$\mathcal{L}_{inh} := \{\varphi : I \to \mathbb{K}^n : \varphi \text{Lösung von } y' = A(x)y + b(x)\}$

Satz 3.5.2. *Es gelten folgende Aussagen:*

(i) $\varphi_1, \varphi_2 \in \mathcal{L}_h \Rightarrow c_1\varphi_1 + c_2\varphi_2 \mathcal{L}_h, c_1, c_2 \in \mathbb{K}$
Beweis: $\varphi_1, \varphi_2 \in \mathcal{L}_h$, *d.h.* $\varphi_i(x) = A(x)\varphi_i(x), i = 1,2 \Rightarrow (c_1\varphi_1(x) + c_2\varphi_2(x))' = c_1\varphi_1'(x) + c_2\varphi_2'(x) = c_1 A(x)\varphi_1(x) + c_2 A(x)\varphi_2(x) = A(x)(c_1\varphi_1(x) + c_2\varphi_2(x))$

(ii) $\varphi \in \mathcal{L}_h, \psi \in \mathcal{L}_{inh} \Rightarrow \varphi + \psi \in \mathcal{L}_{inh}$
Beweis: $\varphi \in \mathcal{L}_h, \psi \in \mathcal{L}_{inh} \Rightarrow \varphi'(x) = A(x)\varphi(x), \psi'(x) = A(x)\psi(x) + b(x) \Rightarrow (\varphi(x) + \psi(x))' = \varphi'(x) + \psi'(x) = A(x)\varphi(x) + A(x)\psi(x) + b(x) = A(x)(\varphi(x) + \psi(x) + b(x)) \Rightarrow \varphi + \psi \in \mathcal{L}_{inh}$

(iii) $\psi_1, \psi_2 \in \mathcal{L}_{inh} \Rightarrow \psi_1 - \psi_2 \in \mathcal{L}_h$
Beweis: $\psi_1, \psi_2 \in \mathcal{L}_{inh}, \psi_i'(x) = A(x)\psi_i(x) + b(x), i = 1,2 \Rightarrow (\psi_1(x) - \psi_2(x))' = \psi_1'(x) - \psi_2'(x) = A(x)\psi_1(x) + b(x) - (A(x)\psi_2(x) + b(x)) = A(x)(\psi_1(x) - \psi_2(x)) \Rightarrow \psi_1 - \psi_2 \in \mathcal{L}_h$ □

3.5.3 Lineare Abhängigkeit und Unabhängigkeit

Definition 3.5.3. *(i) k stetige Funktionen* $\varphi_1, \dots, \varphi_k \in \mathcal{C}(I, \mathbb{K}^n)$ *heissen* linear unabhängig, *g.d.w.:*

$$\forall \alpha_1, \dots, \alpha_k \in \mathbb{K}(\sum_{i=1}^k \alpha_i\varphi_i = 0 \Rightarrow \alpha_1 = \dots = \alpha_k = 0$$

(ii) k stetige Funktionen $\varphi_1, \dots, \varphi_k \in \mathcal{C}(I, \mathbb{K}^n)$ heissen linear abhängig, *g.d.w.:*

$$\exists \alpha_1, \dots, \alpha_k \in \mathbb{K} : \sum_{i=1}^{k} \alpha_i \varphi_i = 0 \wedge \sum_{i=1}^{k} |\alpha_i| > 0$$

Bemerkung: $\sum_{i=1}^{k} \alpha_i \varphi_i = 0 : \forall x \in I : \sum_{i=1}^{k} \alpha_i \varphi_i(x) = 0$

Satz 3.5.4. *(a) Für k Lösungen $\varphi_1, \dots, \varphi_k \in \mathcal{L}_h$ sind folgende Aussagen äquivalent:*

(i) $\varphi_1, \dots, \varphi_k$ sind linear unabhängig über $\mathbb{K}$.

(ii) $\exists x_0 \in I : \varphi_1(x_0), \dots, \varphi_k(x_0) \in \mathbb{K}^n$ sind linear unabhängige Vektoren über $\mathbb{K}$.

(iii) $\forall x \in I : \varphi_1(x), \dots, \varphi_k(x) \in \mathbb{K}^n$ sind linear unabhängige Vektoren über $\mathbb{K}$.

***Beweis**: Der Beweis von (iii) nach (ii) nach (i) ist trivial. Zu zeigen ist daher noch die Richtung von (i) nach (iii). Hier erfolgt der Beweis indirekt. Seien $\varphi_1, \dots, \varphi_k \in \mathcal{L}_h$. Annahme: $\exists x_0 \in I : \varphi_1(x_0), \dots, \varphi_k(x_0)$ sind linear abhängig, d.h. $\exists \alpha_1, \dots, \alpha_k \in \mathbb{K} : \sum_{i=1}^{k} |\alpha_i| > 0 \wedge \sum_{i=1}^{k} \alpha_i \varphi_i(x_0) = 0$. Es gilt $\varphi = \sum_{i=1}^{k} \alpha_i \varphi_i \in \mathcal{L}_h$ mit $\varphi(x_0) = \sum_{i=1}^{k} \alpha_i \varphi_i(x_0) = 0$. Wegen des Existenz- und Eindeutigkeitssatz von Picard-Lindelöf des AWP $y' = A(x)y, y = \varphi(x_0) = 0$ folgt $\varphi = 0$ ist die einzige Lösung des AWP, d.h. $\forall x \in I : \varphi(x) = \sum_{i=1}^{k} \alpha_i \varphi_i(x) = 0 \Rightarrow \varphi_1, \dots, \varphi_k$ sind linear abhängig. ↯zu (i)*

(b) $\dim \mathcal{L}_h = n$

***Beweis**: Man zeigt zuerst, dass die Dimension größer oder gleich n ist: Für die Einheitsvektoren $e_1, \dots, e_n \in \mathbb{K}^n$ existieren Lösungen $\varphi_1, \dots, \varphi_n \in \mathcal{L}_h$ mit $\varphi_i(x_0) = e_i$. Da $e_1, \dots, e_n$ linear unabhängige Vektoren in $\mathbb{K}$ sind, sind auch die Lösungen $\varphi_1, \dots, \varphi_n$ linear unabhängig nach dem Punkt (a)(ii) dieses Satzes. Somit gilt* $\dim \mathcal{L}_h \geq n$.

