

Kapitel 1

Integrale in Anwendungen

1.1 Einfache Berechnung von Integralen

1.1.1 Lineare Modifikation

Wir betrachten den folgenden Satz.

Satz 1.1 *Integration durch lineare Modifikation*

Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ eine integrierbare Funktion mit Stammfunktion $F : \mathbb{R} \rightarrow \mathbb{R}$ und $m, q, x_0, x_E \in \mathbb{R}$ mit $m \neq 0$ und $x_0 < x_E$, dann gilt folgendes.

$$(a) \quad \int f(m \cdot x + q) dx = \frac{1}{m} \cdot F(m \cdot x + q) + c$$

$$(b) \quad \int_{x_0}^{x_E} f(m \cdot x + q) dx = \frac{1}{m} \cdot (F(m \cdot x_E + q) - F(m \cdot x_0 + q))$$

Beweis: Übung. □

Beispiele:

- Wir betrachten ein *unbestimmtes Elementarintegral* und eine *lineare Modifikation*.

$$\begin{aligned} \text{elementar: } F(x) &= \int \cos(x) dx = \sin(x) + c \\ \text{linear modifiziert: } F(x) &= \int \cos(2x + 3) dx = \frac{1}{2} \cdot \sin(2x + 3) + c \end{aligned} \tag{1.1}$$

- Wir betrachten das *unbestimmte Integral*

$$\underline{\underline{F(x)}} = \int (7x - 2)^3 dx = \frac{1}{7} \cdot \frac{1}{4} \cdot (7x - 2)^4 + c = \frac{1}{28} \cdot (7x - 2)^4 + c. \tag{1.2}$$

- Wir betrachten das *unbestimmte Integral*

$$\underline{\underline{F(x)}} = \int 3^{2x+9} dx = \frac{1}{2} \cdot \frac{1}{\ln(3)} \cdot 3^{2x+9} + c = \frac{1}{2 \ln(3)} \cdot 3^{2x+9} + c. \tag{1.3}$$

Bemerkungen:

- i) *Linear modifizierte Elementar-Integrale* machen ca. 90% aller *Integrale* in Alltag, Naturwissenschaft, Technik und Wirtschaft aus.
- ii) Die Methode der *linearen Modifikation* ist ein einfacher Spezialfall der allgemeineren Methode der *Substitution*.

Liste mit den wichtigsten *linear modifizierten Standard-Integralen*:

- $\int (m \cdot x + q)^p dx = \frac{1}{m \cdot (p+1)} \cdot (m \cdot x + q)^{p+1} + c$
- $\int a^{m \cdot x + q} dx = \frac{1}{m \cdot \ln(a)} \cdot a^{m \cdot x + q} + c$
- $\int y_0 \cdot a^{\frac{x-x_0}{\Sigma}} dx = \frac{\Sigma}{\ln(a)} \cdot y_0 \cdot a^{\frac{x-x_0}{\Sigma}} + c$
- $\int A \cdot \sin(\omega \cdot t + \varphi) dt = -\frac{A}{\omega} \cdot \cos(\omega \cdot t + \varphi) + c$

1.1.2 Integration mit Software

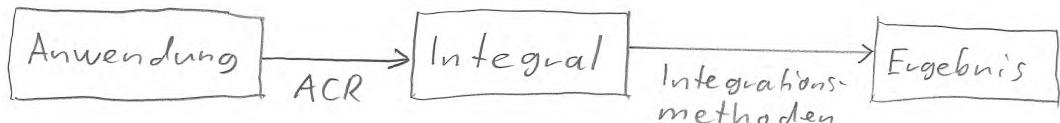
Siehe Demonstrationen mit Python/Numpy & Python/Sympy.

1.2 Archimedes-Cauchy-Riemann-Approximationsprozess

1.2.1 Einleitung

Einleitende Bemerkungen:

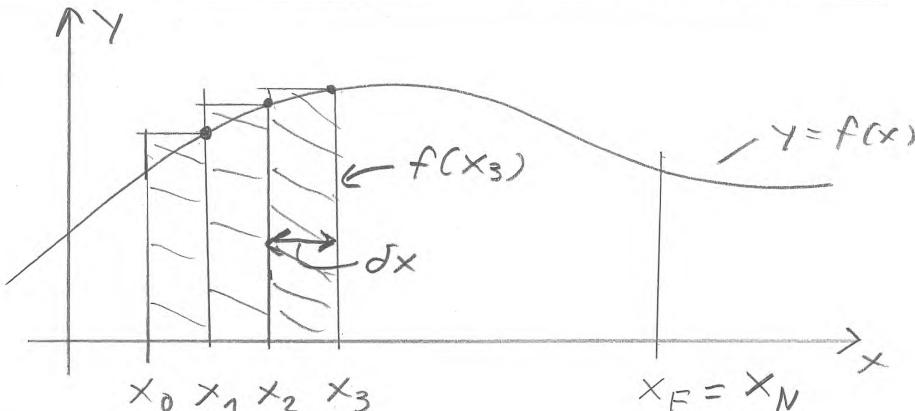
- i) Die Abkürzung ACR-Prozess steht für ARCHIMEDES-CAUCHY-RIEMANN-*Approximationsprozess*.
- ii) Sinn und Zweck eines ACR-Prozess ist nicht die Berechnung von *Integralen* sondern das Auffinden von *Integralen* in der Praxis. Das Konzept ist im folgenden Schema dargestellt.



- iii) Aus den historischen Wurzeln des ACR-Prozesses ergibt sich auch die Erklärung für die Schreibweise mit dem "Integral-Haken".

1.2.2 Historische Methode zur Berechnung von Integralen

Wir betrachten den *Graphen* einer *stetigen Funktion* $f : \mathbb{R} \rightarrow \mathbb{R}$ entlang eines *reellen Intervalls* $[x_0, x_E]$. Die Situation ist im folgenden *x-y-Diagramm* dargestellt.



Schon bevor die NEWTON-LEIBNIZ-*Formel* bekannt war, gab es Methoden zur näherungsweisen Berechnung der *Fläche* A entlang des *reellen Intervalls* $[x_0, x_E]$ zwischen der x -Achse und dem *Graphen* von f . Dabei wurde typischerweise nach den folgenden Schritten vorgegangen.

S1 Man wählt ein $N \in \mathbb{N}^+$ und definiert sich die Werte

$$\delta x := \frac{x_E - x_0}{N} \quad (1.4)$$

$$x_k := x_0 + k \cdot \delta x \quad \text{für } k \in \{1, \dots, N\}. \quad (1.5)$$

S2 Man berechnet für $k \in \{1, \dots, N\}$ die *Flächeninhalte* der *Rechtecke* gemäß

$$\delta A_k := f(x_k) \cdot \delta x. \quad (1.6)$$

S3 Für N gross genug erhält man in guter Näherung

$$A \approx \sum_{k=1}^N \delta A_k = \sum_{k=1}^N f(x_k) \cdot \delta x. \quad (1.7)$$

Unter Verwendung des modernen CAUCHY-Grenzwertbegriffs erhält man

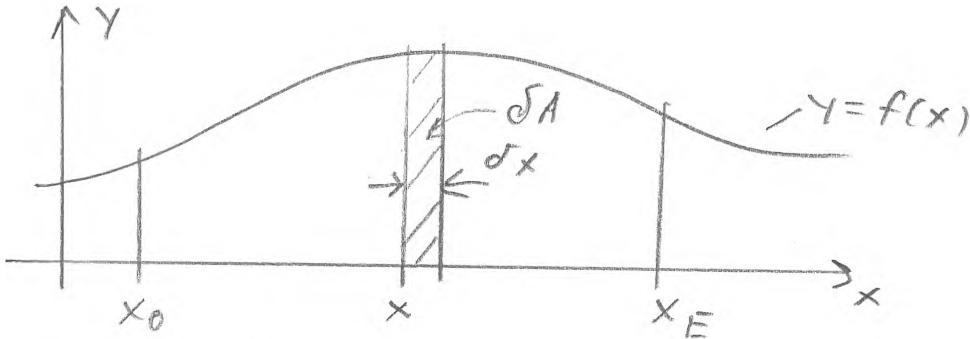
$$A = \int_{x_0}^{x_E} f(x) dx = \lim_{N \rightarrow \infty} \sum_{k=1}^N f(x_k) \cdot \delta x \approx \sum_{k=1}^N f(x_k) \cdot \delta x. \quad (1.8)$$

Bemerkungen:

- i) Aus (1.8) lässt sich die Herkunft des “Integral-Hakens” erkennen: Es ist ein grosses “S” für *Summe*. Analog wurde aus dem Faktor δx das *Differentialsymbol* dx .
- ii) Früher berechnete man *Integrale* im Sinne von (1.8) näherungsweise durch *Summen*.
- iii) Heute berechnet man *Summen* im Sinne von (1.8) durch *Integrale* mit Hilfe der NEWTON-LEIBNIZ-Formel.

1.2.3 Konzeption am Urbeispiel

Wir betrachten den *Graphen* einer *stetigen Funktion* $f : \mathbb{R} \rightarrow \mathbb{R}$ entlang eines *reellen Intervalls* $[x_0, x_E]$. Die Situation ist im folgenden *x-y-Diagramm* dargestellt.



Um die *Fläche* zwischen der *x-Achse* oberhalb des *Intervall* $[x_0, x_E]$ und dem *Graphen* von f zu berechnen, verwenden wir einen ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten einen kleinen Streifen an der Position x der Breite δx . Die Fläche dieses Streifens beträgt

$$\underline{\delta A} \approx \underline{f(x)} \cdot \delta x. \quad (1.9)$$

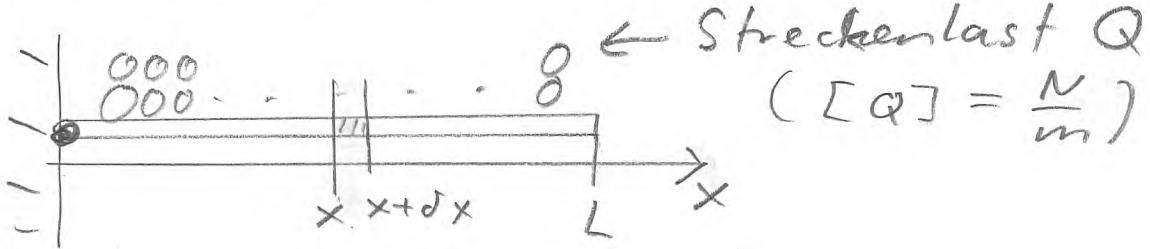
S2 Global: Durch *Integration* über x können wir die gesamte *Fläche* A berechnen. Wir erhalten

$$\underline{\underline{A}} = \int_{x_0}^{x_E} f(x) dx = \left[F(x) \right] \Big|_{x_0}^{x_E} = \underline{\underline{F(x_E) - F(x_0)}}. \quad (1.10)$$

1.2.4 Anwendungen

1.2.4.1 Drehmoment einer Streckenlast

Wir betrachten einen drehbar gelagerten Balken der Länge L , welcher durch eine konstante Streckenlast Q gleichmässig belastet wird. Die Situation ist in der folgendne Skizze dargestellt.



Um das *Drehmoment* der *Streckenlast* zu berechnen, verwenden wir einen ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten ein kleines Teilstück des Balkens an der Position x mit Länge δx . Das *Drehmoment*, welches die *Streckenlast* auf das Teilstück ausübt beträgt

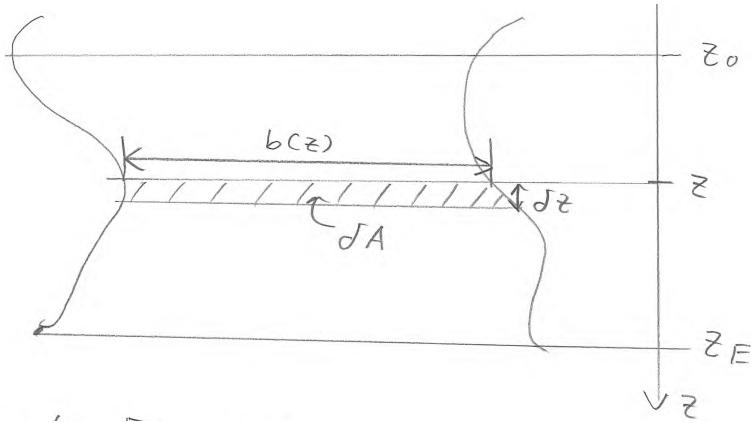
$$\underline{\delta M} \approx r(x) \cdot \underline{\delta F} = x \cdot Q \cdot \delta x. \quad (1.11)$$

S2 Global: Durch *Integration* über x können wir das gesamte *Drehmoment* der *Streckenlast* berechnen. Wir erhalten

$$\underline{\underline{M}} = \int_0^L x \cdot Q \, dx = Q \int_0^L x \, dx = Q \cdot \frac{1}{2} \cdot [x^2] \Big|_0^L = \frac{1}{2} \cdot Q \cdot (L^2 - 0^2) = \underline{\underline{\frac{1}{2} \cdot Q \cdot L^2}}. \quad (1.12)$$

1.2.4.2 Staukräfte

Wir betrachten eine *Staumauer* mit veränderlicher Breite $b(z)$, welche ein Medium mit Dichte ρ staut. Die Situation ist in der folgenden Skizze dargestellt.



Um die *Kraft* des gesamten Mediums auf die Staumauer zu berechnen, verwenden wir einen ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten einen kleinen horizontalen Streifen an der Position z mit Höhe δt . Die *Kraft*, welche das Medium auf diesen Streifen ausübt, beträgt

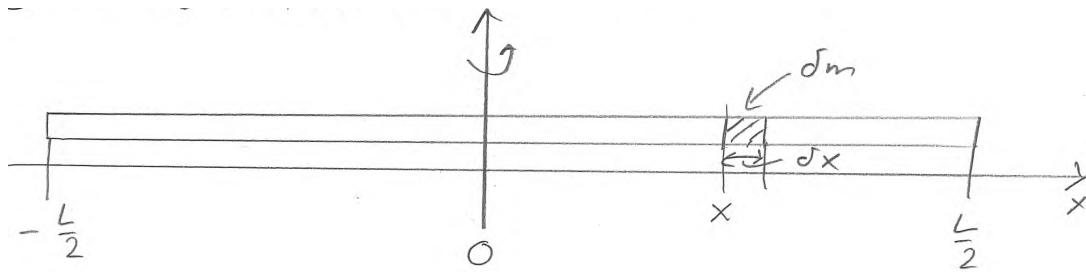
$$\underline{\delta F} \approx p(z) \cdot \underline{\delta A} = \rho \cdot g \cdot z \cdot b(z) \cdot \delta z. \quad (1.13)$$

S2 Global: Durch *Integration* über z können wir die *Kraft* des gesamten Mediums auf die Staumauer berechnen. Wir erhalten

$$\underline{\underline{F}} = \int_{z_0}^{z_E} \rho \cdot g \cdot z \cdot b(z) \, dz = \rho \cdot g \int_{z_0}^{z_E} z \cdot b(z) \, dz = \underline{\underline{\dots}} \quad (1.14)$$

1.2.4.3 Massenträgheitsmomente

Wir betrachten einen dünnen Stab mit *Länge* L und *Masse* m , welcher gedreht werden soll um eine *Drehachse*, die *senkrecht* zum Stab durch seinen *Schwerpunkt* verläuft. Die Situation ist in der folgenden Skizze dargestellt.



Um das *Trägheitsmoment* des *Stabes* zu berechnen, verwenden wir einen **ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess**. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten ein kleines Teilstück des Stabes an der Position x mit Länge δx . Für die Masse δm des Teilstücks gilt

$$\frac{\delta m}{m} = \frac{\delta x}{L} \Rightarrow \delta m = \frac{m}{L} \cdot \delta x. \quad (1.15)$$

Das *Trägheitsmoment* des Teilstücks beträgt

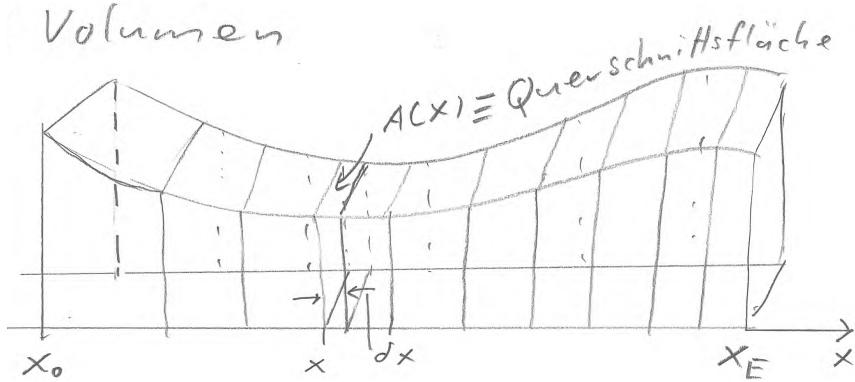
$$\underline{\delta I} \approx r^2(x) \cdot \delta m = x^2 \cdot \frac{m}{L} \cdot \delta x = \frac{m}{L} \cdot x^2 \cdot \delta x. \quad (1.16)$$

S2 Global: Durch *Integration* über x können wir das gesamte *Trägheitsmoment* des *Stabes* berechnen. Wir erhalten

$$\begin{aligned} \underline{\underline{I}} &= \int_{-\frac{L}{2}}^{\frac{L}{2}} \frac{m}{L} \cdot x^2 dx = 2 \cdot \frac{m}{L} \int_0^{\frac{L}{2}} x^2 dx = 2 \cdot \frac{m}{L} \cdot \frac{1}{3} \cdot \left[x^3 \right] \Big|_0^{\frac{L}{2}} = \frac{2}{3} \cdot \frac{m}{L} \cdot \left(\left(\frac{L}{2} \right)^3 - 0^3 \right) \\ &= \frac{2}{3} \cdot \frac{m}{L} \cdot \frac{L^3}{2^3} = \underline{\underline{\frac{1}{12} \cdot m \cdot L^2}} \quad (1.17) \end{aligned}$$

1.2.4.4 Volumen

Wir betrachten einen *Körper* mit veränderlichem *Querschnitt* $A(x)$. Die Situation ist in der folgenden Sizze dargestellt.



Um das *Volumen* des *Körpers* zu berechnen, verwenden wir einen ARCHIMEDES-CAUCHY-RIEMANN-*Approximationsprozess*. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten eine kleine Teilstück des Körpers an der Position x mit *Querschnittsfäche* $A(x)$ und *Länge* δx . Das *Volumen* des Teilstücks beträgt

$$\underline{\delta V} \approx \underline{A(x)} \cdot \underline{\delta x}. \quad (1.18)$$

S2 Global: Durch *Integration* über x können wir das gesamte *Volumen* des Körpers berechnen. Wir erhalten

$$\underline{\underline{V}} = \int_{x_0}^{x_E} A(x) \, dx = \underline{\underline{\dots}}. \quad (1.19)$$

Kapitel 2

Vektoranalysis

2.1 Grundlagen

2.1.1 Skalarfelder

2.1.1.1 Definition

Wir betrachten die folgende Definition.

Definition 2.1 *Reellwertige Funktion in mehreren reellen Variablen*

Seien $n \in \mathbb{N}^+$, $A \subseteq \mathbb{R}^n$ und $B \subseteq \mathbb{R}$. Eine *Funktion* auf A der Form $f : A \rightarrow B$ heisst *reellwertige Funktion in n reellen Variablen*.

Bemerkungen:

- i) Die *reellen Variablen* werden nach dem *Funktionsnamen* f in runden Klammern aufgezählt, jeweils durch ein Semicolon getrennt.

$$n = 1 : \quad f(x) = \dots \tag{2.1}$$

$$n = 2 : \quad f(x; y) = \dots \tag{2.2}$$

$$n = 3 : \quad f(x; y; z) = \dots \tag{2.3}$$

$$\text{allg: } f(x_1; x_2; \dots; x_n) = \dots \tag{2.4}$$

- ii) Wird die *Definitionsmenge* A als *Teilmenge* eines *geometrischen Raumes* interpretiert, dann wird f als *Skalarfeld* auf A bezeichnet.
- iii) Im allgemeinen können die *reellen Variablen* $x_1, x_2, \dots, x_n \in \mathbb{R}$ und der *Funktionswert* $f(x_1; x_2; \dots; x_n)$ beliebige und auch unterschiedliche Masseinheiten tragen.
- iv) Gilt $n \geq 2$ und enthält die *Definitionsmenge* A einen auch noch so kleinen *Würfel* in nD, dann kann f nicht *injektiv* sein.
- v) Aus dem Wert von n kann apriori keine Aussage über die *Surjektivität* der *Funktion* gemacht werden.
- vi) Der Begriff *Beschränktheit* lässt sich direkt von 1D auf nD übertragen.
- vii) Der Begriff *Monotonie* lässt sich nicht von 1D auf nD übertragen.