Zu zeigen ist nun, dass die Dimension kleiner gleich n ist: Annahme: $\psi_1, \dots, \psi_{n+1} \in \mathcal{L}_h$ seien linear unabhängige Lösungen. Dann sind die $n+1$ Vektoren aus dem Körper $\mathbb{K}$ $\psi_1(x_0), \dots, \psi_{n+1}(x_0)$ linear unabhängig. ↯zu $\dim \mathbb{K}^n = n$

□

Bemerkung: Sind $\varphi_1, \dots, \varphi_k$ keine Lösungen des homogenen DGL $y' = A(x)y$, d.h. $\varphi_1, \dots, \varphi_k \notin \mathcal{L}_k$, so ist Satz 3.5.4 i.a. falsch. Beispiel:

$$\varphi_1 = \varphi_1(x) = \begin{pmatrix} x \\ 0 \end{pmatrix} \quad \varphi_2 = \varphi_2(x) = \begin{pmatrix} x^2 \\ 0 \end{pmatrix} \quad \varphi_1, \varphi_2 \in \mathcal{C}\left((0,2), \mathbb{R}^2\right)$$

sind linear unabhängig. Denn $\alpha_1 \varphi_1(x) + \alpha_2 \varphi_2(x) = 0$ und $\alpha_1 \varphi_1 + \alpha_2 \varphi_2 = 0$ und es folgt:

$$\begin{aligned}
0 = \quad & \alpha_1 \varphi_1(\frac{1}{2}) + \alpha_2 \varphi_2(\frac{1}{2}) = \alpha_1 \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} \frac{1}{4} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\alpha_1}{2} + \frac{\alpha_2}{4} \\ 0 \end{pmatrix} \\
0 = \quad & \alpha_1 \varphi_1(1) + \alpha_2 \varphi_2(1) = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ 0 \end{pmatrix} \\
\Rightarrow \quad & \frac{\alpha_1}{2} + \frac{\alpha_2}{4} = 0 \wedge \alpha_1 + \alpha_2 = 0 \\
\Rightarrow \quad & \alpha_1 = \alpha_2 = 0 \\
\Rightarrow \quad & \varphi_1, \varphi_2 \in \mathcal{C}((0,2), \mathbb{R}^2)
\end{aligned}$$

sind linear unabhängig. Aber $\varphi_1(1) = \varphi_2(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, d.h. (a)(iii) ist nicht erfüllt.

Definition 3.5.5. *Ein Lösungsfundamentalsystem der Differentialgleichung*

$$y' = A(x)y \quad A(x) \in \mathfrak{M}_n(\mathbb{K})$$

sind n linear unabhängige Lösungen $\varphi_1, \dots, \varphi_n \in \mathcal{L}_h$. Schreibt man $\varphi_i(x) = \begin{pmatrix} \varphi_{1i}(x) \\ \vdots \\ \varphi_{ni}(x) \end{pmatrix}$, so heißt $\Phi = (\varphi_1, \dots, \varphi_n) = \begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1n} \\ \vdots & \vdots & & \vdots \\ \varphi_{n1} & \varphi_{n2} & \cdots & \varphi_{nn} \end{pmatrix}$ Fundamentalmatrix *von $y' = A(x)y$.*

Bemerkungen:

(a) Nach Satz 3.5.4 gilt: $\varphi_1, \dots, \varphi_n$ linear unabhängig, g.d.w. $\exists x_0 \in I : \det \Phi(x_0) \neq 0 \Leftrightarrow \forall x \in I : \det \Phi(x) \neq 0$

(b) Ist $\Phi = (\varphi_1, \dots, \varphi_n)$ ein Lösungsfundamentalsystem von $y' = A(x)y$, so lässt sich die allgemeine Lösung von $y' = A(x)y$ in der Form

$$\varphi(x) = \Phi(x)c, c \in \mathbb{K}^n, c = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \quad \text{oder } \varphi(x) = c_1\varphi_1(x) + \cdots + c_n\varphi_n(x)$$

schreiben. Man kann Φ selbst als Lösung der folgenden Matrix-DGL $Y' = A(x)Y$ auffassen.

3.5.4 Die inhomogene Differentialgleichung $y' = A(x)y + b(x)$

Satz 3.5.6. *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $A : I \to \mathfrak{M}_n(\mathbb{K}), b : I \to \mathbb{K}^n$ stetig. Dann gilt für $\psi_0 \in \mathcal{L}_{inh} : \mathcal{L}_{inh} = \psi_0 + \mathcal{L}_h := \{\psi_0 + \varphi : \varphi \in \mathcal{L}_j\}$, d.h. $\psi(x) = \psi_0(x) + \varphi(x) + c, c \in \mathbb{K}^n$, wobei Φ eine Lösungsfundamentalmatrix ist.*
Allgemeine Lösung des inhomogenen DGL ist die allgemeine Lösung der homogenen DGL plus einer speziellen Lösung der inhomogenen DGL.
Beweis:

1. Schritt *Zu zeigen ist $\mathcal{L}_{inh} \subset \psi_0 + \mathcal{L}_h$. Sei $\psi \in \mathcal{L}_{inh}$. Daraus folgt nach Satz 3.5.2 $\psi - \psi_0 \in \mathcal{L}_h \Rightarrow \exists \varphi \in \mathcal{L}_h : \psi - \psi_0 = \varphi$.*

2. Schritt *Zu zeigen ist $\psi_0 + \mathcal{L}_h \subset \mathcal{L}_{inh}$. Sei $\psi \in \psi_0 + \mathcal{L}_h$, d.h. $\exists \varphi \in \mathcal{L}_h : \psi = \psi_0 + \varphi \Rightarrow \psi \in \mathcal{L}_{inh} \Rightarrow \psi_0 + \mathcal{L}_h \in \mathcal{L}_{inh}$*

□

Lösen der inhomogenen DGL $y' = A(x)y + b(x)$ mit der Methode der Variation der Konstanten

Sei $\Phi = (\varphi_1, \ldots, \varphi_n)$ ein Lösungsfundamentalsystem von $y' = A(x)y$. Dann gilt für die allgemeine Lösung von $y' = A(x)y : \varphi(x) = \Phi(x), c \in \mathbb{K}^n$.
Variation der Konstanten: Der Ansatz ist $\psi(x) = \Phi(x)c(x)$. Man muss nun differenzieren: $\psi'(x) = \Phi'(x)c(x) + \Phi(x)c'(x) = A(x)\Phi(x)c(x) + \Phi(x)c'(x) = A(x)\psi(x) + \Phi(x)c'(x) = A(x)\psi(x) + b(x) \Rightarrow \Phi(x)c'(x) = b(x)$. Da ein $\Phi^{-1}(x)$ für alle $x \in I$ existiert, folgt $c'(x) = \Phi^{-1}(x)b(x) \Rightarrow c(x) - c(x_0) = \int_{x_0}^{x} \Phi^{-1}(t)b(t)dt \Rightarrow c(x) = c + \int_{x_0}^{x} \Phi^{-1}(t)b(t)dt$

$$\Rightarrow \psi(x) = \Phi(x)c(x) = \Phi(x)c + \int_{x_0}^{x} \Phi^{-1}(t)b(t)dt$$

Allgemeine Lösung der inhomogenen DGL:

$$\boxed{\psi(x) = \Phi(x)c + \Phi(x)\int_{x_0}^{x} \Phi^{-1}(t)b(t0)dt}$$

Beispiel:

$$\begin{aligned} y_1' &= \frac{1}{2x}y_1 - \frac{1}{2x^2} + 1, x > 0 \\ y_2' &= y_1 + x \end{aligned}$$

Gesucht ist die allgemeine Lösung des inhomogenen DGL mit $y = (y_1, y_2)^T$ und $A(x) = \begin{pmatrix} \frac{1}{2x} & -\frac{1}{2x^2} \\ 1 & 0 \end{pmatrix}, b(x) = (1, x)^T$.
Ansatz:

$$\begin{aligned} y_2 &= y_2(x)x^{\alpha}, \alpha \in \mathbb{R} \\ y_1(x) &= y_2'(x) = \alpha x^{\alpha-1} \end{aligned}$$

Einsetzen in die erste Gleichung:

$$\begin{aligned} y_1'(x) = && \alpha(\alpha-1)x^{\alpha-2} &= \frac{1}{2x}\alpha x^{\alpha-1} - \frac{1}{2x^2}x^a \\ \Rightarrow && \alpha(\alpha-1)x^{\alpha-2} &= \left(\frac{\alpha}{2} - \frac{1}{2}\right)x^{\alpha-2}, \forall x > 0 \\ \Rightarrow && \alpha(\alpha-1) &= \frac{\alpha}{2} - \frac{1}{2} \\ \Rightarrow && \alpha^2 - \frac{3}{2}\alpha + \frac{1}{2} &= 0 \\ \Rightarrow && \alpha_{1,2} &= \frac{3}{4} \pm \sqrt{\frac{9}{16} - \frac{1}{2}} = \frac{3}{4} \pm \frac{1}{4} \\ \Rightarrow && \alpha_1 &= 1 \wedge \alpha_2 = \frac{1}{2} \end{aligned}$$

$\alpha = 1 : \varphi_1(x) = (1, x)^T$ und $\alpha = \frac{1}{2} : \varphi_2(x) = \begin{pmatrix} \frac{1}{2}x^{-\frac{1}{2}} \\ x^{\frac{1}{2}} \end{pmatrix}$. Daher sind φ_1, φ_2 linear unabhängig: Denn für $\Phi(x) = (\varphi_1(x), \varphi_2(x)) = \begin{pmatrix} 1 & \frac{1}{2}x^{-\frac{1}{2}} \\ x & x^{\frac{1}{2}} \end{pmatrix}$ gilt $\det \Phi(x) = x^{\frac{1}{2}} - \frac{1}{2}xx^{-\frac{1}{2}} = x^{\frac{1}{2}} - \frac{1}{2}x^{\frac{1}{2}} = \frac{1}{2}x^{\frac{1}{2}} \neq 0$ für $x > 0$. Damit ist $\Phi(x)$ ist Fundamentalmatrix und es gilt für die allgemeine Lösung

$$\varphi(x) = \Phi(x)c = \begin{pmatrix} 1 & \frac{1}{2}x^{-\frac{1}{2}} \\ x & x^{\frac{1}{2}} \begin{pmatrix} x-1 \\ 0 \end{pmatrix} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 + \frac{c_2}{2}x^{-\frac{1}{2}} \\ c_1 x + c_2 x^{\frac{1}{2}} \end{pmatrix} = c_1\varphi_1(x) + c_2\varphi_2(x)$$

Im zweiten Schritt erfolgt wieder die Variation der Konstanten. Dies liefert: $\psi_0(x) = \Phi(x) \int_1^x \Phi^{-1}(t)b(t)dt$. Es gilt

$$\Phi^{-1}(x) = \begin{pmatrix} 2 & -\frac{1}{x} \\ -2x^{\frac{1}{2}} & 2x^{-\frac{1}{2}} \end{pmatrix}$$

$$\int_1^x \Phi^{-1}(t)b(t)dt = \int_1^x \begin{pmatrix} 2 & -\frac{1}{t} \\ -2t^{\frac{1}{2}} & 2t^{-\frac{1}{2}} \end{pmatrix} \begin{pmatrix} 1 \\ t \end{pmatrix} dt = \int_1^x \begin{pmatrix} 1 \\ 0 \end{pmatrix} dt = \begin{pmatrix} x-1 \\ 0 \end{pmatrix}$$

$$\psi(x) = \Phi(x) \begin{pmatrix} x-1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{2}x^{-\frac{1}{2}} \\ x & x^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} x-1 \\ 0 \end{pmatrix} = \begin{pmatrix} x-1 \\ x^2 - x \end{pmatrix}$$

ist die spezielle Lösung des inhomogenen DGL.

Allgemeine Lösung des inhomogenen DGL:

$$\psi(x) = \varphi(x)c + \psi_0(x) = \begin{pmatrix} c_1 + \frac{c_2}{2}x^{-\frac{1}{2}} + x - 1 \\ c_1 x + c_2 x^{\frac{1}{2}} + x^2 - x \end{pmatrix}$$

In Komponenten: $y_1 = y_1(x) = c_1 + \frac{c_2}{2}x^{-\frac{1}{2}} + x - 1$ und $y_2 = y_2(x) = c_1 x + c_2 x^{\frac{1}{2}} + x^2 - x$

3.6 Lineare Differentialgleichungen n-ter Ordnung

3.6.1 Lösungstheorie

Übertragung der Resultate über lineare DGL 1. Ordnung auf lineare DGL n-ter Ordnung. Sei $I \subset \mathbb{R}$ ein Intervall, $a_k, f \in \mathcal{C}(I, \mathbb{K})$ seien stetige Funktionen zwischen $0 \leq k \leq n-1$. Dann heißt $L(y) := y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = f(x)$ *homogene lineare Differentialgleichung n-ter Ordnung*, falls $f = 0$. Ansonsten heißt es inhomogen. L heißt *linearer Differentialausdruck*, d.h. $L(c_1y_1 + c_2y_2) = c_1L(y_1) + c_2L(y_2)$.