Beispiele:

- *Skalarfeld* in 2D:

$$\begin{aligned} f : \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x; y) &\mapsto f(x; y) := x \cdot y. \end{aligned} \tag{2.5}$$

Es gilt dann

$$f(2; 3) = 2 \cdot 3 = 6 \tag{2.6}$$

$$f(-3; 5) = -3 \cdot 5 = -15. \tag{2.7}$$

- *Abstand zum Ursprung* in nD:

$$n = 1 : r(x) := \sqrt{x^2} = |x| \tag{2.8}$$

$$n = 2 : r(x; y) := \sqrt{x^2 + y^2} \tag{2.9}$$

$$n = 3 : r(x; y; z) := \sqrt{x^2 + y^2 + z^2} \tag{2.10}$$

$$\text{allg: } r(x_1; x_2; \dots; x_n) := \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \tag{2.11}$$

- *Höhe* über Meer auf einer topographischen Karte:

$$\begin{aligned} H : [0, \pi] \times [0, 2\pi] &\rightarrow \mathbb{R}_0^+ [\text{m}] \\ (\theta; \varphi) &\mapsto H(\theta; \varphi) := \text{Höhe über Meer an der Position } (\theta; \varphi). \end{aligned} \tag{2.12}$$

- *Volumen* eines *Kreiskegels* als *Funktion* von *Radius* und *Höhe*:

$$\begin{aligned} V : \mathbb{R}_0^+ [\text{m}] \times \mathbb{R}_0^+ [\text{m}] &\rightarrow \mathbb{R}_0^+ [\text{m}^3] \\ (r; h) &\mapsto V(r; h) := \frac{1}{3} \cdot \pi \cdot r^2 \cdot h. \end{aligned} \tag{2.13}$$

- *Gesamtenergie* eines Projektils der *Masse m*:

$$\begin{aligned} E : \mathbb{R} \left[\frac{\text{m}}{\text{s}} \right] \times \mathbb{R} [\text{m}] &\rightarrow \mathbb{R} [\text{J}] \\ (v; h) &\mapsto E(v; h) := \frac{m}{2} \cdot v^2 + m \cdot g \cdot h. \end{aligned} \tag{2.14}$$

- *Zeitunabhängiges Temperaturfeld*:

$$\begin{aligned} T : \mathbb{R}^3 [\text{m}] &\rightarrow \mathbb{R}_0^+ [\text{K}] \\ (x; y; z) &\mapsto T(x; y; z) := \text{Temperatur am Ort } (x; y; z). \end{aligned} \tag{2.15}$$

- *Zeitabhängiges Temperaturfeld*:

$$\begin{aligned} T : \mathbb{R} [\text{s}] \times \mathbb{R}^3 [\text{m}] &\rightarrow \mathbb{R}_0^+ [\text{K}] \\ (t; x; y; z) &\mapsto T(t; x; y; z) := \text{Temperatur zur Zeit } t \text{ am Ort } (x; y; z). \end{aligned} \tag{2.16}$$

2.1.1.2 Level-Mengen

Wir betrachten die folgende Definition.

Definition 2.2 Level-Menge

Seien $n \in \mathbb{N}^+$, $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}$ und f eine reellwertige Funktion der Form $f : A \rightarrow B$. Die Level-Menge zu einem Level $L \in \mathbb{R}$ ist das Urbild

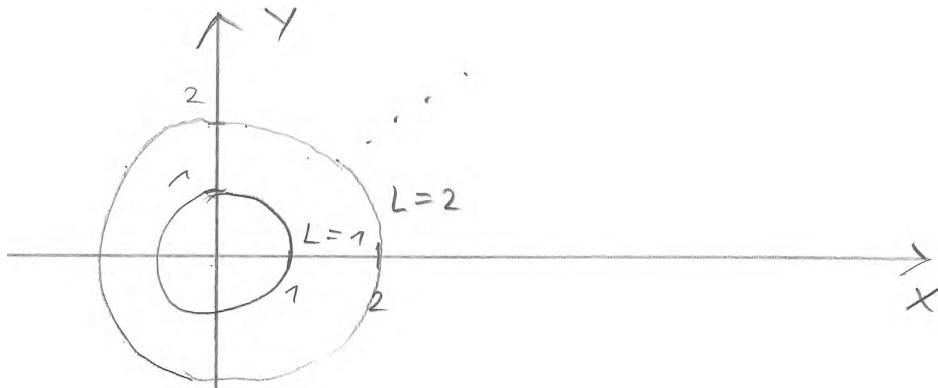
$$f^{-1}(\{L\}) = \{p \in A \mid f(p) = L\}. \quad (2.17)$$

Bemerkungen:

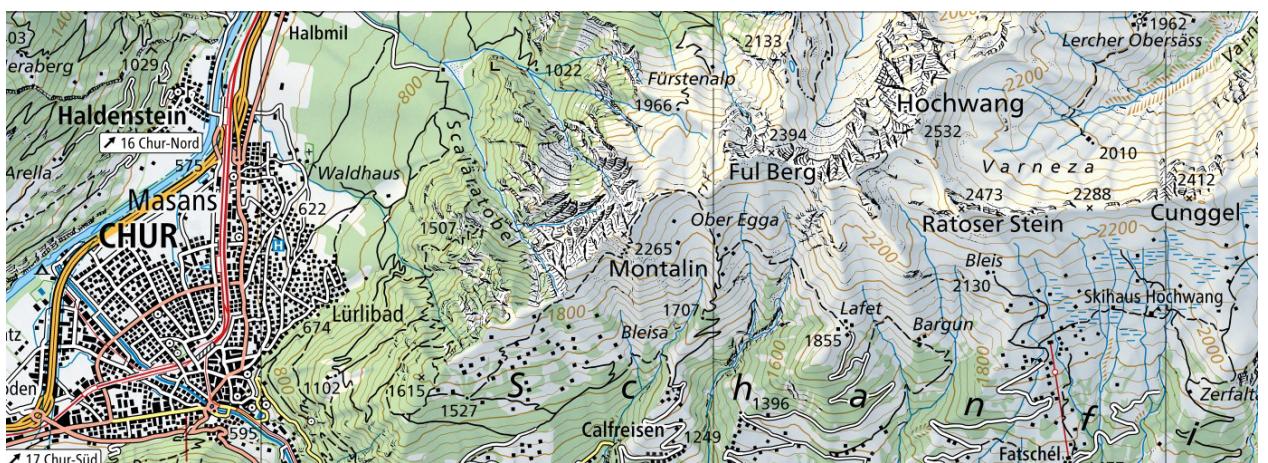
- i) Die Level-Menge zum Level L besteht aus allen Elementen der Definitionsmenge, an welchen die Funktion den Wert des Levels L hat.
- ii) Die Begriffe Level-Menge und Niveau-Menge sind synonym.
- iii) Für $n = 2$ sind die Level-Mengen meistens Kurven in der x - y -Ebene, welche als Level-Linien bzw. Niveau-Linien bezeichnet werden.
- iv) Für $n = 3$ sind die Level-Mengen meistens Flächen in 3D, welche als Level-Flächen bzw. Niveau-Flächen bezeichnet werden.
- v) Insbesondere für $n \in \{2, 3\}$ ist die Struktur der Level-Mengen sehr nützlich für die Visualisierung (Plot) von f .

Beispiele:

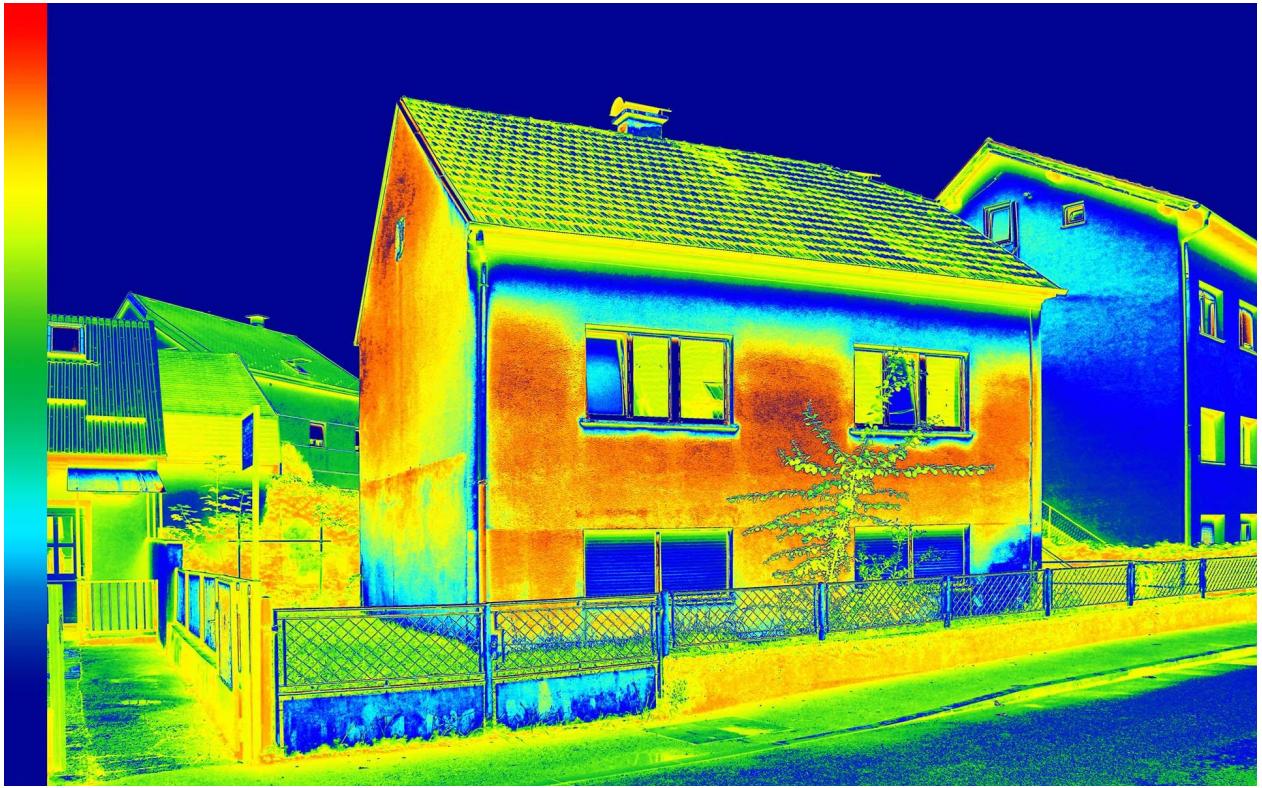
- Die Level-Mengen des Abstandes zum Ursprung sind Sphären mit Mittelpunkt am Ursprung.



- Die Level-Linien der Höhe über Meer sind die Höhenlinien auf einer topographischen Karte.



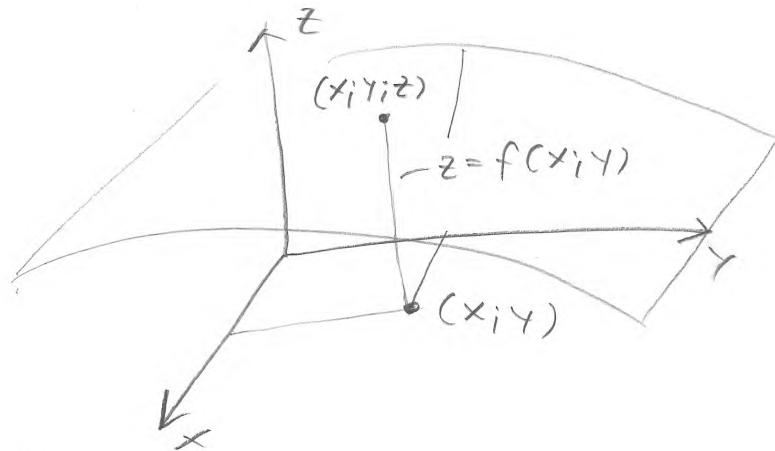
- Die *Level-Mengen* der *Temperatur* werden in der *Thermographie* durch eine Farbskala dargestellt.



2.1.1.3 Visualisierung

Es gibt mehrere Möglichkeiten, um eine *reellwertige Funktion* in mehreren *reellen Variablen* zu visualisieren.

- 1. Funktionsgraph:** Die *Funktionswerte* werden in eine weitere *Dimension* aufgetragen. Wir zeigen eine Beispieldarstellung für eine *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$.



Anwendungen: *Flächen-Plot*.

- 2. Farb-Plot:** Die *Funktionswerte* werden durch unterschiedliche *Farben* gekennzeichnet.
Anwendungen: Thematische Landkarten, *Thermographie*.

- 3. Level-Mengen-Plot:** Die *Level-Mengen* zu verschiedenen *Levels* werden gezeichnet.
Anwendungen: *Höhenlinien*, *Äquipotentiallinien*, *Äquipotentialflächen*, *Atomorbitale*.

Bemerkungen:

- i) Welche Visualisierung am besten geeignet ist, hängt sowohl von der *Dimension* als auch vom Verlauf der *Funktion* ab.
- ii) Nicht jede *Fläche* in 3D ist der *Graph* einer *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Gegenbeispiel: Alle geschlossenen *Flächen*, wie z.B. eine *Sphäre*.
- iii) Jede *Ebene* in 3D kann als *Graph* einer *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ beschrieben werden. Dies geschieht mit Hilfe einer *Normalen-Gleichung*.

2.1.2 Vektorfelder

2.1.2.1 Definition

Wir betrachten die folgende Definition.

Definition 2.3 *Vektorwertige Funktion* in mehreren *reellen Variablen*

Seien $n \in \mathbb{N}^+ \setminus \{1\}$, $A, B \subseteq \mathbb{R}^n$. Eine *Funktion* auf A der Form $\mathbf{v} : A \rightarrow B$ heisst *vektorwertige Funktion* in n *reellen Variablen*.

Bemerkungen:

- i) Die *reellen Variablen* werden nach dem *Funktionsnamen* \mathbf{v} in runden Klammern aufgezählt, jeweils durch ein Semicolon getrennt.

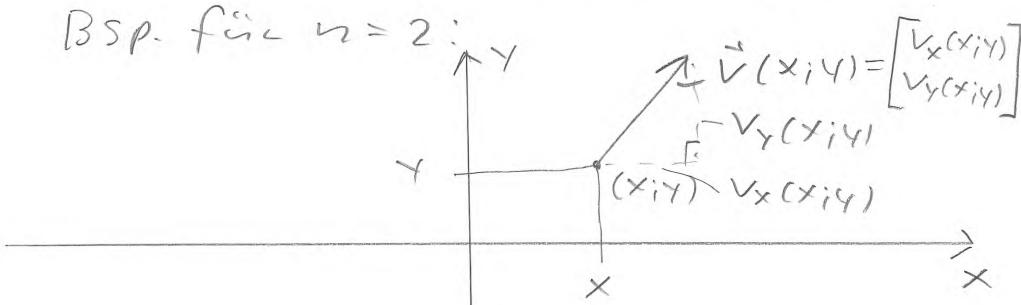
$$n = 2 : \quad \mathbf{v}(x; y) = \begin{bmatrix} v_x(x; y) \\ v_y(x; y) \end{bmatrix} \quad (2.18)$$

$$n = 3 : \quad \mathbf{v}(x; y; z) = \begin{bmatrix} v_x(x; y; z) \\ v_y(x; y; z) \\ v_z(x; y; z) \end{bmatrix} \quad (2.19)$$

$$\text{allg:} \quad \mathbf{v}(x_1; x_2; \dots; x_n) = \begin{bmatrix} v_1(x_1; x_2; \dots; x_n) \\ v_2(x_1; x_2; \dots; x_n) \\ \vdots \\ v_n(x_1; x_2; \dots; x_n) \end{bmatrix} \quad (2.20)$$

- ii) Jede *Komponente* eines *Vektorfeldes* ist eine *reellwertige Funktion* in n *reellen Variablen*.

- iii) Werden die *Definitionsmenge* A und die *Zielmenge* B als Teilmengen eines *geometrischen Raumes* interpretiert, dann wird \mathbf{v} als *Vektorfeld* auf A bezeichnet. In diesem Fall wird an jedem *Punkt* $(x_1; x_2; \dots; x_n) \in A$ ein *Vektor* $\mathbf{v}(x_1; x_2; \dots; x_n)$ "angehängt".



- iv) Im allgemeinen können die *reellen Variablen* $x_1, x_2, \dots, x_n \in \mathbb{R}$ und der *Funktionswert* $\mathbf{v}(x_1; x_2; \dots; x_n)$ beliebige und auch unterschiedliche Masseinheiten tragen.

- v) In der Physik werden oft *zeitabhängige Vektorfelder* verwendet, welche zusätzlich von der *Zeit* als *unabhängige Variable* abhängen.

Anwendungen:

- *Strömungsdynamik*: Die zentrale Grösse ist das *Geschwindigkeitsvektorfeld*

$$\mathbf{v}(t; x; y; z) \equiv \text{Geschwindigkeit des Mediums zur Zeit } t \text{ am Ort } (x; y; z). \quad (2.21)$$

- *Elektrodynamik*: Um *Dichte* und *Bewegung* von *elektrischen Ladungen* und die *Wechselwirkung* mit *elektrischen* und *magnetischen Feldern* zu beschreiben, werden mehrere im allgemeinen *zeitabhängige Skalarfelder* und *zeitabhängige Vektorfelder* verwendet.

$$\rho(t; x; y; z) := \text{Elektrische Ladungsdichte} \quad (2.22)$$

$$\mathbf{v}(t; x; y; z) := \text{Geschwindigkeitsfeld der elektrischen Ladungsträger} \quad (2.23)$$

$$\mathbf{J}(t; x; y; z) := \rho(t; x; y; z) \cdot \mathbf{v}(t; x; y; z) \equiv \text{Stromdichte-Vektorfeld} \quad (2.24)$$

$$\mathbf{E}(t; x; y; z) := \text{Elektrisches Feld} \quad (2.25)$$

$$\mathbf{B}(t; x; y; z) := \text{Magnetisches Feld} \equiv \text{Magnetische Flussdichte} \text{ (veraltet)} \quad (2.26)$$

$$\mathbf{P}(t; x; y; z) := \text{Polarisation} \quad (2.27)$$

$$\mathbf{M}(t; x; y; z) := \text{Magnetisierung} \quad (2.28)$$

$$\mathbf{D}(t; x; y; z) := \text{Dielektrische Verschiebung} \quad (2.29)$$

$$\mathbf{H}(t; x; y; z) := \text{Magnetisches Hilfsfeld} \equiv \text{Magnetfeld} \text{ (veraltet)} \quad (2.30)$$

$$\mathbf{S}(t; x; y; z) := \text{POYNTING- Vektorfeld} \quad (2.31)$$

$$\phi(t; x; y; z) := \text{Skalarpotential} \quad (2.32)$$

$$\mathbf{A}(t; x; y; z) := \text{Vektorpotential} \quad (2.33)$$

- *Differentialgleichungen ersten Grades*: Der Verlauf der verschiedenen *Lösungen* einer ODE ersten Grades kann durch das *Richtungsvektorfeld* $\hat{\mathbf{v}}(x; y)$ beschrieben und visualisiert werden.

2.1.2.2 Eigenschaften

Wir betrachten folgende Definition.

Definition 2.4 Einheitsvektorfeld

Seien $n \in \mathbb{N}^+ \setminus \{1\}$, $A, B \subseteq \mathbb{R}^n$, dann ist ein Vektorfeld $\mathbf{v} : A \rightarrow B$ ein *Einheitsvektorfeld*, falls für alle $p \in A$ gilt

$$\mathbf{v}(p) = \hat{\mathbf{v}}(p). \quad (2.34)$$

Bemerkungen:

- i) Bei einem *Einheitsvektorfeld* wird an jedem Punkt $p \in A$ ein *Einheitsvektor* “angehängt”.
- ii) Wenn ein *Vektorfeld* als *Einheitsvektorfeld* identifiziert ist, dann wird es in der Notation sinnvollerweise durch einen Hut gekennzeichnet.
- iii) Wie alle *Einheitsvektoren*, tragen auch die *Vektoren* eines *Einheitsvektorfeldes* keine Masseneinheit.
- iv) Die Bezeichnungen *Einheitsvektorfeld* und *Richtungsvektorfeld* sind synonym.

Wir betrachten folgende Definition.