Satz 3.6.1. *(i) Sei $\mathcal{L}_h := \{\varphi : I \to \mathbb{K} : L(\varphi) = 0\}$ die Menge der Lösungen des homogenen DGL. Dann ist* $\dim \mathcal{L}_h = n$

(ii) Sei $\mathcal{L}_{inh} := \{\varphi : I \to \mathbb{K} : L(\varphi) = f(x)\}$ die Menge der Lösungen der inhomogenen DGL. Dann gilt für ein $\psi_0 \in \mathcal{L}_{inh} : \mathcal{L}_{inh} = \psi_0 + \mathcal{L}_h$.

(iii) n-Lösungen $\varphi_1, \dots, \varphi_n \in \mathcal{L}_h$ der homogenen DGL sind genau dann linear unabhängig, wenn $\exists x_0 \in I : W(x)^1 := \det \begin{pmatrix} \varphi_1(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & & \varphi_n'(x) \\ \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \dots & \varphi_n^{(n-1)}(x) \end{pmatrix} \neq 0 \Leftrightarrow \forall x \in I : W(x) \neq 0$

Beweis: *Die DGL $L(y) = f(x)$ ist äquivalent zu dem inhomogenen DGL-System 1. Ordnung $y_0 := y$.*

$$\begin{aligned} y_0' &= y_1 & y_0' &= y_1 \\ y_1' &= y_2 & y_0'' &= y_2 \\ &\vdots & &\vdots \\ y_{n-2}' &= y_{n-1} & y_0^{(n-1)} &= y_{n-1} \\ y_{n-1}' &= -a_0(x)y_0 - a_1(x)y_1 - \cdots a_{n-1}(x)y_{n-1} + f(x) \end{aligned}$$

Jeder Lösung $\varphi : I \to \mathbb{K}$ von $L(y) = f(x)$ entspricht einer Lösung $y = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} = \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \varphi^{(n-1)} \end{pmatrix} : I \to \mathbb{K}^n$ des Systems oben. Damit folgen die obenstehenden Behauptungen aus den Sätzen

Definition 3.6.2. *Eine Basis $\varphi_1, \dots, \varphi_n \in \mathcal{L}_h$ der homogenen DGL $L(y) = 0$ heißt* Lösungsfundamentalsystem. *Also: Allgemeine Lösung von $L(y) = 0 : \varphi(x) = c_1\varphi_1(x) + c_2\varphi_2(x) + \cdots + c_n\varphi_n(x)$*

Beispiel: $y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$ besitzt die Lösungen $\varphi_1(x) = x, \varphi_2(x) = \sqrt{x}$. Es gilt φ_1, φ_2 sind linear unabhängig. Denn $W(x) = \det \begin{pmatrix} \varphi_1(x) & \varphi_2(x) \\ \varphi_1'(x) & \varphi_2'(x) \end{pmatrix} = \det \begin{pmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{pmatrix} = \frac{1}{2}x\frac{1}{\sqrt{x}} - \sqrt{x} = -\frac{1}{2}\sqrt{x} \neq= 0$. Somit ist $\{\varphi_1, \varphi_2\}$ ein Lösungsfundamentalsystem. Allgemeine Lösung: $\varphi(x) = c_1\varphi_1(x) + c_2\varphi_2(x) = c_1x + c_2\sqrt{x}$

Bestimmung einer speziellen Lösung von $L(y) = f(x)$ mit der Methode der Variation der Konstanten

[1] heißt Wronskideterminante

1. Schritt $L(y) = f(x)$ auf ein DGL-System 1. Ordnung umschreiben und Methode der Variation der Konstanten für DGL-System 1. Ordnung benutzen:

Sei $\Phi(x) = \begin{pmatrix} \varphi_1(x) & \dots & \varphi_n(x) \\ \varphi_1'(x) & & \varphi_n'(x) \\ \vdots & & \vdots \\ \varphi_1^{(n-1)}(x) & \dots & \varphi_n^{(n-1)} \end{pmatrix}$ Fundamentalsystem zum zugehörigen DGL-System.

2. Schritt Ansatz: $\psi(x) = \Phi(x)c(x)$.

$$\begin{aligned} \psi'(x) &= \Phi'(x)c(x) \\ &= A(x)\psi(x) + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(x) \end{pmatrix} \\ &= A(x)\Phi(x)c(x) + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(x) \end{pmatrix} \\ \Phi(x)c'(x) &= b(x) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ f(x) \end{pmatrix} \end{aligned}$$

wobei

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \dots & 1 \\ -a_0(x) & -a_1(x) & -a_2(x) & \dots & -a_{n-1}(x) \end{pmatrix}$$

Oder in Komponenten $\sum_{i=1}^n c_i'(x)\varphi_i^{(k)}(x) = 0$ für $k = 0, \dots, n-2$

3. Schritt Berechnung der $c_i'(x)$ mit Cramerscher Regel

$c_i'(x) = \frac{W_i(x)}{W(x)}, W_i(x) = \det \begin{pmatrix} \varphi_1(x) & \dots & \varphi_{i-1}(x) & 0 & \varphi_{i+1}(x) & \dots \\ \varphi_1'(x) & \dots & \varphi_{i-1}'(x) & 0 & \varphi_{i+1}'(x) & \dots \\ \vdots & & \vdots & \vdots & \vdots & \\ \varphi_1^{(n-1)}(x) & \dots & \varphi_{i-1}^{(n-1)}(x) & 0 & \varphi_{i+1}^{(n-1)} & \dots \end{pmatrix}$

Integration: $c_i(x) = \int_{x_0}^x \frac{W_i(t)}{W(t)} dt$

Spezielle Lösung von $L(y) = f(x) : \psi_0(x) = \sum_{i=1}^n c_i(x)\varphi_i(x)$, wobei $c_i(x) =$

$\int_{x_0}^{x} \frac{W_i(t)}{W(t)} dt$
Das Anfangswertproblem

$$L(y) = f(x), \varphi(x_0) = y_0, \varphi'(x_0) = y_1, \ldots, \varphi^{(n-1)}(x_0) = y_{n-1}, x_0 \in I$$

hat genau eine Lösung.

3.6.2 Reduktion der Ordnung einer Differentialgleichung $L(y) = 0$ bei Kenntnis einer Lösung

Ansatz: $\psi(x) = \varphi(x)u(x)$, wobei $L(\varphi(x)) = 0$.
Demonstration der Methode anhand einer linearen Differentialgleichung 2. Ordnung:

$$y'' + a(x)y' + b(x) = 0 \tag{3.6}$$

Sei $\varphi \neq 0$ eine Lösung von Gleichung 3.6.
Einsetzen in DGL:
$\varphi u'' + (2\varphi' + a(x)\varphi)u' + \underbrace{u(\varphi'' + a(x)\varphi' + b(x)\varphi)}_{=0} = 0 \Leftrightarrow u'' + \left(\frac{2\varphi'}{\varphi} + a(x)\right) u' = 0$. Setzen $v := u'$. Dann ist

$$\boxed{v' + \left(\frac{2\varphi'}{\varphi} + a(x)\right) v = 0} \tag{3.7}$$

eine lineare DGL 1. Ordnung. Es gilt: Ist $v \neq 0$ Lösung der Gleichung 3.7, so ist $\psi(x) = \varphi(x)u(x)$, wobei $u(x) = \int_{x_0}^{x} v(t)dt$, eine von φ unabhängige Lösung von $y'' + a(x)y' + b(x)y = 0$. Denn es gilt

$$\begin{aligned} W(x) = & \det \begin{pmatrix} \varphi(x) & \psi(x) \\ \varphi'(x) & \psi'(x) \end{pmatrix} = \det \begin{pmatrix} \varphi(x) & \varphi(x)u(x) \\ \varphi'(x) & \varphi'(x)u(x) + \varphi(x)u'(x) \end{pmatrix} \\ = & \ \varphi\varphi' u + \varphi^2 u' - \varphi\varphi' u = \varphi^2 u' \\ = & \ \varphi^2 v \neq 0 \end{aligned}$$

□

3.6.3 Praktisches Vorgehen beim Lösen von linearen DGL 2. Ordnung mit nichtkonstanten Koeffizienten

$$L(y) = y'' + a(x)y' + b(x)y = f(x)$$

Der Lösungsweg ist wie folgt:

1. Schritt Finden einer Lösung der homogenen DGL $L(y) = 0$ durch Potenzreihenansatz.

$$\varphi(x) = \sum_{k=-\infty} \infty a_k x^k \quad \varphi(x) = x^\alpha, \alpha \in \mathbb{R}$$

$a(x)$ und $b(x)$ müssen ebenfalls in Potenzreihen entwickelt werden.

2. Schritt Reduktion der Ordnung von $L(y) = 0$ auf eine DGL 1. Ordnung

3. Schritt Berechnung einer Lösung der inhomogenen DGL (beispielsweise mit Variation der Konstanten oder auch Raten)

3.6.4 Lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten

$$L(y) = y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = f(x)$$

wobei a_i konstant ist.

Satz 3.6.3. *Besitzt das Polynom $P_L(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$ die paarweise voneinander verschiedenen Nullstellen $\lambda_k \in \mathbb{C}, k = 1, \ldots, s$ mit den Vielfachheiten $n_k, 1 \leq k \leq s$. Dann besitzt die homogene DGL $L(y) = 0$ ein Lösungsfundamentalsystem aus folgenden Funktionen*

$$\varphi_{km}(x) := x^m e^{\lambda_k x} \quad m = 0, 1, \ldots, n_k - 1 \quad k = 1, \ldots, s$$

Proof. 1. Schritt Man zeigt, dass $\varphi_{km}(x) = x^m e^{\lambda_k x}$ Lösungen von $L(y) = 0$ sind.
Ansatz: $\varphi(x) := e^{\lambda x}$. Dann gilt $\varphi^{(k)}(x) = \lambda^k e^{\lambda x}$.

$$L(\varphi(x)) = (\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0)e^{'\lambda x} = 0$$

Die obige Gleichung ist genau dann 0, wenn

$$P_L(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0 = 0$$

Somit sind $\varphi_k(x) = e^{\lambda_k x}$ Lösungen von $L(y) = 0$, wobei λ_k Nullstellen von dem Polynom P_L sind.
Sind alle Nullstellen von P_L einfach, d.h. $n_k = 1, k = 1, \ldots, n$, so bilden die Funktionen $\varphi_k(x) = e^{\lambda_k x}, k = 1, \ldots, n$, bereits ein Fundamentalsystem. Ein Problem besteht bei mehrfachen Nullstellen. Dazu berechnet man $L(x^m e^{\lambda x})$:
$L(x^m e^{\lambda x}) = L\left(\frac{\partial^m}{\partial\lambda^m}(e^{\lambda x})\right)$
Vertauschung der Differentiation (nach Satz von Schwartz): $= \frac{\partial^m}{\partial\lambda^n} L(e^{\lambda x}) = \frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)e^{\lambda x})$. Ist $\lambda = \lambda_j$ eine n_j-fache Nullstelle von $P_L, P_L(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{n_k}$, dann gilt $\frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)e^{\lambda x}) = 0$ für $m = 0, 1, \ldots, n_j - 1$. Denn $P_L(\lambda) = Q_j(\lambda)(\lambda - \lambda_j)^{n_j}, Q_j(\lambda) = \prod_{k=1, k\neq j}^s (\lambda - \lambda_j)^{n_k}$.
Leibnitzsche Regel:

$$\begin{aligned}
\frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)e^{\lambda x}) &= \sum_{k=0}^m \binom{m}{k} \frac{\partial^k}{\partial\lambda^k}(P_L(\lambda)) \frac{\partial^{m-k}}{\partial\lambda^{m-k}}(e^{\lambda k}, \frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)) \\
&= \frac{\partial^m}{\partial\lambda^m}((\lambda - \lambda_j)^{n_j} Q_j(\lambda)) \\
&= \sum_{k=0}^m \frac{\partial^k}{\partial\lambda^k}((\lambda - \lambda_j)^{n_j} \frac{\partial^{m-k}}{\partial\lambda^{m-k}}(Q_j(\lambda)), \frac{\partial^k}{\partial\lambda^k}(\lambda - \lambda_j)^{n_j} \\
&= n_j(n_j - 1)\ldots(n_j - k + 1)(\lambda - \lambda_j)^{n_j - k} \\
&\Rightarrow \frac{\partial^k}{\partial\lambda^k}((\lambda - \lambda_j)^{n_j}) \\
&= 0
\end{aligned}$$

Somit ist $\frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)) = 0$ für $m = 0, 1, \ldots, n_j - 1$ und $\frac{\partial^m}{\partial\lambda^m}(P_L(\lambda)e^x) = 0$ für $m = 0, 1, \ldots, n_j - 1$ und $j = 1, \ldots, s$.