Definition 2.5 Homogenes Vektorfeld

Seien $n \in \mathbb{N}^+ \setminus \{1\}$, $A, B \subseteq \mathbb{R}^n$, dann ist ein Vektorfeld $\mathbf{v} : A \rightarrow B$ *homogen*, falls es ein $\mathbf{w} \in B$ gibt, so dass für alle $p \in A$ gilt

$$\mathbf{v}(p) = \mathbf{w}. \quad (2.35)$$

Bemerkungen:

- i) Ein *homogenes Vektorfeld* ist einfach ein *konstantes Vektorfeld*.
- ii) Bei einem *homogenen Vektorfeld* wird an jedem Punkt $p \in A$ der gleiche Vektor \mathbf{w} “angehängt”.

2.1.2.3 Visualisierung

Siehe Demonstration von *Quiver-Plots*.

2.2 Parametrisierte Kurven & Linienintegrale

2.2.1 Parametrisierte Kurven

2.2.1.1 Definition

Wir betrachten die folgende Definition.

Definition 2.6 Parametrisierte Kurve

Seien $n \in \mathbb{N}^+$ und $\tau_0, \tau_E \in \mathbb{R}$ mit $\tau_0 < \tau_E$. Eine *parametrisierte Kurve* ist eine *differentierbare Funktion* der Form

$$\mathbf{s} : [\tau_0, \tau_E] \rightarrow \mathbb{R}^n$$

$$\tau \mapsto \mathbf{s}(\tau) := \begin{bmatrix} s_1(\tau) \\ s_2(\tau) \\ \vdots \\ s_n(\tau) \end{bmatrix}. \quad (2.36)$$

Analog zur *Vektor-Kinematik* in der *Physik* definiert man weitere Begriffe, um eine *parametrisierte Kurve* zu beschreiben. Wir betrachten dazu die folgende Definition.

Definition 2.7 Weitere Begriffe

Seien $n \in \mathbb{N}^+$, $\tau_0, \tau_E \in \mathbb{R}$ mit $\tau_0 < \tau_E$ und $\mathbf{s} : [\tau_0, \tau_E] \rightarrow \mathbb{R}^n$ eine *parametrisierte Kurve*.

(a) Geschwindigkeitsvektor:

$$\mathbf{v}(\tau) := \dot{\mathbf{s}}(\tau)$$

(d) Beschleunigungsvektor:

$$\mathbf{a}(\tau) := \ddot{\mathbf{s}}(\tau)$$

(b) Bahngeschwindigkeit:

$$v(\tau) := |\mathbf{v}(\tau)|$$

(e) Bahnbeschleunigung:

$$a_B(\tau) := \langle \mathbf{a}(\tau), \hat{\mathbf{e}}(\tau) \rangle$$

(c) Bahnvektor für $v(\tau) \neq 0$:

$$\hat{\mathbf{e}}(\tau) := \hat{\mathbf{v}}(\tau)$$

(f) Bahn:

$$B = \mathbf{s}([\tau_0, \tau_E])$$

Bemerkungen:

i) Der *Ortsvektor* $\mathbf{s}(\tau)$ zeigt für jedes τ vom *Ursprung* auf den entsprechenden *Punkt* der *Bahn* der *parametrisierten Kurve* in \mathbb{R}^n .

ii) Für die Masseinheiten erhalten wir

$$[\mathbf{v}] = [v] = \frac{[\mathbf{s}]}{[\tau]} \quad (2.37)$$

$$[\mathbf{a}] = [a_B] = \frac{[\mathbf{s}]}{[\tau]^2}. \quad (2.38)$$

iii) Der *Kurvenparameter* τ kann je nach *parametrisierter Kurve* ganz unterschiedliche *geometrische* oder *physikalische* Bedeutungen haben.

- *physikalische Zeit (Kinematik)*
- *Länge*
- *Winkel*
- eine beliebige “Absteckung” der *Bahn*

iv) Die *Bahn* ist die *Menge* aller von der *parametrisierten Kurve* durchlaufenen *Punkte*, d.h. sie ist die *Kurve* im klassischen *geometrischen* Sinn.

v) Der *Bahnvektor* ist der *Richtungsvektor (Einheitsvektor)* entlang der *Bahn*.

vi) Die *Parametrisierung* $s(\tau)$ muss nicht *injektiv* sein, aber soll so gewählt werden, dass gilt $v(\tau) > 0$. Ausnahmen an einzelnen *Punkten* sind jedoch möglich. An diesen muss der *Bahnvektor* \hat{e} von Hand definiert werden.

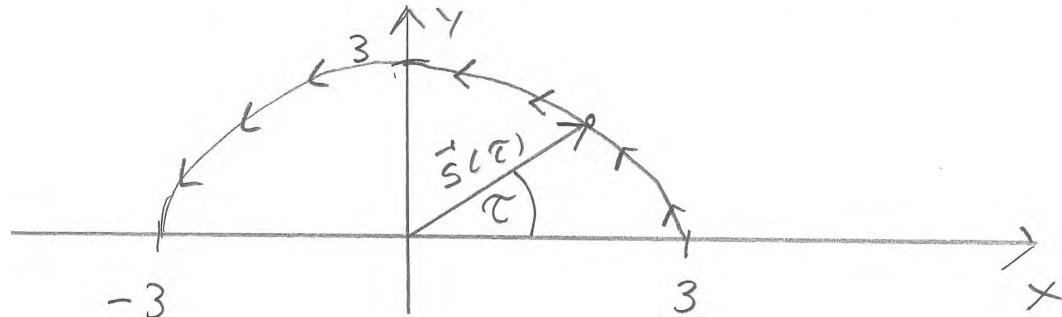
vii) Mehrere unterschiedliche *parametrisierte Kurven* können die gleiche *Bahn* haben.

Beispiele:

- Wir betrachten die *parametrisierte Kurve*

$$s(\tau) = \begin{bmatrix} 3 \cdot \cos(\tau) \\ 3 \cdot \sin(\tau) \end{bmatrix} \quad \text{für } \tau \in [0, \pi]. \quad (2.39)$$

Die *Bahn* dieser *parametrisierten Kurve* ist ein *Halbkreis* mit *Mittelpunkt* am *Ursprung* und *Radius* $r = 3$.



- Wir betrachten die *parametrisierte Kurve*

$$s(\tau) = \begin{bmatrix} 3 \cdot \cos(\pi \cdot \tau) \\ 3 \cdot \sin(\pi \cdot \tau) \end{bmatrix} \quad \text{für } \tau \in [0, 1]. \quad (2.40)$$

Die *Bahn* ist die gleiche wie beim ersten Beispiel.

- Wir betrachten die *parametrisierte Kurve*

$$s(\tau) = \begin{bmatrix} 3 \cdot \cos(\pi \cdot \tau^2) \\ 3 \cdot \sin(\pi \cdot \tau^2) \end{bmatrix} \quad \text{für } \tau \in [0, 1]. \quad (2.41)$$

Die *Bahn* ist die gleiche wie beim ersten Beispiel.

- Wir betrachten die *parametrisierte Kurve*

$$s(\tau) = \begin{bmatrix} 3 \\ 2 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \end{bmatrix} \cdot \tau \quad \text{für } \tau \in [0, 4]. \quad (2.42)$$

Die *Bahn* ist eine *gerade Strecke* zwischen den *Punkten* $(3; 2)$ und $(-5; 6)$.

2.2.1.2 Standard-Parametrisierungen

Die folgenden *Parametrisierungen* werden sehr häufig verwendet.

1. Für *Geraden* und *Strecken* in \mathbb{R}^n .

$$\mathbf{s}(\tau) = \mathbf{s}_0 + \mathbf{v} \cdot \tau \quad (2.43)$$

$$\mathbf{s}(\tau) = \mathbf{s}_0 + \mathbf{v} \cdot (\tau - \tau_0) \quad (2.44)$$

$$\mathbf{s}(\tau) = \mathbf{s}_0 + (\mathbf{s}_E - \mathbf{s}_0) \cdot \tau \quad \text{für } \tau \in [0, 1] \quad (2.45)$$

2. Für einen *Kreis* in \mathbb{R}^2 mit *Mittelpunkt* M und *Radius* $r > 0$.

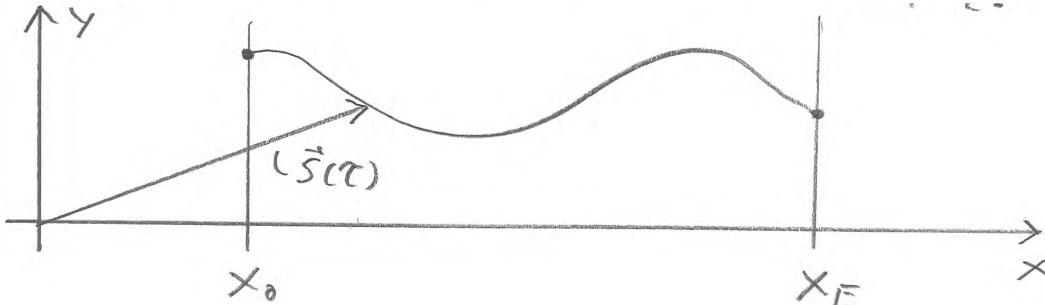
$$\mathbf{s}(\tau) = \mathbf{M} + \begin{bmatrix} r \cdot \cos(\tau) \\ r \cdot \sin(\tau) \end{bmatrix} \quad \text{für } \tau \in [0, 2\pi] \quad (2.46)$$

$$\mathbf{s}(\tau) = \mathbf{M} + \begin{bmatrix} r \cdot \cos(\omega \cdot \tau) \\ r \cdot \sin(\omega \cdot \tau) \end{bmatrix} \quad \text{für } \tau \in [0, T] \quad \text{mit } \omega = \frac{2\pi}{T} \quad (2.47)$$

3. Für den *Graph* einer *Funktion* $f : [x_0, x_E] \rightarrow \mathbb{R}$.

$$\mathbf{s}(\tau) = \begin{bmatrix} \tau \\ f(\tau) \end{bmatrix} \quad \text{für } \tau \in [x_0, x_E] \quad (2.48)$$

Die Situation ist im folgenden *x-y*-Diagramm dargestellt.



2.2.1.3 Bogenlänge

Wir betrachten die folgende Definition.

Definition 2.8 Bogenlänge

Seien $n \in \mathbb{N}^+$, $\tau_0, \tau_E \in \mathbb{R}$ mit $\tau_0 < \tau_E$ und $\mathbf{s} : [\tau_0, \tau_E] \rightarrow \mathbb{R}^n$ eine *parametrisierte Kurve* mit *Bahngeschwindigkeit* $v(\tau)$. Die *Bogenlänge* der *parametrisierten Kurve* ist die *reelle Zahl*

$$\Delta s := \int_{\tau_0}^{\tau_E} v(\tau) d\tau. \quad (2.49)$$

Bemerkungen:

- i) Für die Masseinheit erhalten wir

$$[\underline{\underline{\Delta s}}] = [v] \cdot [\tau] = [\mathbf{v}] \cdot [\tau] = \frac{[\mathbf{s}]}{[\tau]} \cdot [\tau] = [\underline{\underline{\mathbf{s}}}] \quad (2.50)$$

- ii) Man kann zeigen, dass die *Bogenlänge* Δs nicht von der Wahl der *Parametrisierung* sondern nur von der *Bahn* abhängt (sofern mehrfache Durchläufe auch mehrfach gerechnet werden).
- iii) In 2D und 3D stimmt die *Bogenlänge* mit dem klassischen Begriff der *Weglänge* aus der *Elementargeometrie* überein.
- iv) Wählt man für den *Kurvenparameter* die *geometrische Weglänge* $\tau = s$ entlang der *Bahn* ("Kilometrierung"), dann gilt

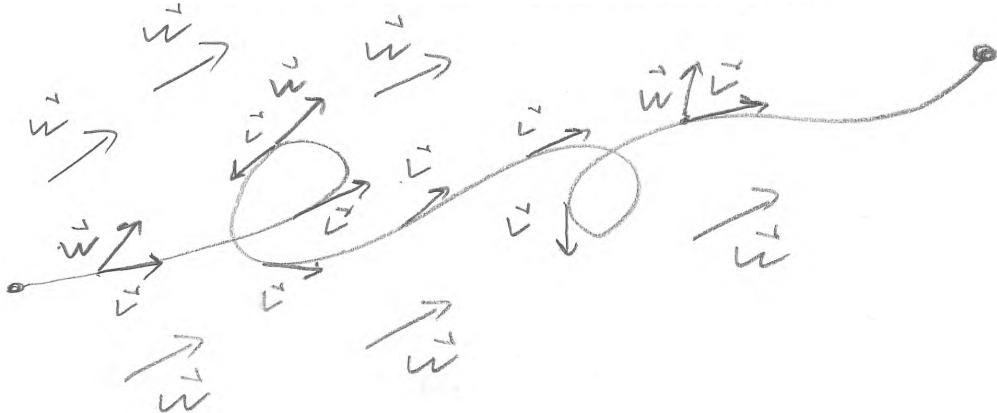
$$v(s) = 1 \quad (2.51)$$

$$\underline{\Delta s} = \int_{\tau_0}^{\tau_E} v(\tau) d\tau = \int_{s_0}^{s_E} v(s) ds = \int_{s_0}^{s_E} 1 ds = \underline{s_E - s_0}. \quad (2.52)$$

- v) In der Praxis lässt sich das *Integral* in (2.49) nur in einfachen Fällen *analytisch* berechnen. Meistens muss auf *numerische Integration* zurückgegriffen werden.

2.2.2 Linienintegrale

Wir betrachten eine *parametrisierte Kurve* im Bereich eines *Vektorfeldes*. Die Situation ist in der folgenden Skizze dargestellt.



Wir betrachten die folgende Definition.

Definition 2.9 Linienintegral

Seien $n \in \mathbb{N}^+$, $\tau_0, \tau_E \in \mathbb{R}$ mit $\tau_0 < \tau_E$, $\mathbf{s} : [\tau_0, \tau_E] \rightarrow \mathbb{R}^n$ eine *parametrisierte Kurve* mit *Geschwindigkeitsvektor* $\mathbf{v}(\tau)$ und $\mathbf{w} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein *Vektorfeld*. Das *Linienintegral* des *Vektorfeldes* \mathbf{w} entlang der *Kurve* $\mathbf{s}(\tau)$ ist die *reelle Zahl*

$$I := \boxed{\int_{\tau_0}^{\tau_E} \langle \mathbf{w}, \mathbf{v} \rangle d\tau}. \quad (2.53)$$

Bemerkungen:

- i) Die Begriffe *Linienintegral* und *Kurvenintegral* sind synonym.
- ii) Für die Masseinheit erhalten wir

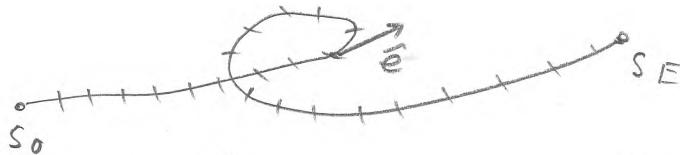
$$\underline{\underline{[I]}} = [\langle \mathbf{w}, \mathbf{v} \rangle] \cdot [\tau] = [\mathbf{w}] \cdot [\mathbf{v}] \cdot [\tau] = [\mathbf{w}] \cdot \frac{[\mathbf{s}]}{[\tau]} \cdot [\tau] = \underline{\underline{[\mathbf{w}] \cdot [\mathbf{s}]}}, \quad (2.54)$$

iii) Man kann zeigen, dass das *Linienintegral* I bis auf das *Vorzeichen* nicht von der Wahl der *Parametrisierung* sondern nur von der *Bahn* abhängt (sofern mehrfache Durchläufe auch mehrfach gerechnet werden).

iv) Wählt man für den *Kurvenparameter* die *geometrische Weglänge* $\tau = s$ entlang der *Bahn* (“Kilometrierung”), dann gilt

$$\mathbf{v}(s) = v(s) \cdot \hat{\mathbf{e}}(s) = 1 \cdot \hat{\mathbf{e}}(s) = \hat{\mathbf{e}}(s). \quad (2.55)$$

Die Situation ist in der folgenden Skizze dargestellt.



Somit folgt

$$\underline{I} = \int_{\tau_0}^{\tau_E} \langle \mathbf{w}, \mathbf{v} \rangle d\tau = \underline{\int_{s_0}^{s_E} \langle \mathbf{w}, \hat{\mathbf{e}} \rangle ds}. \quad (2.56)$$

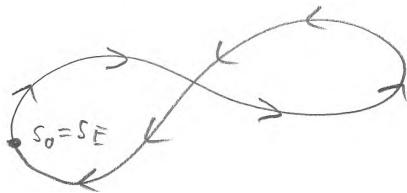
v) Um ein *Linienintegral* auszurechnen, müssen die *Koordinaten* der *Punkte* entlang der *parametrisierten Kurve* im *Vektorfeld* eingesetzt werden. Vollständig ausgeschrieben mit allen Abhängigkeiten ergibt dies

$$I = \int_{\tau_0}^{\tau_E} \langle \mathbf{w}(\mathbf{s}(\tau)), \mathbf{v}(\tau) \rangle d\tau = \int_{s_0}^{s_E} \langle \mathbf{w}(\mathbf{s}(s)), \hat{\mathbf{e}}(s) \rangle ds. \quad (2.57)$$

vi) In der Literatur findet man für *Linienintegrale* entlang einer *Kurve* γ die Schreibweisen

$$I = \int_{\tau_0}^{\tau_E} \langle \mathbf{w}, \mathbf{v} \rangle d\tau = \int_{s_0}^{s_E} \langle \mathbf{w}, \hat{\mathbf{e}} \rangle ds = \int_{\gamma} \langle \mathbf{w}, \hat{\mathbf{e}} \rangle ds = \int_{\gamma} \mathbf{w} \cdot ds. \quad (2.58)$$

vii) Ein *Linienintegral* über eine *geschlossene Kurve* heisst *Zirkulation* des *Vektorfeldes* entlang der betreffenden *Kurve*.



In diesem Fall verwendet man das *Ring-Integralzeichen*, d.h.

$$\Upsilon = \oint_{\tau_0}^{\tau_E} \langle \mathbf{w}, \mathbf{v} \rangle d\tau = \oint_{s_0}^{s_E} \langle \mathbf{w}, \hat{\mathbf{e}} \rangle ds = \oint_{\gamma} \mathbf{w} \cdot ds. \quad (2.59)$$

viii) Das *Linienintegral* eines *Vektorfeldes* \mathbf{w} entlang einer *Kurve* ist das *Integral* des Anteils von \mathbf{w} , welcher entlang $\hat{\mathbf{e}}$ und somit entlang der *Kurve* zeigt. Der zu $\hat{\mathbf{e}}$ bzw. zur *Kurve* senkrechte Anteil von \mathbf{w} ist für das *Linienintegral* irrelevant.

Anwendungen:

- Mechanik: *Arbeitsintegral* gemäss

$$\Delta W = \int_{s_0}^{s_E} \langle \mathbf{F}(s), \hat{\mathbf{e}} \rangle \, ds. \quad (2.60)$$

- Elektrodynamik: *Elektrische Spannung* gemäss

$$U = \int_{s_0}^{s_E} \langle \mathbf{E}, \hat{\mathbf{e}} \rangle \, ds. \quad (2.61)$$

Wir betrachten den folgenden Satz.