2. Schritt Die $\varphi_{km}(x) = x^m e^{\lambda_k x}, m = 0, 1, \ldots, n_k - 1, n = n_1 + \cdots + n_s, k = 1, \ldots, s$ sind stets linear unabhängig.

Lemma 3.6.4. *Seien $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ verschiedene Zahlen und $P_1, \ldots, P_r$ Polynome auf einem offenen Intervall $I \subset \mathbb{R}$. Dann gilt:*

$$\sum_{j=1}^{r} P_j(x) e^{\lambda_j x} = 0 \, \forall x \in I \Rightarrow P_1 = P_2 = \cdots = P_r = 0$$

Beweis durch Induktion: $r = 1 : P_1(x)e^{\lambda_1 x} = 0 \, \forall x \in I \Rightarrow P_1(x) = 0 \, \forall x \in I \Rightarrow P_1 = 0$.
Induktionsschritt mit $(r-1) \Rightarrow r$: Sei $\sum_{j=1}^{r} P_j(x)e^{\lambda_j x} = 0$. Damit folgt $\sum_{j=1}^{r-1} P_j(x)e^{(\lambda_j - \lambda_r)x} + P_r(x) = 0$. Sei nun m der Grad des Polynoms plus 1. Dann gilt folgendes (nach m-maliger Differentiation nach x):

$$\begin{aligned}
0 &= \frac{\partial^m}{\partial x^m} \left(\sum_{j=1}^{r-1} P_j(x) e^{(\lambda_j - \lambda_r)x} + P_r(x) \right) \\
&= \frac{\partial^m}{\partial x^m} \left(\sum_{j=1}^{r-1} P_j(x) e^{(\lambda_j - \lambda_r)x} \right) \\
&= \sum_{j=1}^{r-1} \frac{\partial^m}{\partial x^m} (P_j(x) e^{(\lambda_j - \lambda_r)x}) \\
&= \sum_{j=1}^{r-1} Q_j(x) e^{(\lambda_j - \lambda_r)x} \\
&\textit{wobei } Q_j - (\lambda_j - \lambda_r)^m P_j(x) + R_j(x)
\end{aligned}$$

R_j ist ebenfalls ein Polynom. Aus der Induktionsvoraussetzung folgt nun: $Q_j(x) = 0 \, \forall x \in I$, d.h. $Q_1 = Q_2 = \cdots = Q_{r-1} = 0 \Rightarrow P_j = R_j = 0, j = 1, \ldots, r-1 \Rightarrow 0 = \sum_{j=1}^{r} P_j(x)e^{\lambda_j x} = P_r(x)e^{\lambda_r x} \Rightarrow P_r = 0$ auf I.

Jetzt zur linearen Unabhängigkeit von $\varphi_{jk}(x) = x^k e^{\lambda_j x}$ mit $k = 0, 1, \ldots, n_j - 1, j = 1, \ldots, s$ gilt für alle $x \in \mathbb{R}$:

$$\begin{aligned}
\sum_{j=1}^{s} \sum_{k=0}^{n_j - 1} c_{jk} \varphi_{jk}(x) &= \sum_{j=1}^{s} \sum_{k=0}^{n_j - 1} c_{jk} x^k e^{\lambda_j x} \\
&= \sum_{j=1} s \left(\sum_{k=0}^{n_j - 1} c_{jk} x^k \right) e^{\lambda_j x} = 0
\end{aligned}$$

Nach dem Lemma folgt $c_{jk} = 0$.

□

Praktisches Vorgehen anhand eines Beispiels Man finde die allgemeine reelle Lösung der DGL $x + 2\mu x + \omega_0^2 x = 0$ $(\omega_0 > 0, \mu \geq 0)$ der gedämpften Schwingung. Man sucht die Lösungen der Form $t^m e^{\lambda_k t}$.
Ansatz: $\varphi(t) = e^{\lambda t}, \varphi'(t) = \lambda e^{\lambda t}, \varphi''(\lambda^2 e^{\lambda t})$ Einsetzen in DGL $L(\varphi(t)) = (\lambda^2 + 2\mu\lambda + \omega_0^2)e^{\lambda t} = 0 \Leftrightarrow \lambda^2 + 2\mu\lambda + \omega_0^2$. Nullstellen: $\lambda = -\mu \pm \sqrt{\mu^2 - \omega_0^2}$.

1. Fall $0 \leq \mu < \omega_0 : \lambda_{1,2} = -\mu \pm i\omega, \omega = \sqrt{\omega_0^2 - \mu^2} > 0$
   Lösungsfundamentalsystem: $\varphi_1(t) = e^{\lambda_1 t} = e^{-\mu t} e^{i\omega t}, \varphi_2(t) = e^{\lambda_2 t} = e^{-\mu t} e^{-i\omega t}$
   $e^{i\omega t} = \cos(\omega t) + i \sin(\omega t)$. Betrachten $\begin{array}{rcl} \psi_1(t) & = & \Re\varphi_1(t) = e^{-\mu t}\cos(\omega t) \\ \psi_2(t) & = & \Im\varphi_2(t) = e^{-\mu t}\sin(\omega t) \end{array}$. Es gilt $L(\psi_1(t)) = L(\psi_2(t)) = 0$. Nun ist noch zu zeigen, dass $\{\psi_1, \psi_2\}$ ein Lösungsfundamentalsystem ist:

$$\begin{aligned} \det \begin{pmatrix} \psi_1(t) & \psi_2(t) \\ \psi_1'(t) & \psi_2'(t) \end{pmatrix} &= \det \begin{pmatrix} e^{-\mu t}\cos(\omega t) & e^{-\mu t}\sin(\omega t) \\ -\mu e^{-\mu t}\cos(\omega t) & -\mu e^{-\mu t} sin(\omega t) \\ -e^{-\mu t}\omega\sin(\omega t) & +e^{-\mu t}\omega\cos(\omega t) \end{pmatrix} \\ &= -\mu e^{-2\mu t}\cos(\omega t)\sin(\omega t) + \omega e^{-2\mu t}\cos^2(\omega t) \\ &\quad + \mu e^{-2\mu t}\cos(\omega t)\sin(\omega t) + \omega e^{-2\mu t}\sin^2(\omega t) \\ &= \omega e^{-2\mu t} \neq 0 \end{aligned}$$