Satz 2.1 *Linienintegral bei konstantem Anteil entlang der Bahn*

Seien $n \in \mathbb{N}^+$, $\tau_0, \tau_E \in \mathbb{R}$ mit $\tau_0 < \tau_E$, $\mathbf{s} : [\tau_0, \tau_E] \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve mit Bahnvektor $\hat{\mathbf{e}}(\tau)$ und Bogenlänge Δs sowie $\mathbf{w} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Vektorfeld. Gilt entlang der Bahn der Kurve $\langle \mathbf{w}, \hat{\mathbf{e}} \rangle =: C \equiv \text{konst.}$, dann beträgt das Linienintegral von \mathbf{w} entlang der Kurve

$$I = C \cdot \Delta s. \quad (2.62)$$

Beweis: Für das Linienintegral von \mathbf{w} entlang der Kurve erhalten wir

$$\begin{aligned} I &= \int_{\tau_0}^{\tau_E} \langle \mathbf{w}, \mathbf{v} \rangle \, d\tau = \int_{s_0}^{s_E} \langle \mathbf{w}, \hat{\mathbf{e}} \rangle \, ds = \int_{s_0}^{s_E} C \, ds = C \int_{s_0}^{s_E} 1 \, ds = C \cdot [s] \Big|_{s_0}^{s_E} = C \cdot (s_E - s_0) \\ &= \underline{C \cdot \Delta s}. \end{aligned} \quad (2.63)$$

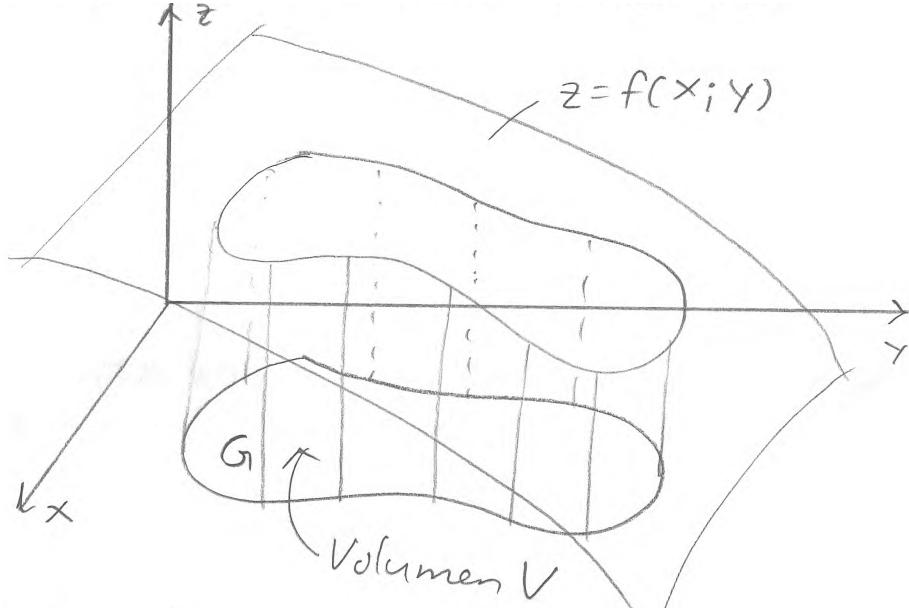
Damit haben wir den Satz bewiesen. □

2.3 Mehrfach-Integrale

2.3.1 Zweifach-Integrale

2.3.1.1 Definition

Wir betrachten ein *Gebiet* $G \subset \mathbb{R}^2$ und eine *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Der *Graph* von f ist im folgenden x - y - z -Diagramm dargestellt.



Wir betrachten die folgende Definition.

Definition 2.10 Integration über ein Gebiet

Seien $G \subset \mathbb{R}^2$ ein *Gebiet* und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine *integrierbare Funktion*. Das *Integral* von f über das *Gebiet* G ist

$$\boxed{\int_G f \, dA := \text{Volumen zwischen } G \text{ und dem Graphen von } f.} \quad (2.64)$$

Bemerkungen:

i) Wie in 1D lassen sich präzise Bedingungen definieren, die an eine *Teilmenge* $G \subset \mathbb{R}^2$ und an eine *Funktion* f gestellt werden müssen, damit das *Integral* existiert. Ohne im Detail auf diese Voraussetzungen einzugehen, bezeichnen wir eine *Teilmenge* $G \subset \mathbb{R}^2$ als *Gebiet* und eine *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ als *integrierbar*, falls alle notwendigen Kriterien erfüllt sind.

ii) Für die Masseinheit erhalten wir

$$\left[\int_G f \, dA \right] = [V] = [A] \cdot [f] = [x] \cdot [y] \cdot [f] = [x] \cdot [y] \cdot [z]. \quad (2.65)$$

iii) Ein *Integral* über ein *Gebiet* in 2D wird synonym auch als *Zweifach-Integral* oder *Doppel-Integral* bezeichnet.

iv) In der Literatur findet man auch die alternativen Schreibweisen

$$\begin{aligned}
 I &= \int_G f \, dA = \int_G f \, dG = \int_G f \, d\mathbb{R}^2 = \int_G f \, dx^2 = \int_G f \, d^2x = \int_G f \, d\mu \\
 &= \iint_G f \, dA = \iint_G f \, dG = \iint_G f \, d\mathbb{R}^2 = \iint_G f \, dx^2 = \iint_G f \, d^2x = \iint_G f \, d\mu \\
 &= \int_G f(x; y) \, dx \, dy = \iint_G f(x; y) \, dx \, dy.
 \end{aligned} \tag{2.66}$$

v) *Integrale über Gebiete* in 2D erhält man aus einem ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess in 2D gemäss folgenden Schritten.

S1 Lokal: Der Beitrag zur Grösse I eines kleinen Flächenstücks δA auf der x - y -Ebene ist

$$\underline{\delta I} \approx \dots \approx \underline{f(x; y) \cdot \delta A}. \tag{2.67}$$

S2 Global: Durch *Integration* über das *Gebiet* G können wir die gesamte Grösse I berechnen. Wir erhalten

$$\underline{\underline{I}} = \int_G f \, dA = \underline{\dots} \tag{2.68}$$

2.3.1.2 Elementare Rechenregeln

Wie in 1D ist die *Linearität* die wichtigste Eigenschaft des *Integrals*. Wir betrachten dazu den folgenden Satz.

Satz 2.2 *Linearität* des *Integrals* in 2D

Seien $G \subset \mathbb{R}^2$ ein *Gebiet*, $g, h : \mathbb{R}^2 \rightarrow \mathbb{R}$ integrierbare Funktionen und $a, b \in \mathbb{R}$. Dann gelten die folgenden Rechenregeln.

(a) Faktor-Regel:

$$\boxed{\int_G a \cdot g \, dA = a \int_G g \, dA.} \tag{2.69}$$

(b) Summen-Regel:

$$\boxed{\int_G (g + h) \, dA = \int_G g \, dA + \int_G h \, dA.} \tag{2.70}$$

(c) Linearität:

$$\boxed{\int_G (a \cdot g + b \cdot h) \, dA = a \int_G g \, dA + b \int_G h \, dA.} \tag{2.71}$$

Wie in 1D kann auch in 2D ein *Integral* in mehrere zerlegt werden. Dazu betrachten wir den folgenden Satz.

Satz 2.3 Zerlegungssatz in 2D

Seien $G, H \subset \mathbb{R}^2$ Gebiete und $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine integrierbare Funktion, dann gilt

$$\int_{G \cup H} f \, dA = \int_G f \, dA + \int_H f \, dA - \int_{G \cap H} f \, dA. \quad (2.72)$$

Wir betrachten den folgenden Satz.

Satz 2.4 Flächensatz

Sei $G \subset \mathbb{R}^2$ ein Gebiet mit Flächeninhalt $A > 0$, dann gilt

$$A = \int_G 1 \, dA. \quad (2.73)$$

Bemerkungen:

i) Im Spezialfall, dass der Integrand den konstanten Wert 1 hat, ist das Volumen zwischen dem Gebiet G in der x - y -Ebene und dem Graph des Integranden gerade der Flächeninhalt von G .

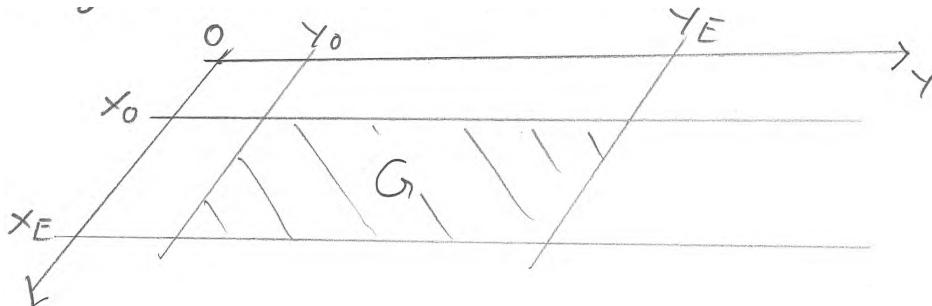
ii) Man prüft leicht nach, dass der berechnete Flächeninhalt die korrekte Maßeinheit hat. Es gilt

$$\underline{\underline{[A]}} = \left[\int_G 1 \, dA \right] = [x] \cdot [y] \cdot [1] = [x] \cdot [y] \cdot 1 = \underline{\underline{[x] \cdot [y]}}. \quad (2.74)$$

iii) Der Flächensatz stellt eine fundamentale Verbindung her zwischen dem Begriff Integral aus der Analysis und dem Begriff Flächeninhalt aus der Geometrie.

2.3.1.3 Integration über Rechtecke

Ein besonders einfacher Fall liegt vor, wenn das Gebiet G ein achsenparalleles Rechteck in der x - y -Ebene ist.



Wir betrachten dazu den folgenden Satz.

Satz 2.5 FUBINI-Satz für Rechtecke

Seien $x_0, x_E, y_0, y_E \in \mathbb{R}$ mit $x_0 < x_E$ und $y_0 < y_E$, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine integrierbare Funktion sowie G das Rechteck

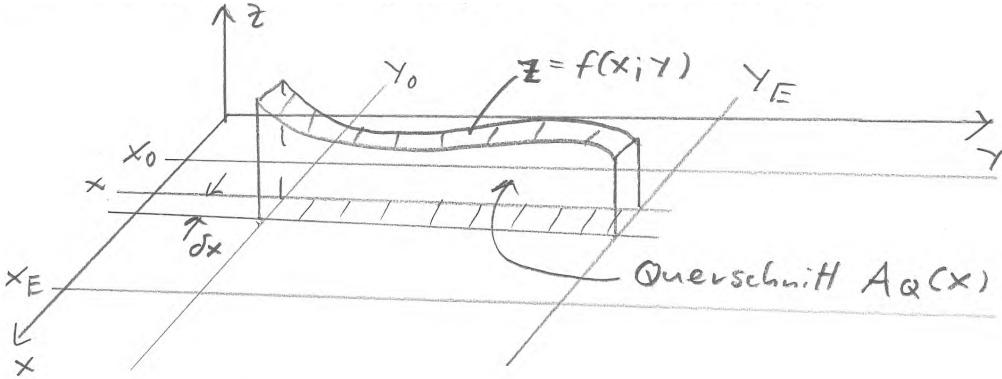
$$G := [x_0, x_E] \times [y_0, y_E]. \quad (2.75)$$

Dann gilt

$$\int_G f \, dA = \int_{y_0}^{y_E} \int_{x_0}^{x_E} f(x; y) \, dx \, dy = \int_{x_0}^{x_E} \int_{y_0}^{y_E} f(x; y) \, dy \, dx. \quad (2.76)$$

Beweis: Um das *Volumen* zwischen dem *Rechteck* G und dem *Graphen* von f zu berechnen verwenden wir einen **ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess**. Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten einen kleinen Streifen des *Rechtecks* G mit Breite $\delta x > 0$, wie im folgenden x - y - z -Diagramm dargestellt.



Die *Querschnittsfläche* $A_Q(x)$ können wir berechnen durch das *bestimmte Integral*

$$A_Q(x) = \int_{y_0}^{y_E} f(x; y) dy. \quad (2.77)$$

Das *Volumen* zwischen dem kleinen Streifen und dem *Graphen* von f ist

$$\underline{\delta I} \approx \underline{A_Q(x)} \cdot \delta x. \quad (2.78)$$

S2 Global: Durch *Integration* über x können wir das gesamte *Volumen* zwischen dem *Rechteck* G und dem *Graphen* von f berechnen. Wir erhalten

$$\underline{\underline{I}} = \int_{x_0}^{x_E} A_Q(x) dx = \int_{x_0}^{x_E} \underline{\int_{y_0}^{y_E} f(x; y) dy} dx. \quad (2.79)$$

Durch Vertauschen der Rollen von x und y erhalten wir auf analoge Weise die zweite Version. Damit haben wir den Satz bewiesen. \square

Bemerkungen:

- i) Durch den *FUBINI-Satz* kann ein *Integral* über ein *Gebiet* in 2D auf die Berechnung von 2 verschachtelten *Integralen* in 1D zurückgeführt werden und umgekehrt.
- ii) Bei der Anwendung des *FUBINI-Satzes* werden die *Zweifach-Integrale* in der Reihenfolge von innen nach aussen berechnet.
- iii) Bei der *Integration* über ein *Rechteck* kann die *Integrationsreihenfolge* vertauscht werden.
- iv) Der *FUBINI-Satz* ist ein allgemein gültiges Prinzip, dessen Aussage sowohl auf allgemeinere *Gebiete* als auch auf *Gebiete* in nD erweitert werden kann.

Je nach Struktur des *Integranden* kann die *Integration* über ein *Rechteck* weiter vereinfacht werden. Wir betrachten dazu den folgenden Satz.

Satz 2.6 Spezialfälle der *Integration* über ein *Rechteck*.

Seien $x_0, x_E, y_0, y_E \in \mathbb{R}$ mit $x_0 < x_E$ und $y_0 < y_E$ sowie G das *Rechteck*

$$G := [x_0, x_E] \times [y_0, y_E]. \quad (2.80)$$

Dann gilt folgendes.

(a) Konstanten-Regel: Für alle $C \in \mathbb{R}$ ist

$$\boxed{\int_G C \, dA = C \cdot (x_E - x_0) \cdot (y_E - y_0).} \quad (2.81)$$

(b) Separation-Regel: Für zwei integrierbare Funktionen $g, h : \mathbb{R} \rightarrow \mathbb{R}$ ist

$$\boxed{\int_G g(x) \cdot h(y) \, dA = \int_{x_0}^{x_E} g(x) \, dx \cdot \int_{y_0}^{y_E} h(y) \, dy.} \quad (2.82)$$

Beweis: Gemäss FUBINI-Satz und Faktor-Regel gilt

$$\begin{aligned} \underline{\underline{\int_G g(x) \cdot h(y) \, dA}} &= \int_{y_0}^{y_E} \int_{x_0}^{x_E} g(x) \cdot h(y) \, dx \, dy = \int_{y_0}^{y_E} h(y) \cdot \int_{x_0}^{x_E} g(x) \, dx \, dy \\ &= \underline{\underline{\int_{x_0}^{x_E} g(x) \, dx \cdot \int_{y_0}^{y_E} h(y) \, dy.}} \end{aligned} \quad (2.83)$$

Wir zeigen mehrere Varianten, um die *Konstanten-Regel* zu beweisen.

Variante 1: Gemäss FUBINI-Satz und Faktor-Regel gilt

$$\begin{aligned} \underline{\underline{\int_G C \, dA}} &= \int_{y_0}^{y_E} \int_{x_0}^{x_E} C \, dx \, dy = C \int_{y_0}^{y_E} \int_{x_0}^{x_E} 1 \, dx \, dy = C \int_{y_0}^{y_E} \left[x \right] \Big|_{x_0}^{x_E} \, dy = C \int_{y_0}^{y_E} (x_E - x_0) \, dy \\ &= C \cdot (x_E - x_0) \int_{y_0}^{y_E} 1 \, dy = C \cdot (x_E - x_0) \cdot \left[y \right] \Big|_{y_0}^{y_E} \\ &= \underline{\underline{C \cdot (x_E - x_0) \cdot (y_E - y_0)}}. \end{aligned} \quad (2.84)$$

Variante 2: Gemäss Faktor-Regel und Separation-Regel gilt

$$\begin{aligned} \underline{\underline{\int_G C \, dA}} &= C \int_G 1 \, dA = C \int_G 1 \cdot 1 \, dA = C \int_{x_0}^{x_E} 1 \, dx \cdot \int_{y_0}^{y_E} 1 \, dy = C \cdot \left[x \right] \Big|_{x_0}^{x_E} \cdot \left[y \right] \Big|_{y_0}^{y_E} \\ &= \underline{\underline{C \cdot (x_E - x_0) \cdot (y_E - y_0)}}. \end{aligned} \quad (2.85)$$

Variante 3: Gemäss Faktor-Regel und Flächensatz gilt

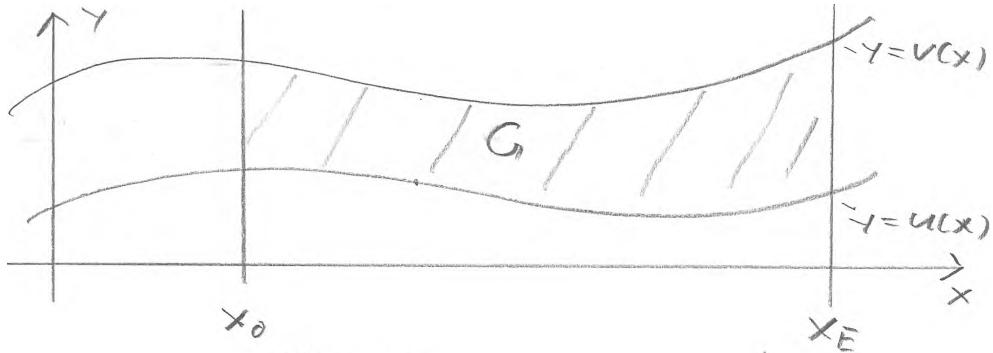
$$\underline{\underline{\int_G C \, dA}} = C \int_G 1 \, dA = C \cdot A = C \cdot \Delta x \cdot \Delta y = \underline{\underline{C \cdot (x_E - x_0) \cdot (y_E - y_0)}}. \quad (2.86)$$

Damit haben wir den Satz bewiesen. □

2.3.1.4 Integration über allgemeine Gebiete

Mit Hilfe des FUBINI-Satzes können auch *Integrale* über allgemeinere *Gebiete* $G \subset \mathbb{R}$ berechnet. Wir zeigen im folgenden eine Auswahl von besonders häufig auftretenden Fällen.

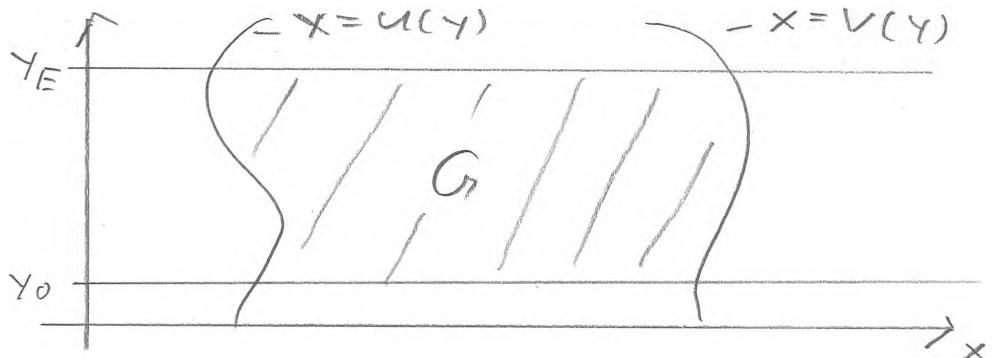
1. Wir betrachten ein in y -Richtung durch *Graphen* begrenztes *Gebiet*, wie im folgendem x - y -Diagramm dargestellt.



Gemäss FUBINI-Satz erhalten wir in diesem Fall

$$\int_G f \, dA = \int_{x_0}^{x_E} \int_{u(x)}^{v(x)} f(x; y) \, dy \, dx. \quad (2.87)$$

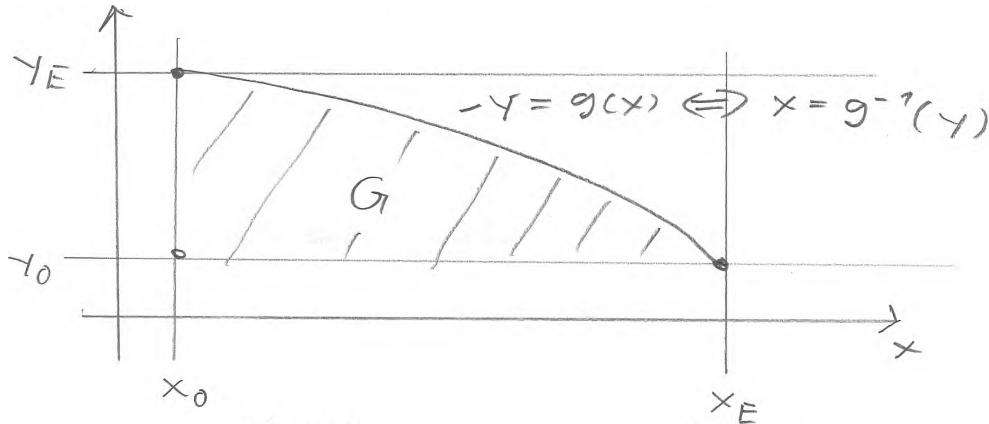
2. Wir betrachten ein in x -Richtung durch *Graphen* begrenztes *Gebiet*, wie im folgendem x - y -Diagramm dargestellt.



Gemäss FUBINI-Satz erhalten wir in diesem Fall

$$\int_G f \, dA = \int_{y_0}^{y_E} \int_{u(y)}^{v(y)} f(x; y) \, dx \, dy. \quad (2.88)$$

3. Wir betrachten ein *dreieckartiges Gebiet*, wie im folgendem x - y -Diagramm dargestellt.



Gemäss FUBINI-Satz erhalten wir in diesem Fall

$$\int_G f \, dA = \int_{x_0}^{x_E} \int_{y_0}^{g(x)} f(x; y) \, dy \, dx = \int_{y_0}^{y_E} \int_{x_0}^{g^{-1}(y)} f(x; y) \, dx \, dy. \quad (2.89)$$

4. Ein komplizierteres Gebiet kann meistens in einfachere *Gebiete* zerlegt und das *Integral* mit Hilfe des *Zerlegungssatzes* berechnet werden.