2. Fall $\mu = \omega_0$: Nullstellen $\lambda_1 = \lambda_2 = -\mu$. Damit ist $\varphi_1(t) = e^{-\mu t}, \varphi_2(t) = te^{-\mu t}$ ein Fundamentalsystem.

3. Fall $\mu > \omega_0$: Dann ist $\lambda_1 = - + \sqrt{\mu^2 - \omega_0^2}, \lambda_2 = -\mu - \sqrt{\mu_2 - \omega_0^2} < 0$. Man setzt: $\mu_j := -\lambda_j > 0, j = 1, 2$. Dann ist $\varphi_1(t) = e^{\lambda_1 t} = e^{-\mu_1 t}, \varphi_2(t) = e^{\lambda_2 t} e^{-\mu_2 t}$ ist Fundamentalsystem.

Zusammenfassend hat man für die allgemeine Lösung von $x'' + 2\mu x' + \omega_0^2 x = 0$:

$$\varphi(t) = c_1 e^{-\mu t}\cos(\omega t) + c_2 e^{-\mu t}\sin(\omega t) \tag{3.8}$$

$$\varphi(t) = c_1 e^{-\mu t} + c_2 t e^{-\mu t} \text{ für } \mu = \omega_0 \tag{3.9}$$

$$\varphi(t) = c_1 e^{-\mu_1 t} + c_2 e^{-\mu_2 t} \text{für } \omega_0 < \mu \tag{3.10}$$

$$\mu_1 = -\lambda_1 = +\mu - \sqrt{\mu^2 - \omega_0^2}$$

Satz 3.6.5. *Sei $G \subset \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ ein Gradientenfeld. Dann erhält man alle Skalarfelder (Stammfunktionen) $\Phi : G \to \mathbb{R}$ von f, d.h. $f = \ grad\Phi$, in der Form*

$$\Phi = \Phi_0 + c \qquad c \in \mathbb{R}$$

mit $f = \ grad\Phi_0$.
***Beweis**: Aus $f = \ grad\Phi_0 \Rightarrow \ grad\Phi = \ grad\Phi_0 + c = \ grad\Phi_0 = f$. Umgekehrt sei $f = \ grad\Phi = \ grad\Phi_0 \Rightarrow \ grad(\Phi - \Phi_0) = 0 \Rightarrow (\Phi - \Phi_0)' = 0$. Da g ein Gebiet ist, folgt $\Phi - \Phi_0 = c$.* □

Satz 3.6.6. *Sei $G \subset \mathbb{R}^n$ ein Gebiet und $f : G \to \mathbb{R}^n$ ein stetiges Vektorfeld. Dann gilt, dass f genau dann ein Gradientenfeld ist, $\int_\gamma f(x)gx$ wegunabhängig ist.*
Beweis:

„⇒“ *Sei $f = \; grad\Phi = \Phi'$. Seien $A, B \in G$ zunächst zwei Punkte, die durch einen stetig differenzierbaren Weg $\gamma : [a, b] \to \mathbb{R}^n$ mit $\gamma(a) = A$ und $\gamma(b) = B$ verbunden sind. Dann ist: $\int_\gamma f(x)dx = \int_a^b f(\gamma(t))\gamma'(t)dt = \int_a^b \Phi'(\gamma(t))\gamma'(t)dt = \int_a^b \frac{d}{dt}(\Phi(\gamma(t)))dt = \Phi(\gamma(b)) - \Phi(\gamma(a)) = \Phi(A) - \Phi(B)$.*
Sieen $A, B \in G$ beliebige und γ ein stückweise stetig differenzierbarer Weg, der A und B verbindet. Ein solcher existiert, da G polygonzugzusammenhängend ist. Also $\gamma = \gamma_1 \oplus \gamma_2 \oplus \cdots \oplus \gamma_n$, wobei γ_k stetig differenzierbar. Damit ist $\int_\gamma \; grad\Phi(x)dx = \sum_{k=1}^n \int_{\gamma_k} \; grad\Phi(x)dx$.

„⇐“ *Sei $\xi \in G$ und $B_\varepsilon(\xi) \subset G$ eine offene ε-Kugel. Für beliebige $h \in \mathbb{R}^n$ mit $\xi + h \in B_\varepsilon(\xi)$ ist $\sigma(t) = \xi + th \in B_\varepsilon(\xi)$ für $0 \leq t \leq 1$. Setzen $\Phi(x) := \int_A^x f(y)dy, A \in G$ fest, $x \in G$ beliebig. Dann ist $\Phi(\xi + h) - \Phi(\xi) = \int_A^{\xi+h} f(y)dy - \int_A^\xi f(y)dy = \int \xi^{\xi+h} f(y)dy = \int_\sigma f(y)dy = \int_0^1 f(\sigma(t))\sigma'(t)dt = \int_0^1 f(\xi + th)hdt$. Ferner hat man $\int_\sigma f(\xi)dy = \int_0^1 f(\xi)\sigma'(t)dt = \int_0^1 f(\xi)hdt = f(\xi)h$. Somit hat man $|\Phi(\xi + h) - \Phi(\xi) - f(\xi)h| = |\int_\sigma f(y) - f(\xi))dy| \leq \max_{y \in \{\xi + th : 0 \leq t \leq 1} ||f(y) - f(\xi)||_2 L(\sigma)$. Da f stetig ist, folgt $\max_{y \in \{\xi + th : 0 \leq t \leq 1} ||f(y) - f(\xi)||_2 \xrightarrow{h \to 0} 0$. Damit gilt $f(\xi) = \Phi(\xi) = \; grad\Phi(\xi)$.*

□

Folgerung: Es gilt $\int_\gamma \mathrm{grad}\Phi dx = \Phi(B) - \Phi(A), \gamma(a) = A, \gamma(b) = B$ für jeden stetig differenzierbaren Weg γ in einem Gebiet G und Anfangspunkt A und Endpunkt B und jedes stetig differenzierbare Skalarfeld Φ.