Bemerkungen:

- i) Bei allen Anwendungen des FUBINI-Satzes werden die *Zweifach-Integrale* in der Reihenfolge von innen nach aussen berechnet. Dabei dürfen die *Grenzen* des *inneren Integrals* von der *Integrationsvariablen* des *äußeren Integrals* abhängen.
- ii) Bei der *Integration* über ein *dreieckartiges Gebiet* kann die *Integrationsreihenfolge* auf einfache Weise vertauscht werden.
- iii) Es gibt eine Analogie zwischen den *Grenzen* eines *Zweifach-Integrals* und den Bewegungen des Stifts in einem Plotter.



Dabei kann man sich vorstellen, dass der Stift des Plotters das *Gebiet* auf dem Papier komplett "überstreichen" muss.

<i>Inneres Integral</i>	schnelle hin und her Bewegungen des Stifts entlang des Balkens
<i>Äußeres Integral</i>	langsame Bewegung des Balkens entlang des Papiers

Beim Vertauschen der *Integrationsreihenfolge* wird das Papier um einen *rechten Winkel* gedreht.

2.3.2 Dreifach-Integrale

Die Theorie der *Integrale* in 3D ist ganz analog zur Theorie in 2D. Exemplarisch betrachten wir dazu den folgenden Satz.

Satz 2.7 FUBINI-Satz für Quader

Seien $x_0, x_E, y_0, y_E, z_0, z_E \in \mathbb{R}$ mit $x_0 < x_E$, $y_0 < y_E$ und $z_0 < z_E$, $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ eine integrierbare Funktion sowie Q der Quader

$$Q := [x_0, x_E] \times [y_0, y_E] \times [z_0, z_E]. \quad (2.90)$$

Dann gilt

$$\int_Q f \, dV = \int_{z_0}^{z_E} \int_{y_0}^{y_E} \int_{x_0}^{x_E} f(x; y; z) \, dx \, dy \, dz. \quad (2.91)$$

Bemerkungen:

- i) Durch den FUBINI-Satz kann ein *Integral* über ein *Gebiet* in 3D auf die Berechnung von 3 verschachtelten *Integralen* in 1D zurückgeführt werden und umgekehrt.
- ii) Bei der Anwendung des FUBINI-Satzes werden die *Dreifach-Integrale* in der Reihenfolge von innen nach aussen berechnet.
- iii) Bei der *Integration* über einen *Quader* kann die *Integrationsreihenfolge* vertauscht werden.

Integrale über *Gebiete* in 3D erhält man aus einem ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess in 3D gemäss folgenden Schritten.

S1 Lokal: Der Beitrag zur Grösse I eines kleinen *Volumenstücks* δV im x - y - z -Raum ist

$$\underline{\delta I} \approx \dots \approx \underline{f(x; y; z)} \cdot \underline{\delta V}. \quad (2.92)$$

S2 Global: Durch *Integration* über das *Gebiet* G können wir die gesamte Grösse I berechnen. Wir erhalten

$$\underline{\underline{I}} = \int_G f \, dV = \underline{\underline{\dots}}. \quad (2.93)$$

Anwendungen:

- *Masse:* Wir betrachten einen Körpers K mit *Dichte* $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$.

S1 Lokal: Die *Masse* eines kleinen *Volumenstücks* δV im x - y - z -Raum ist

$$\underline{\delta m} \approx \rho \cdot \underline{\delta V} = \underline{\rho(x; y; z)} \cdot \underline{\delta V}. \quad (2.94)$$

S2 Global: Durch *Integration* über den *Körper* K können wir seine gesamte *Masse* berechnen. Wir erhalten

$$\underline{\underline{m}} = \int_K \rho \, dV. \quad (2.95)$$

- **Trägheitsmoment:** Wir betrachten einen Körpers K mit *Dichte* $\rho : \mathbb{R}^3 \rightarrow \mathbb{R}$. Der *Abstand* eines *Punktes* von einer vorgegebenen *Drehachse* sei $r : \mathbb{R}^3 \rightarrow \mathbb{R}$.

S1 Lokal: Das *Trägheitsmoment* eines kleinen *Volumenstücks* δV im x - y - z -Raum ist

$$\underline{\delta I} \approx r^2 \cdot \delta m \approx r^2 \cdot \rho \cdot \delta V = \underline{r^2(x; y; z) \cdot \rho(x; y; z) \cdot \delta V}. \quad (2.96)$$

S2 Global: Durch *Integration* über den *Körper* K können wir sein gesamtes *Trägheitsmoment* berechnen. Wir erhalten

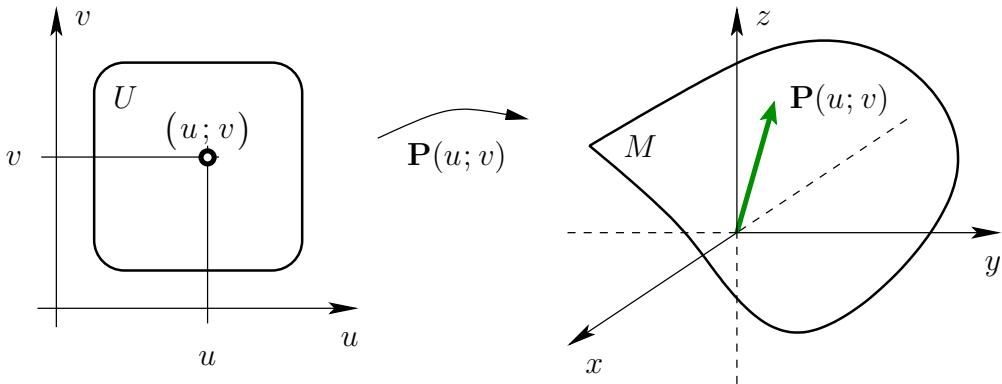
$$I = \underline{\int_K r^2 \cdot \rho \, dV}. \quad (2.97)$$

2.4 Parametrisierte Flächen & Flussintegrale

2.4.1 Parametrisierte Flächen

2.4.1.1 Definition

Wir betrachten eine allgemeine *Fläche* $M \subset \mathbb{R}^3$, welche durch zwei *Parameter* bzw. *Koordinaten* u und v aus einem *Gebiet* $U \subseteq \mathbb{R}^2$ parametrisiert wird. Die Situation ist in der folgenden Skizze dargestellt.



Definition 2.11 Parametrisierte Fläche

Sei $U \subseteq \mathbb{R}^2$ ein *Gebiet* und $M \subset \mathbb{R}^3$ ein *Fläche*. Eine *Parametrisierung* von M ist eine *Funktion* der Form

$$\boxed{\begin{aligned} \mathbf{P} : U &\rightarrow \mathbb{R}^3 \\ (u; v) &\mapsto \mathbf{P}(u; v) = \begin{bmatrix} x(u; v) \\ y(u; v) \\ z(u; v) \end{bmatrix}, \end{aligned}} \quad (2.98)$$

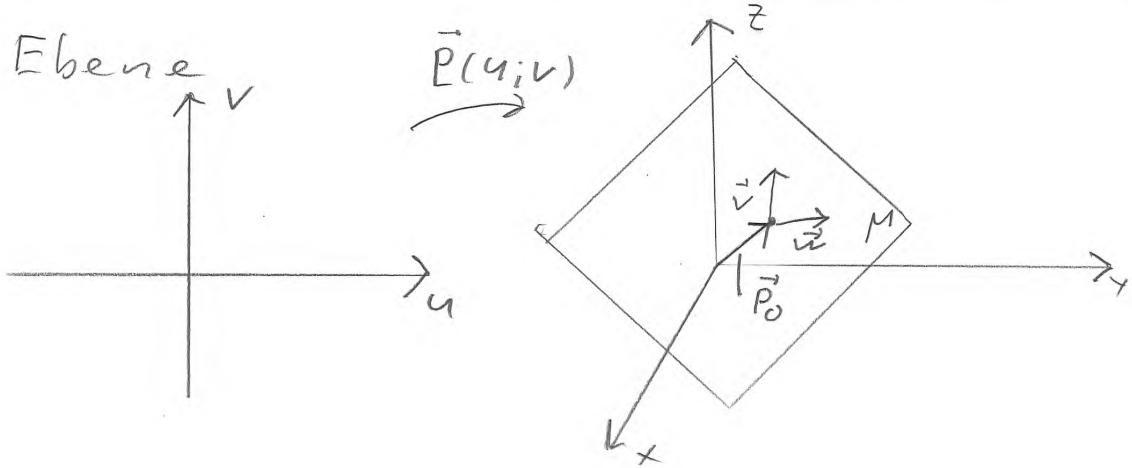
so dass $M = \mathbf{P}(U)$.

Bemerkungen:

- i) In Anlehnung an die Geographie werden U *Karten-Gebiet* oder *Karte* und die *Parameter* u und v auch *Koordinaten* genannt.
- ii) Erfahrungsgemäss können viele Rechnungen im Umgang mit der *Parametrisierung* einfach gehalten werden, wenn man, sofern möglich, für das *Karten-Gebiet* U ein *Rechteck* wählt.
- iii) Weil die *Parametrisierung* \mathbf{P} von einem *Gebiet* in 2D in den *Raum* in 3D führt, kann sie als *Funktion* niemals *surjektiv* sein.
- iv) Die *Parametrisierung* \mathbf{P} soll als *Funktion* so “*injektiv* wie möglich” gewählt werden. Solange nur mehrere einzelne *Punkte* oder *Kurvenstücke* aus dem *Karten-Gebiet* auf den gleichen *Punkt* in M abgebildet werden, ergeben sich in der Regel keine Probleme.
- v) Die *Parameter* bzw. *Koordinaten* u und v können je nach *Fläche* ganz unterschiedliche Bedeutungen und Masseinheiten haben, z.B. *Längen*, *Winkel*, etc..

Beispiele:

- Eine Ebene, welche durch den Punkt \mathbf{P}_0 verläuft und aufgespannt wird durch die beiden Vektoren \mathbf{v} und \mathbf{w} .

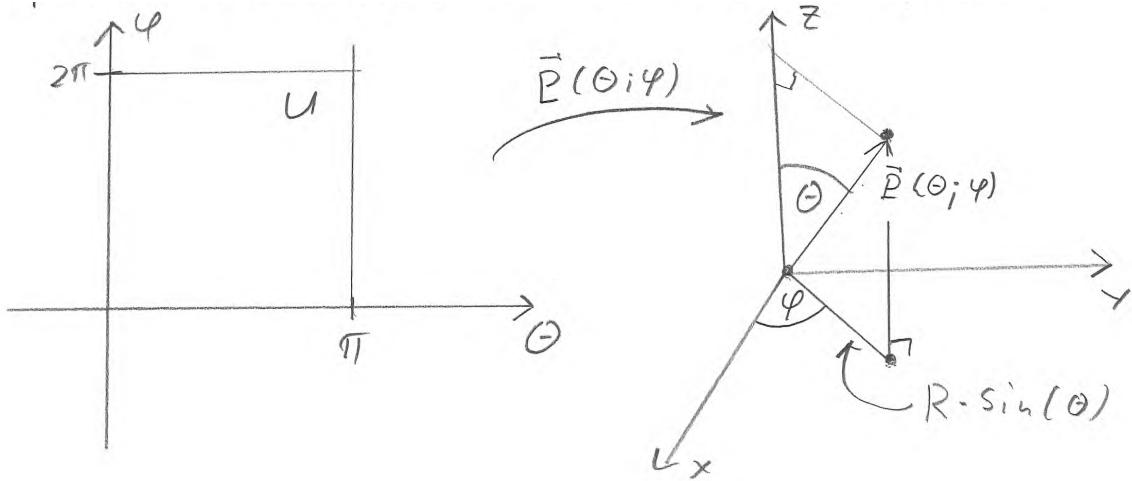


Das Karten-Gebiet ist $U = \mathbb{R}^2$ und die übliche Parametrisierung lautet

$$\mathbf{P}(u; v) = \mathbf{P}_0 + u \cdot \mathbf{v} + v \cdot \mathbf{w}. \quad (2.99)$$

Diese Parametrisierung ist eine bijektive Abbildung zwischen der ganzen u - v -Ebene und der zu parametrisierenden Ebene M in 3D.

- Eine Sphäre mit Radius $R > 0$ und Mittelpunkt am Ursprung.



Das Karten-Gebiet ist $U = [0, \pi] \times [0, 2\pi[$ und die übliche Parametrisierung lautet

$$\mathbf{P}(\theta; \varphi) = \begin{bmatrix} x(\theta; \varphi) \\ y(\theta; \varphi) \\ z(\theta; \varphi) \end{bmatrix} = \begin{bmatrix} R \cdot \sin(\theta) \cdot \cos(\varphi) \\ R \cdot \sin(\theta) \cdot \sin(\varphi) \\ R \cdot \cos(\theta) \end{bmatrix}. \quad (2.100)$$

Diese Parametrisierung ist nicht injektiv, weil die Strecken $\{0\} \times [0, 2\pi[$ und $\{\pi\} \times [0, 2\pi[$ jeweils als Ganzes auf den Nord- bzw. Südpol der Sphäre abgebildet werden.

2.4.1.2 Koordinatenbasis-Vektorfelder

Wir betrachten folgende Definition.

Definition 2.12 Koordinatenbasis-Vektorfelder

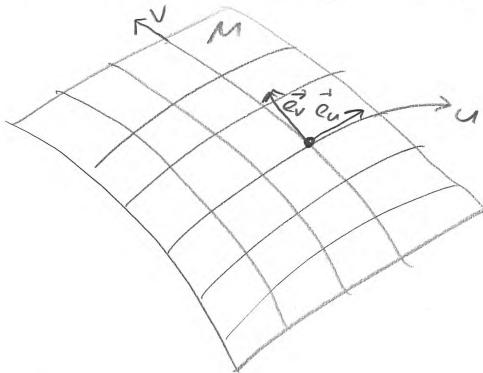
Seien $U \subseteq \mathbb{R}^2$ und $\mathbf{P} : U \rightarrow \mathbb{R}^3$ eine parametrisierte Fläche mit Koordinaten u und v . Die Koordinatenbasis-Vektorfelder sind die Vektorfelder

$$\begin{aligned} \mathbf{e}_u : U &\rightarrow \mathbb{R}^3 \\ (u; v) &\mapsto \mathbf{e}_u(u; v) := \mathbf{P}_{,u}(u; v) \end{aligned} \quad \text{und} \quad \begin{aligned} \mathbf{e}_v : U &\rightarrow \mathbb{R}^3 \\ (u; v) &\mapsto \mathbf{e}_v(u; v) := \mathbf{P}_{,v}(u; v), \end{aligned} \quad (2.101)$$

wobei mit $\mathbf{P}_{,u}$ und $\mathbf{P}_{,v}$ die Ableitungen von \mathbf{P} nach den Koordinaten u bzw. v gemeint sind.

Bemerkungen:

- i) Die Koordinatenbasis-Vektorfelder sind Funktionen der Koordinaten u und v und in diesem Sinne wirklich Vektorfelder.
- ii) Die Koordinatenbasis-Vektorfelder werden bezeichnet durch $\{\mathbf{e}_u, \mathbf{e}_v\}$ oder $\{\mathbf{e}_1, \mathbf{e}_2\}$.
- iii) An jedem Punkt von M zeigen die Koordinatenbasis-Vektorfelder je in eine tangentiale Richtung. Sie spannen daher die Tangentialebene an M auf.



2.4.1.3 Normalen-Vektor

Wir betrachten folgende Definition.

Definition 2.13 Normalen-Vektor

Der Normalen-Vektor einer parametrisierten Fläche ist das GRASSMANN-Vektor-Produkt der Koordinatenbasis-Vektorfelder, d.h.

$$\mathbf{n} := \mathbf{e}_u \times \mathbf{e}_v. \quad (2.102)$$

Bemerkungen:

- i) Der Normalen-Vektor steht senkrecht auf der Fläche M .
- ii) Die Länge des Normalen-Vektors hängt ab von der Parametrisierung.

iii) Durch *Normierung* erhält man aus dem *Normalen-Vektor* den *Einheitsnormalen-Vektor*

$$\hat{\mathbf{n}} := \pm \frac{1}{|\mathbf{n}|} \cdot \mathbf{n}. \quad (2.103)$$

iv) Für geschlossene *Flächen* M wird das *Vorzeichen* des *Einheitsnormalen-Vektors* so gewählt, dass $\hat{\mathbf{n}}$ überall nach aussen zeigt.

2.4.1.4 Metrik

Wir betrachten folgende Definition.

Definition 2.14 *Metrik*

Die *Metrik* einer *parametrisierten Fläche* ist die *GRAM-Matrix* der *Koordinatenbasis-Vektorfelder*, d.h.

$$G := \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} := \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle \end{bmatrix}. \quad (2.104)$$

Bemerkungen:

i) Weil das *Skalar-Produkt symmetrisch* ist, gilt dies auch für G , d.h.

$$G^T = G. \quad (2.105)$$

ii) Weil das *Skalar-Produkt positiv definit* ist, gilt dies auch für G , d.h.

$$g := \det(G) = g_{11} \cdot g_{22} - g_{21} \cdot g_{12} > 0. \quad (2.106)$$

iii) In vielen Anwendungen stehen die *Koordinatenbasis-Vektorfelder senkrecht* aufeinander. Allgemein gilt

$$\mathbf{e}_1 \perp \mathbf{e}_2 \Leftrightarrow G \text{ ist } \textit{diagonal}. \quad (2.107)$$

2.4.1.5 Mass-Funktion

Wir betrachten folgende Definition.

Definition 2.15 *Mass-Funktion*

Seien G die *Metrik* einer *parametrisierten Fläche* und

$$g := \det(G), \quad (2.108)$$

dann ist die *Mass-Funktion* die *Wurzel* \sqrt{g} .

Wir betrachten folgenden Satz.

Satz 2.8 *Normalen-Vektor & Mass-Funktion*

Seien \mathbf{n} der *Normalen-Vektor* und \sqrt{g} die *Mass-Funktion* einer *parametrisierten Fläche*, dann gilt

$$\sqrt{g} = |\mathbf{n}|. \quad (2.109)$$

Beweis: Es seien \mathbf{e}_1 und \mathbf{e}_2 die *Koordinatenbasis-Vektorfelder* der *parametrisierten Fläche* mit Zwischenwinkel $\alpha := \angle(\mathbf{e}_1; \mathbf{e}_2) \in [0, \pi]$. Somit folgt $\sin(\alpha) \geq 0$ und

$$\begin{aligned}
\underline{\underline{g}} &= \sqrt{\det(G)} = \sqrt{g_{11} \cdot g_{22} - g_{12} \cdot g_{21}} = \sqrt{\langle \mathbf{e}_1, \mathbf{e}_1 \rangle \cdot \langle \mathbf{e}_2, \mathbf{e}_2 \rangle - \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \cdot \langle \mathbf{e}_2, \mathbf{e}_1 \rangle} \\
&= \sqrt{|\mathbf{e}_1|^2 \cdot |\mathbf{e}_2|^2 - \langle \mathbf{e}_1, \mathbf{e}_2 \rangle^2} = \sqrt{|\mathbf{e}_1|^2 \cdot |\mathbf{e}_2|^2 - |\mathbf{e}_1|^2 \cdot |\mathbf{e}_2|^2 \cdot \cos^2(\alpha)} \\
&= \sqrt{|\mathbf{e}_1|^2 \cdot |\mathbf{e}_2|^2 \cdot (1 - \cos^2(\alpha))} = \sqrt{|\mathbf{e}_1|^2 \cdot |\mathbf{e}_2|^2 \cdot \sin^2(\alpha)} = |\mathbf{e}_1| \cdot |\mathbf{e}_2| \cdot \sin(\alpha) \\
&= |\mathbf{e}_1 \times \mathbf{e}_2| = |\underline{\mathbf{n}}|. \tag{2.110}
\end{aligned}$$

Damit haben wir den Satz bewiesen. \square

Bemerkungen:

i) Wegen $g > 0$ kann die *Wurzel* immer in den *reellen Zahlen* gezogen werden und es gilt

$$\sqrt{g} > 0. \tag{2.111}$$

ii) Für den *Einheitsnormalen-Vektor* folgt

$$\hat{\mathbf{n}} := \pm \frac{1}{\sqrt{g}} \cdot \mathbf{n}. \tag{2.112}$$

2.4.2 Flächenintegrale

Sei $\mathbf{P} : U \rightarrow \mathbb{R}^3$ die *Parametrisierung* einer *Fläche* M . Mit Hilfe der *Parametrisierung*, welche alle *Punkte* von M durch *Koordinatenpaare* $(u; v) \in U$ beschreibt, kann eine *Funktion* des Typs

$$f : M \rightarrow \mathbb{R} \quad (2.113)$$

aufgefasst werden als *Funktion* der Form

$$\begin{aligned} f &: U \rightarrow \mathbb{R} \\ (u; v) &\mapsto f(u; v) := f(\mathbf{P}(u; v)). \end{aligned} \quad (2.114)$$

Wir betrachten folgende Definition.

Definition 2.16 *Flächenintegral* über eine *parametrisierte Fläche*

Seien $U \subseteq \mathbb{R}^2$, $\mathbf{P} : U \rightarrow \mathbb{R}^3$ eine *parametrisierte Fläche* mit *Mass-Funktion* \sqrt{g} und $f : M \rightarrow \mathbb{R}$, dann ist das *Integral* von f über M definiert durch

$$\boxed{\int_M f \, dA := \int_U f \sqrt{g} \, dU.} \quad (2.115)$$

Wir betrachten folgenden Satz.

Satz 2.9 *Flächeninhalt* einer *parametrisierten Fläche*

Seien $U \subseteq \mathbb{R}^2$ und $\mathbf{P} : U \rightarrow \mathbb{R}^3$ die *Parametrisierung* einer *parametrisierten Fläche* $M = \mathbf{P}(U)$ mit *Mass-Funktion* \sqrt{g} , dann lässt sich der *Flächeninhalt* von M berechnen durch

$$\boxed{A = \int_M 1 \, dA = \int_U \sqrt{g} \, dU.} \quad (2.116)$$

Bemerkungen:

i) Für die *Masseinheit* erhalten wir

$$\boxed{\left[\int_M f \, dA \right] = [f] \cdot [A] = [f] \cdot [\sqrt{g}] \cdot [u] \cdot [v].} \quad (2.117)$$

ii) Ist das *Karten-Gebiet* U ein *Rechteck* der Form

$$U = [u_0, u_E] \times [v_0, v_E], \quad (2.118)$$

dann lässt sich ein *Flächenintegral* berechnen durch

$$\boxed{I = \int_M f \, dA := \int_U f \sqrt{g} \, dU = \int_{u_0}^{u_E} \int_{v_0}^{v_E} f(u; v) \sqrt{g(u; v)} \, dv \, du.} \quad (2.119)$$

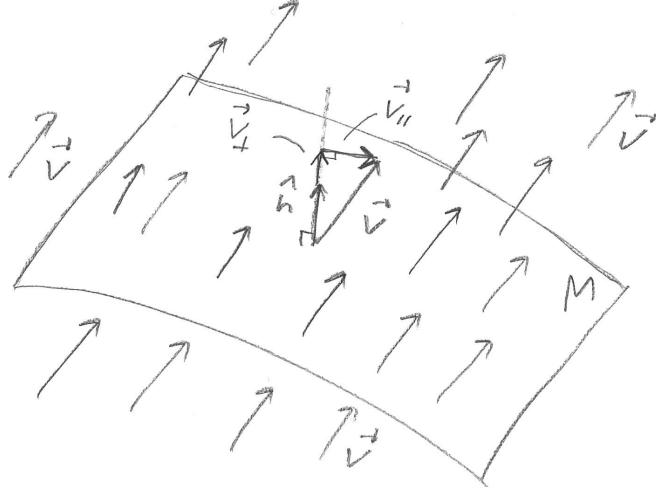
iii) Ein kleines *Flächenstück* auf M kann im Sinne eines *ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess* ausgedrückt werden durch

$$\boxed{\underline{\delta A} \approx \sqrt{g} \cdot \delta U = \sqrt{g} \cdot \delta u \cdot \delta v.} \quad (2.120)$$

Die *Mass-Funktion* ist also gerade der *Umrechnungsfaktor* für den *Flächeninhalt* δU eines kleinen *Flächenstücks* im *Kartengebiet* U auf das entsprechende kleine *Flächenstück* δA auf M .

2.4.3 Flussintegrale

Wir betrachten eine *parametrisierte Fläche* im Bereich eines *Vektorfeldes*. Die Situation ist in der folgenden Skizze dargestellt.



Wir betrachten die folgende Definition.

Definition 2.17 Fluss eines Vektorfeldes

Seien M eine *parametrisierte Fläche* mit *Einheitsnormalen-Vektor* $\hat{\mathbf{n}}$ und $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein *integrierbares Vektorfeld*. Der *Fluss* des *Vektorfeldes* \mathbf{v} durch die *Fläche* M ist

$$\Phi := \int_M \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA. \quad (2.121)$$

Bemerkungen:

- i) Die Begriffe *Flux*, *Fluss* und *Flussintegral* sind synonym.
- ii) Für die Masseinheit erhalten wir

$$[\underline{\Phi}] = [\langle \mathbf{v}, \hat{\mathbf{n}} \rangle] \cdot [A] = [\mathbf{v}] \cdot [\hat{\mathbf{n}}] \cdot [A] = [\mathbf{v}] \cdot 1 \cdot [A] = [\underline{\mathbf{v}}] \cdot [\underline{A}]. \quad (2.122)$$

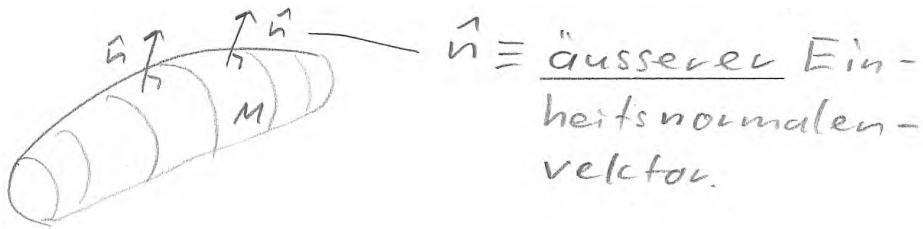
- iii) Man kann zeigen, dass das *Flussintegral* Φ bis auf das *Vorzeichen* nicht von der Wahl der *Parametrisierung* sondern nur von der *Fläche* M abhängt (sofern mehrfache Durchläufe auch mehrfach gerechnet werden).
- iv) Um ein *Flussintegral* auszurechnen, müssen die *Koordinaten* der *Punkte* entlang der *parametrisierten Fläche* im *Vektorfeld* eingesetzt werden. Vollständig ausgeschrieben mit allen Abhängigkeiten ergibt dies

$$\Phi = \int_{u_0}^{u_E} \int_{v_0}^{v_E} \langle \mathbf{v}(x(u; v); y(u; v); z(u; v)), \hat{\mathbf{n}}(u; v) \rangle \cdot \sqrt{g(u; v)} \, dv \, du. \quad (2.123)$$

- v) In der Literatur findet man für *Flussintegrale* durch eine *Fläche* M die Schreibweisen

$$\Phi = \int_M \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA = \int_M \mathbf{v} \cdot d\mathbf{A}. \quad (2.124)$$

- vi)** Ein *Flussintegral* über eine *geschlossene Fläche* heisst *Perforation* des *Vektorfeldes* durch die betreffende *Fläche*.



In diesem Fall verwendet man das *Ring-Integralzeichen*, d.h.

$$\Phi = \oint_M \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA. \quad (2.125)$$

- vii)** Der *Fluss* eines *Vektorfeldes* \mathbf{v} durch eine *Fläche* ist das *Integral* des Anteils von \mathbf{v} , welcher entlang $\hat{\mathbf{n}}$ und somit *senkrecht* zur *Fläche* zeigt. Der zu $\hat{\mathbf{n}}$ *senkrechte* bzw. zur *Fläche parallel* Anteil von \mathbf{v} ist für den *Fluss* irrelevant.

Wir betrachten den folgenden Satz.

Satz 2.10 *Fluss bei konstantem Anteil senkrecht zur Fläche*

Seien M eine parametrisierte *Fläche* mit *Einheitsnormalen-Vektor* $\hat{\mathbf{n}}$ und *Flächeninhalt* A sowie $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein *Vektorfeld*. Gilt entlang der *Fläche* $\langle \mathbf{v}, \hat{\mathbf{n}} \rangle =: C \equiv \text{konst.}$, dann beträgt der *Fluss* von \mathbf{v} durch die *Fläche*

$$\boxed{\Phi = C \cdot A.} \quad (2.126)$$

Beweis: Für den *Fluss* von \mathbf{v} durch die *Fläche* erhalten wir

$$\underline{\Phi} = \int_M \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA = \int_M C \, dA = C \int_M 1 \, dA = \underline{\underline{C \cdot A.}} \quad (2.127)$$

Damit haben wir den Satz bewiesen. \square

Wir betrachten den folgenden Satz.

Satz 2.11 *Perforation eines homogenen Vektorfeldes*

Die *Perforation* eines *homogenen Vektorfeldes* verschwindet.

Anwendungen:

- *Strömungsdynamik*:

$$\Phi_v = \int_M \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA \equiv \text{Volumen-Fluss des Mediums durch die Fläche} \left[\frac{\text{m}^3}{\text{s}} \right]. \quad (2.128)$$

- *Elektrodynamik*: Wir betrachten die folgenden Fälle getrennt.

Fall 1: M ist eine *geschlossene Fläche*. Dann gilt

$$\Phi_E = \oint_M \langle \mathbf{E}, \hat{\mathbf{n}} \rangle \, dA = \frac{1}{\epsilon_0} \cdot Q_{eg} \quad (2.129)$$

$$\Phi_B = \oint_M \langle \mathbf{B}, \hat{\mathbf{n}} \rangle \, dA = 0. \quad (2.130)$$

Fall 2: M ist eine Fläche mit geschlossener Randkurve γ . Dann gilt

$$\Phi_J = I \equiv \text{Stromstärke durch die Fläche [A]} \quad (2.131)$$

$$\Phi_S = P \equiv \text{Leistung durch die Fläche [W]} \quad (2.132)$$

$$U_{\text{ind}} = \Upsilon_E = \oint_{\gamma} \langle \mathbf{E}, \hat{\mathbf{e}} \rangle \, ds = - \left(\int_M \langle \mathbf{B}, \hat{\mathbf{n}} \rangle \, dA \right)^\bullet = -\dot{\Phi}_B. \quad (2.133)$$

2.5 Mehrfach-Differentiale

2.5.1 Partielle Ableitungen

Wir betrachten die folgende Definition.

Definition 2.18 Partielle Ableitung

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine reellwertige Funktion. Die partiellen Ableitungen von f sind die Ableitungen von f nach jeweils einer der n Variablen, wobei die anderen als Konstanten betrachtet werden.

Bemerkungen:

i) Eine reellwertige Funktion heisst differenzierbar, wenn alle partiellen Ableitungen existieren und stetig sind.

ii) Wie die Ableitung in 1D können auch die partiellen Ableitungen in nD mit Hilfe des NEWTON-Differenzenquotienten definiert werden gemäss

$$f_{,\mu}(x_1; x_2; \dots; x_n) := \lim_{\delta s \rightarrow 0} \frac{f(x_1; x_2; \dots; x_\mu + \delta s; \dots; x_n) - f(x_1; x_2; \dots; x_n)}{\delta s}. \quad (2.134)$$

iii) Für die Masseinheit erhalten wir

$$[f_{,\mu}] = \frac{[f]}{[x_\mu]}. \quad (2.135)$$

iv) Die partiellen Ableitungen beschreiben an jedem Punkt die Steigungen des Funktionsgrafen in Richtung der Koordinatenachsen.

v) In der Literatur sind folgende Schreibweisen gebräuchlich

$$f_{,\mu} = f_{,x_\mu} = f_{x_\mu} = \frac{\partial f}{\partial x_\mu} = \frac{\partial}{\partial x_\mu} f = \partial_\mu f. \quad (2.136)$$

2.5.2 Gradient

Wir betrachten die folgende Definition.

Definition 2.19 Gradient

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differentierbare reellwertige Funktion. Der Gradient von f ist das Vektorfeld

$$\nabla f := \begin{bmatrix} f_{,1} \\ f_{,2} \\ \vdots \\ f_{,n} \end{bmatrix}. \quad (2.137)$$

Bemerkungen:

i) Der Gradient ist eine allgemeine Konstruktion in nD.

ii) Ist f ein Skalarfeld, dann ist ∇f ein Vektorfeld.

Beispiele:

- Wir betrachten

$$f(x; y) := x^2 \cdot y^2. \quad (2.138)$$

Der Gradient von f ist

$$\underline{\nabla} f = \begin{bmatrix} f_{,1} \\ f_{,2} \end{bmatrix} = \begin{bmatrix} 2x \cdot y^2 \\ x^2 \cdot 2y \end{bmatrix} = \begin{bmatrix} 2xy^2 \\ 2x^2y \end{bmatrix}. \quad (2.139)$$

- Wir betrachten

$$f(x; y; z) := x^2 \cdot y + z. \quad (2.140)$$

Der Gradient von f ist

$$\underline{\nabla} f = \begin{bmatrix} f_{,1} \\ f_{,2} \\ f_{,3} \end{bmatrix} = \begin{bmatrix} 2x \cdot y + 0 \\ x^2 \cdot 1 + 0 \\ 0 + 1 \end{bmatrix} = \begin{bmatrix} 2xy \\ x^2 \\ 1 \end{bmatrix}. \quad (2.141)$$

Wir betrachten den folgenden Satz.

Satz 2.12 Elementare Rechenregeln für Gradienten

Seien $n \in \mathbb{N}^+$, $g, h : \mathbb{R}^n \rightarrow \mathbb{R}$ differentierbare Funktionen und $a, b \in \mathbb{R}$, dann gelten die folgenden Rechenregeln.

(a) Faktor-Regel:

$$\boxed{\nabla(a \cdot g) = a \cdot \nabla g}$$

(c) Linearität:

$$\boxed{\nabla(a \cdot g + b \cdot h) = a \cdot \nabla g + b \cdot \nabla h}$$

(b) Summen-Regel:

$$\boxed{\nabla(g + h) = \nabla g + \nabla h}$$

(d) Produkt-Regel:

$$\boxed{\nabla(g \cdot h) = (\nabla g) \cdot h + g \cdot \nabla h}$$

Wir betrachten den folgenden Satz.

Satz 2.13 Ketten-Regeln für Gradienten

Seien $n \in \mathbb{N}^+$, dann gelten folgende Ketten-Regeln.

(a) Ketten-Regel A: Für differentierbare $g : \mathbb{R} \rightarrow \mathbb{R}$ und $h : \mathbb{R}^n \rightarrow \mathbb{R}$ gilt

$$\boxed{f(x^1; \dots; x^n) := g(h(x^1; \dots; x^n)) \Rightarrow \nabla f = g'(h(x^1; \dots; x^n)) \cdot \nabla h.}$$

(b) Ketten-Regel B: Für differentierbare $g : \mathbb{R}^n \rightarrow \mathbb{R}$ und $\mathbf{h} : \mathbb{R} \rightarrow \mathbb{R}^n$ gilt

$$\boxed{f(x) := g(\mathbf{h}(x)) \Rightarrow f'(x) = \langle \nabla g(\mathbf{h}(x)), \mathbf{h}'(x) \rangle.}$$

2.5.3 Hesse-Matrix

Wir betrachten die folgende Definition.

Definition 2.20 HESSE-Matrix

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweifach differentierbare reellwertige Funktion. Die HESSE-Matrix von f ist das Vektorfeld

$$\nabla^2 f := \begin{bmatrix} f_{,1,1} & f_{,1,2} & \cdots & f_{,1,n} \\ f_{,2,1} & f_{,2,2} & \cdots & f_{,2,n} \\ \vdots & \vdots & \vdots & \vdots \\ f_{,n,1} & f_{,n,2} & \cdots & f_{,n,n} \end{bmatrix}. \quad (2.142)$$

Beispiele:

- Wir betrachten

$$f(x; y) := x^2 \cdot y^2. \quad (2.143)$$

Der Gradient von f ist

$$\nabla f = \begin{bmatrix} f_{,1} \\ f_{,2} \end{bmatrix} = \begin{bmatrix} 2x \cdot y^2 \\ x^2 \cdot 2y \end{bmatrix} = \begin{bmatrix} 2xy^2 \\ 2x^2y \end{bmatrix}. \quad (2.144)$$

Die HESSE-Matrix von f ist

$$\underline{\underline{\nabla^2 f}} = \begin{bmatrix} f_{,1,1} & f_{,1,2} \\ f_{,2,1} & f_{,2,2} \end{bmatrix} = \begin{bmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}. \quad (2.145)$$

Wir betrachten den folgenden Satz.

Satz 2.14 SCHWARZ-CLAIRAUT-YOUNG-Satz

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweifach differentierbare reellwertige Funktion mit HESSE-Matrix $H \in \mathbb{M}(n, n, \mathbb{R})$, dann gilt

$$H^T = H. \quad (2.146)$$

Bemerkungen:

- Die Symmetrie der HESSE-Matrix gemäss SCHWARZ-CLAIRAUT-YOUNG-Satz ist äquivalent zur Tatsache, dass die partiellen Ableitungen vertauscht werden dürfen, d.h. für alle $\mu, \nu \in \{1, \dots, n\}$ gilt

$$f_{,\nu,\mu} = f_{,\mu,\nu}. \quad (2.147)$$

- Weil die HESSE-Matrix symmetrisch ist, ist sie diagonalisierbar, d.h. ähnlich zu einer diagonalen Matrix.

Wir betrachten die folgende Definition.

Definition 2.21 LAPLACE-Ableitung

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweifach differentierbare reellwertige Funktion. Die LAPLACE-Ableitung von f ist

$$\Delta f := \text{tr}(\nabla^2 f) = f_{,1,1} + f_{,2,2} + \dots + f_{,n,n}. \quad (2.148)$$

2.5.4 Divergenz

Wir betrachten die folgende Definition.

Definition 2.22 Divergenz

Seien $n \in \mathbb{N}^+$ und $\mathbf{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein differentierbares Vektorfeld mit Komponenten

$$\mathbf{v}(x^1; \dots; x^n) = \begin{bmatrix} v^1(x^1; \dots; x^n) \\ \vdots \\ v^n(x^1; \dots; x^n) \end{bmatrix}. \quad (2.149)$$

Die Divergenz von \mathbf{v} ist

$$\boxed{\operatorname{div}(\mathbf{v}) := v^1_{,1} + v^2_{,2} + \dots + v^n_{,n}} \quad (2.150)$$

Beispiele:

- Wir betrachten

$$\mathbf{v}(x; y) := \begin{bmatrix} x \cdot y^2 \\ x^3 \cdot y^3 \end{bmatrix}. \quad (2.151)$$

Die Divergenz von \mathbf{v} ist

$$\boxed{\operatorname{div}(\mathbf{v}) = v^1_{,1} + v^2_{,2} = (x \cdot y^2)_{,x} + (x^3 \cdot y^3)_{,y} = 1 \cdot y^2 + x^3 \cdot 3 \cdot y^2 = \underline{\underline{y^2}} \cdot (1 + 3x^3)}. \quad (2.152)$$

Bemerkungen:

- i) Die Divergenz eines Vektorfeldes ist eine allgemeine Konstruktion in nD.
- ii) Die Divergenz eines Vektorfeldes ist ein Skalarfeld.
- iii) Die Divergenz eines Vektorfeldes ist ein Mass für dessen Quellendichte.
- iv) Ein Vektorfeld \mathbf{v} heisst quellenfrei, falls gilt

$$\operatorname{div}(\mathbf{v}) = 0. \quad (2.153)$$

Wir betrachten den folgenden Satz.

Satz 2.15 Elementare Rechenregeln für Divergenzen

Seien $n \in \mathbb{N}^+$, $\mathbf{v}, \mathbf{w} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ differentiablae Vektorfelder, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differentiabre Funktion und $a, b \in \mathbb{R}$, dann gelten die folgenden Rechenregeln.

(a) Faktor-Regel:

$$\boxed{\operatorname{div}(a \cdot \mathbf{v}) = a \cdot \operatorname{div}(\mathbf{v})}$$

(c) Linearität:

$$\boxed{\operatorname{div}(a \cdot \mathbf{v} + b \cdot \mathbf{w}) = a \cdot \operatorname{div}(\mathbf{v}) + b \cdot \operatorname{div}(\mathbf{w})}$$

(b) Summen-Regel:

$$\boxed{\operatorname{div}(\mathbf{v} + \mathbf{w}) = \operatorname{div}(\mathbf{v}) + \operatorname{div}(\mathbf{w})}$$

(d) Produkt-Regel:

$$\boxed{\operatorname{div}(f \cdot \mathbf{v}) = \langle \nabla f, \mathbf{v} \rangle + f \cdot \operatorname{div}(\mathbf{v})}$$

2.5.5 Rotation

Wir betrachten die folgende Definition.