Ein praktisches Differenzierbarkeitskriterium dafür, dass ein Vektorfeld ein Gradientenfeld ist:

(i) Notwendiges Kriterium
Sei $f : \begin{pmatrix} f_1 \\ \cdots \\ f_n \end{pmatrix} : G \to \mathbb{R}^n$ ein stetig differenzierbares Gradientenfeld auf einer offenen Menge G. Dann gilt notwendigerweise: $\frac{\partial f_j}{\partial x_k} = \frac{\partial f_k}{\partial x_j}, k_{xj} = 1, \ldots, n$. (Integrabilitätskriterien)
Beweis: (Schwatzscher Satz): $f = \;$ grad $\Phi = \Phi'$, da f stetig differenzierbar. Damit ist Φ zweimal stetig differenzierbar. Wegen dem Schwatzschen Satz folgt nun, $\frac{\partial f_j}{\partial x_k} = \frac{\partial^2 \Phi}{\partial x_k \partial x_j} = \frac{\partial^2 \Phi}{\partial x_j \partial x_k} = \frac{\partial f_k}{\partial x_j}$. Dann ist f ein Gradientenfeld.

(ii) Hinreichende Bedingungen
Sei $G \subset \mathbb{R}^n$ eine offene und sternförmige Menge, d.h. $\exists a \in G \forall x \in G \forall t \in [0, 1]$: $a + tx \in G$, und $f = \begin{pmatrix} f_1 \\ \cdots \\ f_n \end{pmatrix} : G \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld mit

$\frac{\partial f_j}{\partial x_k} = \frac{\partial f_k}{\partial x_j}, j,k = 1,\ldots,n$. Dann ist f ein Gradientenfeld, d.h. $\exists \Phi : G \to \mathbb{R} : f =$ grad Φ.
Beweis: Sei o.B.d.A. $a = 0, \sigma(t) = tx$, wobei t zwischen 0 und 1 liegt. $\Phi(x)$ sei $\int_\sigma f(y)dy$.

$$
\begin{aligned}
\Phi(x) &= \int_0^1 f(tx)xdt = \int_0^1 \sum_{j=1}^n f_j(tx)x_j dt \\
\frac{\partial \Phi(x)}{\partial x_k} &= \int_0^1 \frac{\partial}{\partial x_k}(f(tx)x)dt \\
&= \int_0^1 \frac{\partial}{\partial x_k}\left(\sum_{j=1}^n f_j(tx)x_j\right) dt \sum_{j1}^n \frac{\partial}{\partial x_k}(f_j(tx)x_j)dt \\
&= \sum_{j=1}^n \left(\frac{\partial f_j(tx)}{\partial x_k}x_j + f_j(tx)\frac{\partial x_j}{\partial x_k}\right) \\
&= f_k(tx) + \sum_{j=1}^n \frac{\partial f_j(tx)}{\partial x_k}x_j = f_k(tx) + \sum_{j=1}^n \frac{\partial f_j(tx)}{\partial u_k}\frac{\partial u_k}{\partial x_k}x_j \\
&= f_k(tx) + \sum_{j=1}^n \frac{\partial f_j(tx)}{\partial u_k}tx_j = f_k(tx) + \sum_{j=1}^n \frac{\partial f_k(tx)}{\partial u_j}tx_j \\
&= \frac{d}{dt}(tf_k(tx)) \\
\text{Somit ist } \frac{\partial \Phi(x)}{\partial x_k} &= \int_0^1 \frac{d}{dt}(ff_k(tx))dt = f_k(x)
\end{aligned}
$$

3.6.5 Praktische Bestimmung von Φ

Sei o.B.d.A. $f =$ grad $\Phi, \Phi = \Phi(x_1, x_2), f_1 = \frac{\partial \Phi}{\partial x}, f_2 = \frac{\partial \Phi}{\partial x_2}$.

1. Man wählt einen Polygonzug γ in Richtung der Koordinatenachsen in G. $\Phi(x) = \int_a^x f(y)dy$ mit $a = (a_1, a_2)$ und $x = (x_1, x_2)$. $\Phi(x) = \int_\gamma f(y)dy = \int_{\gamma_1} f(y)dy + \int_{\gamma_2} f(y)dy$ mit $\gamma = \gamma_1 \oplus \gamma_2, \gamma_1(t) = (t, a_2), \gamma_2(t) = (x_1, t)$.

Index

Abbildung
lineare, 10
offene, 32
Ableitung, 12
partielle, 14
Auflösbarkeit
lokale, 29

Banachraum, 5
Bolzano/Weierstrass
Satz von, 8

definit, 25
Diffeomorphismus, 32
Differential, 12
Differentialausdruck
linearer, 63
Differentialgleichung
einfache, 41
homogene, 63
lineare, 63
Differentialgleichungen
partielle, 42
Differentialgleichungssystem
homogenes lineares, 46
Doppelpunkt, 38

endlichdimensional, 8

Fläche, 39
Fréchet-Ableitung, 12
Fundamentalmatrix, 61
Fundamentalsystem, 49

Gebiet, 20, 69
Gradient, 17

Homogenität, 5

implizit, 28
indefinit, 25
Isoklinen, 42

Konvergenz
gleichmässige, 7
punktweise, 7
Kurve, 37

Lösung
allgemeine, 43
spezielle, 43
Lösungsfundamentalsystem, 64
Lagrange Multiplikator, 33
Linienelement, 42
Lipschitzbedingung, 55

Matrix
Hessesche, 26
Maximum
isoliertes, 25
lokales, 25
Maximumnorm, 5
Menge
sternförmig, 70
Minimum
isoliertes, 25
lokales, 25

Niveaufläche, 39
Norm, 5
euklidische, 5

Operatorennorm, 11

Polygonzug, 20

Raum

normierter, 5
regulär, 38
Richtungsableitung, 14
Richtungsfeld, 42

Schnittwinkel, 38
semidefinit, 25
singulär, 38
Summennorm, 5

Tangentialebene, 39
Tangentialvektor, 38

Wronskideterminante, 64