Definition 2.23 *Rotation* in 2D

Sei $\mathbf{v} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein *differentierbares Vektorfeld* mit *Komponenten*

$$\mathbf{v}(x^1; x^2) = \begin{bmatrix} v^1(x^1; x^2) \\ v^2(x^1; x^2) \end{bmatrix}. \quad (2.154)$$

Die *Rotation* von \mathbf{v} ist

$$\text{rot}(\mathbf{v}) := v^2_{,1} - v^1_{,2}. \quad (2.155)$$

Wir betrachten die folgende Definition.

Definition 2.24 *Rotation* in 3D

Sei $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein *differentierbares Vektorfeld* mit *Komponenten*

$$\mathbf{v}(x^1; x^2; x^3) = \begin{bmatrix} v^1(x^1; x^2; x^3) \\ v^2(x^1; x^2; x^3) \\ v^3(x^1; x^2; x^3) \end{bmatrix}. \quad (2.156)$$

Die *Rotation* von \mathbf{v} ist

$$\text{rot}(\mathbf{v}) := \begin{bmatrix} v^3_{,2} - v^2_{,3} \\ v^1_{,3} - v^3_{,1} \\ v^2_{,1} - v^1_{,2} \end{bmatrix}. \quad (2.157)$$

Bemerkungen:

- i) Die *Rotation* eines *Vektorfeldes* ist eine spezielle Konstruktion in 2D und 3D.
- ii) Die *Rotation* eines *Vektorfeldes* in 2D ist ein *Skalarfeld*.
- iii) Die *Rotation* eines *Vektorfeldes* in 3D ist ein *Vektorfeld* in 3D.
- iv) Die *Rotation* eines *Vektorfeldes* ist ein Mass für dessen *Wirbeldichte*.
- v) Ein *Vektorfeld* \mathbf{v} heisst *wirbelfrei*, falls gilt

$$\text{rot}(\mathbf{v}) = 0. \quad (2.158)$$

- vi) In 3D steht $\text{rot}(\mathbf{v})$ *senkrecht* auf der “*Wirbel-Ebene*” von \mathbf{v} , analog zum *Drehimpuls* einer *rotierenden Scheibe*.
- vii) Im Englischen wird die *Rotation* als *curl* bezeichnet und entsprechend auch so in Formeln notiert, d.h. durch $\text{curl}(\mathbf{v})$.

Wir betrachten den folgenden Satz.

Satz 2.16 Elementare Rechenregeln für *Rotationen*

Seien $n \in \{2, 3\}$, $\mathbf{v}, \mathbf{w} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ differentierbare Vektorfelder, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differentierbare Funktion und $a, b \in \mathbb{R}$, dann gelten die folgenden Rechenregeln.

(a) *Faktor-Regel:*

$$\text{rot}(a \cdot \mathbf{v}) = a \cdot \text{rot}(\mathbf{v})$$

(c) *Linearität:*

$$\text{rot}(a \cdot \mathbf{v} + b \cdot \mathbf{w}) = a \cdot \text{rot}(\mathbf{v}) + b \cdot \text{rot}(\mathbf{w})$$

(b) *Summen-Regel:*

$$\text{rot}(\mathbf{v} + \mathbf{w}) = \text{rot}(\mathbf{v}) + \text{rot}(\mathbf{w})$$

(d) *Produkt-Regel* für $n = 3$:

$$\text{rot}(f \cdot \mathbf{v}) = \nabla f \times \mathbf{v} + f \cdot \text{rot}(\mathbf{v})$$

2.5.6 Weiteres zu Gradient, Divergenz und Rotation

2.5.6.1 Kombinierte Rechenregeln

Wir betrachten den folgenden Satz.

Satz 2.17 Kombinierte Rechenregeln für *Gradient, Divergenz & Rotation*

Seien $n \in \mathbb{N}^+$, $\mathbf{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein differentierbares Vektorfeld und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine differentierbare Funktion, dann gelten die folgenden Rechenregeln.

(a) *Divergenz eines Gradienten:*

$$\text{div}(\nabla f) = \Delta f$$

(c) *Divergenz einer Rotation für $n = 3$:*

$$\text{div}(\text{rot}(\mathbf{v})) = 0$$

(b) *Rotation eines Gradienten für $n \in \{2, 3\}$:* (d) *Rotation einer Rotation für $n = 3$:*

$$\text{rot}(\nabla f) = 0$$

$$\text{rot}(\text{rot}(\mathbf{v})) = \nabla \text{div}(\mathbf{v}) - \Delta \mathbf{v}$$

Beweis: Siehe Übungen. □

Wir betrachten den folgenden Satz.

Satz 2.18 Rechenregeln für *Vektor-Produkte* in 3D

Seien $\mathbf{v}, \mathbf{w} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ differentierbare Vektorfelder und $g, h : \mathbb{R}^3 \rightarrow \mathbb{R}$ differentierbare Funktionen, dann gelten die folgenden Rechenregeln.

(a) *Divergenz eines Vektor-Produkts:*

$$\text{div}(\mathbf{v} \times \mathbf{w}) = \langle \text{rot}(\mathbf{v}), \mathbf{w} \rangle - \langle \mathbf{v}, \text{rot}(\mathbf{w}) \rangle$$

(b) *Rotation eines Vektor-Produkts:*

$$\text{rot}(\mathbf{v} \times \mathbf{w}) = \nabla_{\mathbf{w}} \mathbf{v} - \nabla_{\mathbf{v}} \mathbf{w} + \text{div}(\mathbf{w}) \cdot \mathbf{v} - \text{div}(\mathbf{v}) \cdot \mathbf{w}$$

(c) *Divergenz eines Vektor-Produkts von Gradienten:*

$$\text{div}(\nabla g \times \nabla h) = 0$$

Bemerkungen:

- i) Jeder *Gradient* ist *wirbelfrei*.
- ii) Jede *Rotation* ist *quellenfrei*.
- iii) In *kartesischen Koordinaten* ist $\Delta \mathbf{v}$ *komponentenweise* zu berechnen.
- iv) Die Terme $\nabla_w \mathbf{v}$ und $\nabla_v \mathbf{w}$ bezeichnen sogenannte *Richtungsableitungen* (siehe nächste Abschnitte.)

2.5.6.2 Nabla-Operator

Wir betrachten die folgende Definition.

Definition 2.25 *Nabla-Operator* in nD

Sei $n \in \mathbb{N}^+$, dann ist der *Nabla-Operator* in \mathbb{R}^n der *Differentialoperator*

$$\boxed{\nabla := \begin{bmatrix} \partial_1 \\ \vdots \\ \partial_n \end{bmatrix}} \quad (2.159)$$

Bemerkungen:

- i) Der *Nabla-Operator* ist ein abstrakter *Differentialoperator*, welcher erst bei seiner Anwendung auf eine *Funktion* bzw. ein *Vektorfeld* eine sinnvolle *mathematische Grösse* ergibt.
- ii) Durch den *Nabla-Operator* können die *Divergenz* in nD und die *Rotation* in 3D durch *Vektor-Operationen* ausgedrückt werden. Insbesondere in älterer Literatur findet man die Schreibweisen

$$\text{div}(\mathbf{v}) = \langle \nabla, \mathbf{v} \rangle = \nabla \cdot \mathbf{v} \quad (2.160)$$

$$\text{rot}(\mathbf{v}) = \nabla \times \mathbf{v}. \quad (2.161)$$

2.5.6.3 Anwendungen

Viele wichtige Formeln in der *Physik* werden durch *Divergenz* bzw. *Rotation* von *Vektorfeldern* ausgedrückt.

- *Strömungsdynamik*: Beschreibt \mathbf{v} das *Geschwindigkeitsvektorfeld* eines *inkompressiblen Mediums* (z.B. Wasser), dann ist es *quellenfrei*, d.h.

$$\text{div}(\mathbf{v}) = 0. \quad (2.162)$$

- *Elektrodynamik*: Die MAXWELL-Gleichungen beschreiben jeweils *Divergenz* und *Rotation* des *E-Feldes* und *B-Feldes*. Es gilt

$$\left. \begin{array}{l} \text{div}(\mathbf{E}) = \frac{1}{\epsilon_0} \cdot \rho \\ \text{div}(\mathbf{B}) = 0 \end{array} \right| \quad \left. \begin{array}{l} \text{rot}(\mathbf{E}) = -\dot{\mathbf{B}} \\ \text{rot}(\mathbf{B}) = \epsilon_0 \cdot \mu_0 \cdot \dot{\mathbf{E}} + \mu_0 \cdot \mathbf{J}. \end{array} \right| \quad (2.163)$$

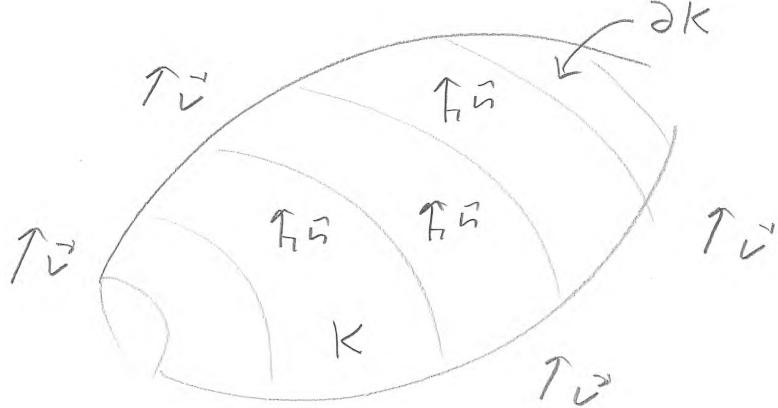
Der *Ladungs-Erhaltungssatz* kann ausgedrückt werden durch die *Kontinuitätsgleichung*

$$\dot{\rho} + \text{div}(\mathbf{J}) = 0. \quad (2.164)$$

2.6 Hauptsätze der Vektoranalysis

2.6.1 Gauss-Integralsatz

Wir betrachten einen Körper $K \subset \mathbb{R}^3$ mit Oberfläche ∂K , äusserem Einheitsnormalen-Vektorfeld $\hat{\mathbf{n}}$ im Bereich eines Vektorfeldes $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Die Situation ist in der folgenden Skizze dargestellt.



Wir betrachten den folgenden Satz.

Satz 2.19 GAUSS-Integralsatz in 3D

Seien $K \subset \mathbb{R}^3$ ein Körper mit Oberfläche ∂K , äusserem Einheitsnormalen-Vektorfeld $\hat{\mathbf{n}}$ im Bereich eines differentierbaren Vektorfeldes $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, dann gilt

$$\oint_{\partial K} \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA = \Phi_{\mathbf{v}} = \int_K \operatorname{div}(\mathbf{v}) \, dV. \quad (2.165)$$

Bemerkungen:

i) Der GAUSS-Integralsatz besagt die *Gleichheit* eines *Flussintegrals* mit einem *Volumenintegral*. Es gilt

$$\text{Perforation von } \mathbf{v} \equiv \text{Summe aller eingeschlossenen Quellen von } \mathbf{v}. \quad (2.166)$$

ii) Der GAUSS-Integralsatz kann allgemein in nD formuliert werden.

iii) Der GAUSS-Integralsatz ist eine Verallgemeinerung der NEWTON-LEIBNIZ-Formel in nD.

iv) Der GAUSS-Integralsatz lässt sich bezogen auf beide Seiten der *Gleichung* sinnvoll anwenden.

v) Der GAUSS-Integralsatz etabliert die Interpretation der *Divergenz* als *Quellendichte* eines *Vektorfeldes*.

vi) Gemäss GAUSS-Integralsatz verschwindet offenbar die *Perforation* jedes *quellenfreien Vektorfeldes* durch eine beliebige Oberfläche ∂V . Aus $\operatorname{div}(\mathbf{v}) = 0$ folgt

$$\underline{\underline{\Phi}_{\mathbf{v}}} = \oint_{\partial K} \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \, dA = \int_K \operatorname{div}(\mathbf{v}) \, dV = \int_K 0 \, dV = \underline{\underline{0}}. \quad (2.167)$$

Damit haben wir auch den Spezialfall aus Satz 2.11 bewiesen.

Anwendungen:

- *Strömungsdynamik*: Aus der *Inkompressibilität* einer *Flüssigkeit* folgt die *Quellenfreiheit* des *Geschwindigkeitsvektorfeldes* \mathbf{v} und somit das Verschwinden der *Perforation* von \mathbf{v} durch eine beliebige *Oberfläche* ∂V .
- *Elektrodynamik*: Die MAXWELL-Gleichungen für die *Divergenz* des *E-Feldes* und *B-Feldes* lauten

$$\operatorname{div}(\mathbf{E}) = \frac{1}{\varepsilon_0} \cdot \rho \quad (2.168)$$

$$\operatorname{div}(\mathbf{B}) = 0. \quad (2.169)$$

Für die *Perforation* des *E-Feldes* und *B-Feldes* durch eine beliebige *Oberfläche* ∂V folgt aus dem *GAUSS-Integralsatz*

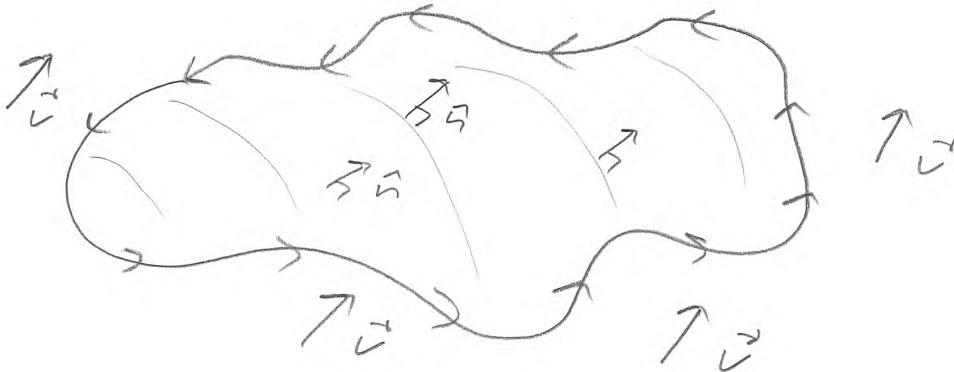
$$\underline{\underline{\Phi_E}} = \oint_{\partial V} \langle \mathbf{E}, \hat{\mathbf{n}} \rangle \, dA = \int_V \operatorname{div}(\mathbf{E}) \, dV = \int_V \frac{1}{\varepsilon_0} \cdot \rho \, dV = \frac{1}{\varepsilon_0} \int_V \rho \, dV = \frac{1}{\varepsilon_0} \cdot Q_{\text{eg}} \quad (2.170)$$

$$\underline{\underline{\Phi_B}} = \oint_{\partial V} \langle \mathbf{B}, \hat{\mathbf{n}} \rangle \, dA = \int_V \operatorname{div}(\mathbf{B}) \, dV = \int_V 0 \, dV = \underline{\underline{0}}. \quad (2.171)$$

- *Erhaltungssätze*
- *Volumen-Berechnungen*
- *Geometrische Analysis*

2.6.2 Stokes-Integralsatz

Wir betrachten ein *Gebiet* $G \subset \mathbb{R}^3$ mit *Randkurve* ∂G , welche das *Einheitsnormalen-Vektorfeld* $\hat{\mathbf{n}}$ *rechts umläuft* im Bereich eines *Vektorfeldes* $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$. Die Situation ist in der folgenden Skizze dargestellt.



Wir betrachten den folgenden Satz.

Satz 2.20 STOKES-Integralsatz in 3D

Seien $G \subset \mathbb{R}^3$ ein *Gebiet* mit *Randkurve* ∂G , welche das *Einheitsnormalen-Vektorfeld* $\hat{\mathbf{n}}$ *rechts umläuft* im Bereich eines *Vektorfeldes* $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, dann gilt

$$\oint_{\partial G} \langle \mathbf{v}, \hat{\mathbf{e}} \rangle \, ds = \Upsilon_{\mathbf{v}} = \int_G \langle \operatorname{rot}(\mathbf{v}), \hat{\mathbf{n}} \rangle \, dA. \quad (2.172)$$

Bemerkungen:

- i) Der STOKES-*Integralsatz* besagt die *Gleichheit* eines *Linienintegrals* mit einem *Flussintegral*. Es gilt

$$\boxed{\text{Zirkulation von } \mathbf{v} \equiv \text{Summe aller eingeschlossenen Wirbel von } \mathbf{v}.} \quad (2.173)$$

- ii) Der STOKES-*Integralsatz* kann allgemein in 2D und 3D formuliert werden.
- iii) Der STOKES-*Integralsatz* ist eine Verallgemeinerung der NEWTON-LEIBNIZ-*Formel* in 2D bzw. 3D.
- iv) Der STOKES-*Integralsatz* lässt sich bezogen auf beide Seiten der *Gleichung* sinnvoll anwenden.
- v) Der STOKES-*Integralsatz* etabliert die Interpretation der *Rotation* als *Wirbeldichte* eines *Vektorfeldes*.
- vi) Gemäss STOKES-*Integralsatz* verschwindet offenbar die *Zirkulation* jedes *wirbelfreien Vektorfeldes* entlang einer beliebigen *Randkurve* ∂G . Aus $\text{rot}(\mathbf{v}) = 0$ folgt

$$\underline{\underline{\Upsilon_{\mathbf{v}}}} = \oint_{\partial G} \langle \mathbf{v}, \hat{\mathbf{e}} \rangle \, ds = \int_G \langle \text{rot}(\mathbf{v}), \hat{\mathbf{n}} \rangle \, dA = \int_G \langle 0, \hat{\mathbf{n}} \rangle \, dA = \int_G 0 \, dA = \underline{\underline{0}} \quad (2.174)$$

Anwendungen:

- *Elektrodynamik*: Die MAXWELL-*Gleichungen* für die *Rotation* des *E-Feldes* und *B-Feldes* lauten

$$\text{rot}(\mathbf{E}) = -\dot{\mathbf{B}} \quad (2.175)$$

$$\text{rot}(\mathbf{B}) = \epsilon_0 \cdot \mu_0 \cdot \dot{\mathbf{E}} + \mu_0 \cdot \mathbf{J}. \quad (2.176)$$

In einer *statischen Situation*, d.h. für $\dot{\mathbf{E}} = \dot{\mathbf{B}} = 0$ vereinfachen sich diese *Gleichungen* zu

$$\text{rot}(\mathbf{E}) = 0 \quad (2.177)$$

$$\text{rot}(\mathbf{B}) = \mu_0 \cdot \mathbf{J}. \quad (2.178)$$

Für die *Zirkulation* des *E-Feldes* und *B-Feldes* entlang einer beliebigen *Randkurve* ∂G folgt aus dem STOKES-*Integralsatz*

$$\underline{\underline{\Upsilon_{\mathbf{E}}}} = \oint_{\partial G} \langle \mathbf{E}, \hat{\mathbf{e}} \rangle \, ds = \int_G \langle \text{rot}(\mathbf{E}), \hat{\mathbf{n}} \rangle \, dA = \int_G \langle 0, \hat{\mathbf{n}} \rangle \, dA = \int_G 0 \, dA = \underline{\underline{0}} \quad (2.179)$$

$$\begin{aligned} \underline{\underline{\Upsilon_{\mathbf{B}}}} &= \oint_{\partial G} \langle \mathbf{B}, \hat{\mathbf{e}} \rangle \, ds = \int_G \langle \text{rot}(\mathbf{B}), \hat{\mathbf{n}} \rangle \, dA = \int_G \langle \mu_0 \cdot \mathbf{J}, \hat{\mathbf{n}} \rangle \, dA = \mu_0 \int_G \langle \mathbf{J}, \hat{\mathbf{n}} \rangle \, dA \\ &= \underline{\underline{\mu_0 \cdot I_{eg}}}. \end{aligned} \quad (2.180)$$

- *Erhaltungssätze*
- *Flächen-Berechnungen*
- *Geometrische Analysis*

2.6.3 Potential-Sätze

Wir betrachten den folgenden Satz.

Satz 2.21 Skalarpotential in 3D

Seien $n \in \{2, 3\}$ und $\mathbf{v} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein differentierbares, wirbelfreies Vektorfeld, dann gibt es ein Skalarfeld $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$, so dass gilt

$$\boxed{\mathbf{v} = \nabla \phi.} \quad (2.181)$$

Bemerkungen:

- i) Jedes wirbelfreie Vektorfeld ist ein Gradientenfeld.
- ii) Das Skalarfeld ϕ heisst Skalarpotential oder Potential des Vektorfeldes \mathbf{v} .
- iii) Das Skalarpotential ist eine Verallgemeinerung der Stammfunktion auf wirbelfreie Vektorfelder in 2D und 3D.
- iv) Für jedes wirbelfreie Vektorfeld \mathbf{v} gibt es unendlich viele Möglichkeiten ein Skalarpotential zu wählen.

Wir betrachten den folgenden Satz.

Satz 2.22 Vektorpotential in 3D

Sei $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein differentierbares, quellenfreies Vektorfeld, dann gibt es ein Vektorfeld $\mathbf{A} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, so dass gilt

$$\boxed{\mathbf{v} = \text{rot}(\mathbf{A})}. \quad (2.182)$$

Bemerkungen:

- i) Jedes quellenfrei Vektorfeld ist ein Rotationsfeld.
- ii) Das Vektorfeld \mathbf{A} heisst Vektorpotential des Vektorfeldes \mathbf{v} .
- iii) Das Vektorpotential ist eine Verallgemeinerung der Stammfunktion auf quellenfreie Vektorfelder in 3D.
- iv) Für jedes quellenfreie Vektorfeld \mathbf{v} gibt es unendlich viele Möglichkeiten ein Vektorpotential zu wählen.

Anwendungen:

- Elektrodynamik: Die MAXWELL-Gleichungen beschreiben jeweils Divergenz und Rotation des E-Feldes und B-Feldes. Es gilt

$$\begin{array}{l|l} \text{div}(\mathbf{E}) = \frac{1}{\epsilon_0} \cdot \rho & \text{rot}(\mathbf{E}) = -\dot{\mathbf{B}} \\ \text{div}(\mathbf{B}) = 0 & \text{rot}(\mathbf{B}) = \epsilon_0 \cdot \mu_0 \cdot \dot{\mathbf{E}} + \mu_0 \cdot \mathbf{J}. \end{array} \quad (2.183)$$

In einer statischen Situation, d.h. für $\dot{\mathbf{E}} = \dot{\mathbf{B}} = 0$ vereinfachen sich diese Gleichungen zu

$$\begin{array}{l|l} \text{div}(\mathbf{E}) = \frac{1}{\epsilon_0} \cdot \rho & \text{rot}(\mathbf{E}) = 0 \\ \text{div}(\mathbf{B}) = 0 & \text{rot}(\mathbf{B}) = \mu_0 \cdot \mathbf{J}. \end{array} \quad (2.184)$$

Weil also das E-Feld wirbelfrei und das B-Feld quellenfrei ist, gibt es entsprechend ein Skalarpotential ϕ und ein Vektorpotential \mathbf{A} , so dass

$$\mathbf{E} = -\nabla \phi \quad \text{und} \quad \mathbf{B} = \text{rot}(\mathbf{A}). \quad (2.185)$$

2.6.4 Zerlegungssatz

Wir betrachten den folgenden Satz.

Satz 2.23 *Zerlegungssatz für Vektorfelder in 3D*

Jedes *differentierbare Vektorfeld* $\mathbf{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ lässt sich zerlegen in eine *Summe* aus einem *wirbelfreien Vektorfeld* $\mathbf{q} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, einem *quellenfreien Vektorfeld* $\mathbf{w} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ und einem *homogenen Vektorfeld* $\mathbf{h} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ gemäss

$$\boxed{\mathbf{v} = \mathbf{w} + \mathbf{q} + \mathbf{h}.} \quad (2.186)$$

Bemerkungen:

- i) Bei einer solchen Zerlegung ist \mathbf{w} ein reines *Wirbelfeld* und \mathbf{q} ein reines *Quellenfeld*.
- ii) Für jedes *differentierbare Vektorfeld* gibt es *unendlich viele* Möglichkeiten eine Zerlegung der Form (2.186) zu wählen.
- iii) Gemäss den *Potential-Sätzen* hat \mathbf{q} ein *Skalarpotential* ϕ und \mathbf{w} ein *Vektorpotential* \mathbf{A} , so dass gilt

$$\mathbf{q} = \nabla\phi \quad \text{und} \quad \mathbf{w} = \text{rot}(\mathbf{A}). \quad (2.187)$$

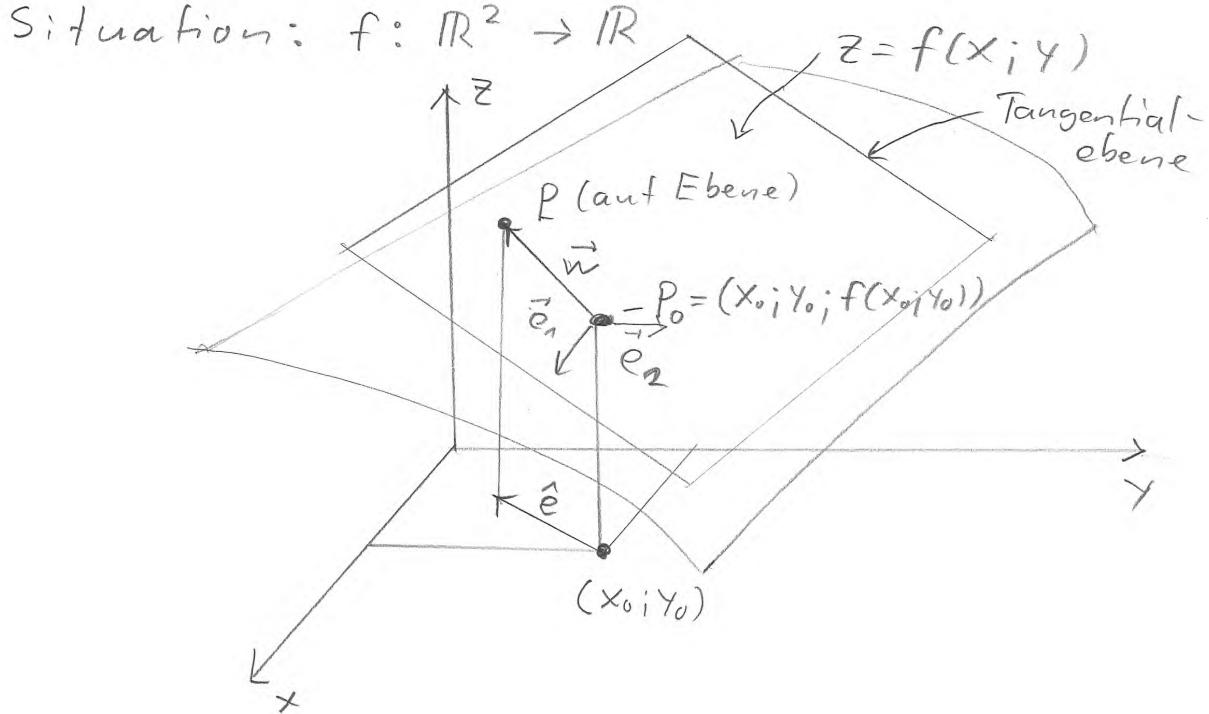
Mit Hilfe dieser *Potentiale* lässt sich die Zerlegung (2.186) schreiben gemäss

$$\underline{\mathbf{v}} = \mathbf{w} + \mathbf{q} + \mathbf{h} = \underline{\text{rot}(\mathbf{A})} + \underline{\nabla\phi} + \underline{\mathbf{h}}. \quad (2.188)$$

2.7 Funktionsdiskussion in mehreren Variablen

2.7.1 Erste Richtungsableitung

Wir betrachten die folgende Situation.



Wir betrachten den folgenden Satz.

Satz 2.24 Steigung entlang eines Funktionsgraphen in 2D

Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ differentierbar, $P_0 = (x_0; y_0; f(x_0; y_0)) \in \mathbb{R}^3$ und $\hat{\mathbf{e}} \in \mathbb{R}^2$ ein Einheitsvektor, dann hat ein Weg auf dem Funktionsgraph von f mit horizontaler Richtung $\hat{\mathbf{e}}$ am Punkt P_0 die Steigung

$$m = \langle \hat{\mathbf{e}}, \nabla f(x_0; y_0) \rangle. \quad (2.189)$$

Beweis: Wir betrachten Funktionsgraphen den von f als parametrisierte Fläche mit Parametrisierung

$$\mathbf{Q}(x; y) := \begin{bmatrix} x \\ y \\ f(x; y) \end{bmatrix}. \quad (2.190)$$

Für die Koordinatenbasis-Vektorfelder erhalten wir

$$\mathbf{e}_1 = \mathbf{Q}_{,x} = \begin{bmatrix} 1 \\ 0 \\ f_{,x} \end{bmatrix} \quad \text{und} \quad \mathbf{e}_2 = \mathbf{Q}_{,y} = \begin{bmatrix} 0 \\ 1 \\ f_{,y} \end{bmatrix}. \quad (2.191)$$

Daraus erhalten wir den Normalen-Vektor

$$\mathbf{n} = \mathbf{e}_1 \times \mathbf{e}_2 = \begin{bmatrix} 1 \\ 0 \\ f_{,x} \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ f_{,y} \end{bmatrix} = \begin{bmatrix} 0 \cdot f_{,y} - f_{,x} \cdot 1 \\ f_{,x} \cdot 0 - 1 \cdot f_{,y} \\ 1 \cdot 1 - 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} -f_{,x} \\ -f_{,y} \\ 1 \end{bmatrix}. \quad (2.192)$$

Wir betrachten $\Delta x, \Delta y \in \mathbb{R}$, so dass gilt

$$\hat{\mathbf{e}} = \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \quad (2.193)$$

Ferner seien $P := (x; y; z)$ der Punkt oberhalb des Punktes $(x_0 + \Delta x; y_0 + \Delta y)$ auf der Tangentialebene E_0 an den Graphen von f im Punkt P_0 und

$$\mathbf{w} := \mathbf{P} - \mathbf{P}_0 = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - f(x_0; y_0) \end{bmatrix} = \begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}. \quad (2.194)$$

Weil der Vektor \mathbf{w} die Punkte P und P_0 auf der Tangentialebene E_0 verbindet, liegt er selbst auf der Tangentialebene E_0 und es muss gelten

$$\begin{aligned} 0 &= \langle \mathbf{w}, \mathbf{n} \rangle = \left\langle \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}, \begin{bmatrix} -f_{,x} \\ -f_{,y} \\ 1 \end{bmatrix} \right\rangle = -\Delta x \cdot f_{,x} - \Delta y \cdot f_{,y} + \Delta z \cdot 1 \\ &= \Delta z - (\Delta x \cdot f_{,x} + \Delta y \cdot f_{,y}) = \Delta z - \left\langle \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}, \begin{bmatrix} f_{,x} \\ f_{,y} \end{bmatrix} \right\rangle = \Delta z - \langle \hat{\mathbf{e}}, \nabla f \rangle. \end{aligned} \quad (2.195)$$

Daraus folgt

$$\Delta z = \langle \hat{\mathbf{e}}, \nabla f \rangle \quad (2.196)$$

und für die gesuchte Steigung erhalten wir

$$\underline{m} = \frac{\Delta z}{\Delta s} = \frac{\Delta z}{|\hat{\mathbf{e}}|} = \frac{\langle \hat{\mathbf{e}}, \nabla f \rangle}{1} = \underline{\underline{\langle \hat{\mathbf{e}}, \nabla f \rangle}}. \quad (2.197)$$

Damit haben wir den Satz bewiesen. \square

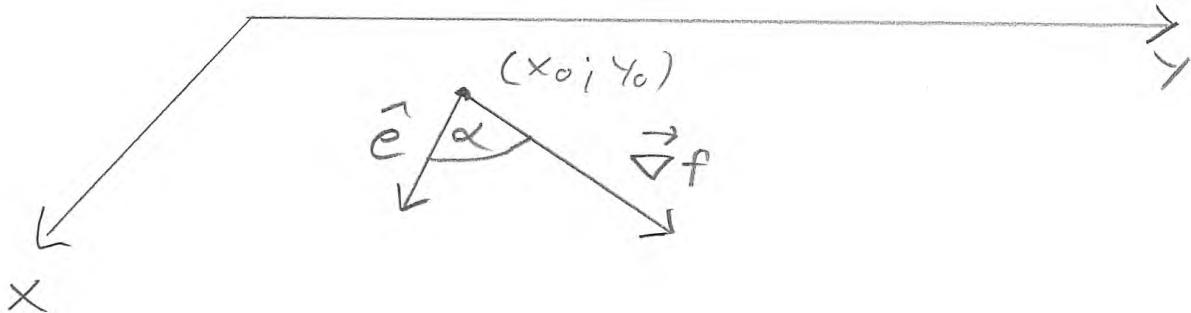
Wir betrachten die folgende Definition.

Definition 2.26 Richtungsableitung in nD

Seien $n \in \mathbb{N}^+$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differentierbar und $\hat{\mathbf{e}} \in \mathbb{R}^n$ ein Einheitsvektor, dann ist die Richtungsableitung von f in Richtung $\hat{\mathbf{e}}$ die reelle Zahl

$$\boxed{\nabla_{\hat{\mathbf{e}}} f := \langle \hat{\mathbf{e}}, \nabla f \rangle}. \quad (2.198)$$

Wir betrachten die Situation in der x - y -Ebene gemäss folgende Skizze.



Wir betrachten den folgenden Satz.

Satz 2.25 Eigenschaften des Gradienten in 2D

Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ differentierbar, $\hat{\mathbf{e}} \in \mathbb{R}^2$ ein Einheitsvektor und $\alpha = \angle(\nabla f; \hat{\mathbf{e}})$, dann gilt folgendes.

- (a) $\nabla_{\hat{\mathbf{e}}} f \in [-|\nabla f|, +|\nabla f|]$
- (b) $\alpha = 0 \Leftrightarrow \hat{\mathbf{e}} \parallel +\nabla f \Leftrightarrow \nabla_{\hat{\mathbf{e}}} f = +|\nabla f|$ maximal
- (c) $\alpha = \pi \Leftrightarrow \hat{\mathbf{e}} \parallel -\nabla f \Leftrightarrow \nabla_{\hat{\mathbf{e}}} f = -|\nabla f|$ minimal
- (d) $\alpha = \frac{\pi}{2} \Leftrightarrow \hat{\mathbf{e}} \perp \nabla f \Leftrightarrow \nabla_{\hat{\mathbf{e}}} f = 0$
- (e) $\nabla f = 0 \Leftrightarrow$ Tangentialebene an den Graphen von f verläuft horizontal

Beweis: Für die Richtungsableitung von f in Richtung $\hat{\mathbf{e}}$ erhalten wir

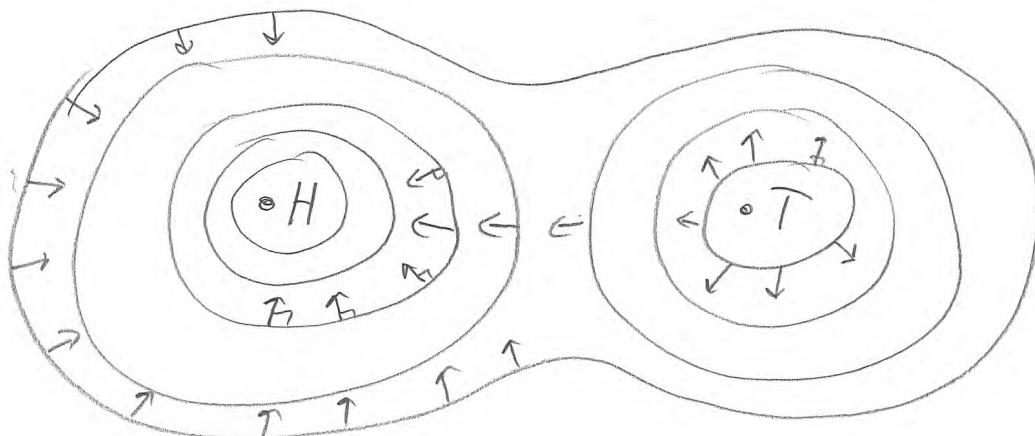
$$\underline{\nabla_{\hat{\mathbf{e}}} f = \langle \hat{\mathbf{e}}, \nabla f \rangle} = |\hat{\mathbf{e}}| \cdot |\nabla f| \cdot \cos(\alpha) = 1 \cdot |\nabla f| \cdot \cos(\alpha) = \underline{\cos(\alpha) \cdot |\nabla f|}. \quad (2.199)$$

Daraus folgen unmittelbar alle Aussagen und wir haben den Satz bewiesen. \square

Bemerkungen:

- i) Der Gradient ∇f zeigt an jedem Punkt der x - y -Ebene gerade in die Richtung, in welcher die Steigung des Graphen von f maximal ist.
- ii) Die Länge $|\nabla f|$ des Gradienten ist an jedem Punkt der x - y -Ebene gerade die maximale Steigung des Graphen von f .
- iii) Der Gradient steht an jedem Punkt der x - y -Ebene senkrecht auf der Level-Linie von f durch diesen Punkt.
- iv) Dort wo der Gradient verschwindet verläuft die Tangentialebene an den Graphen von f horizontal. Entsprechend hat f dort einen Hoch-Punkt, Tief-Punkt oder Sattel-Punkt.
- v) Definition und Eigenschaften sowohl des Gradienten als auch der Richtungsableitung gelten ganz analog auch in nD.

Zeichnet man den Gradienten und die Level-Linien von f in der x - y -Ebene ein, dann ergibt sich ein Bild wie im folgenden Beispiel.



2.7.2 Zweite Richtungsableitung

Wir betrachten den folgenden Satz.

Satz 2.26 HESSE-Formel in nD

Seien $n \in \mathbb{N}^+$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ eine zweifach differentierbare Funktion mit HESSE-Matrix $H := \nabla^2 f$ und $\hat{\mathbf{v}}, \hat{\mathbf{w}} \in \mathbb{R}^n$ zwei Einheitsvektoren, dann gilt

$$\nabla_{\hat{\mathbf{w}}\hat{\mathbf{v}}}^2 f = \nabla_{\hat{\mathbf{w}}}(\nabla_{\hat{\mathbf{v}}} f) = \langle \hat{\mathbf{w}}, H \cdot \hat{\mathbf{v}} \rangle. \quad (2.200)$$

Beweis: Siehe Übungen. □

Bemerkungen:

i) Für $\hat{\mathbf{w}} = \hat{\mathbf{v}}$ folgt

$$\nabla_{\hat{\mathbf{v}}\hat{\mathbf{v}}}^2 f = \langle \hat{\mathbf{v}}, H \cdot \hat{\mathbf{v}} \rangle. \quad (2.201)$$

Dies ist eine *quadratische Form* in $\hat{\mathbf{v}}$ mit H als *Matrix*.

ii) Wegen der *Symmetrien* von H und des *Skalar-Produkts kommutieren* die *zweiten Richtungsableiten*, d.h. es gilt

$$\nabla_{\hat{\mathbf{v}}\hat{\mathbf{w}}}^2 f = \nabla_{\hat{\mathbf{w}}\hat{\mathbf{v}}}^2 f. \quad (2.202)$$

iii) Für die *Standard-Einheitsvektoren* $\hat{\mathbf{e}}_1, \dots, \hat{\mathbf{e}}_n \in \mathbb{R}^n$ entlang der *Koordinatenachsen* gilt

$$\nabla_{\hat{\mathbf{e}}_\mu \hat{\mathbf{e}}_\nu}^2 f = f_{,\nu,\mu} = H_{\nu\mu}. \quad (2.203)$$

iv) Die *zweiten Richtungsableitungen* verschwinden genau dann für alle *Richtungen*, wenn gilt $H = 0$.

2.7.3 Lokale Extrema

Wir betrachten die folgende Definition.

Definition 2.27 Kritische Stelle

Seien $n \in \mathbb{N}^+$ und $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differentierbar, dann heisst ein Punkt $P \in \mathbb{R}^n$ *kritische Stelle* von f , wenn gilt

$$\nabla f(P) = 0. \quad (2.204)$$

Bemerkungen:

- i) Zur Bestimmung der *kritischen Stellen* von f muss das *Gleichungssystem* (2.204) gelöst werden.
- ii) An einer *kritischen Stelle* einer *Funktion* kann sich ein *Tief-Punkt*, ein *Hoch-Punkt* oder ein *Sattel-Punkt* befinden.
- iii) Für $n = 2$ verläuft die *Tangentialebene* an den *Graphen* von f an einer *kritischen Stelle horizontal*.

Wir betrachten den folgenden Satz.

Satz 2.27 *Lokale Extrema in nD*

Seien $n \in \mathbb{N}^+$, $f : \mathbb{R}^n \rightarrow \mathbb{R}$ zweifach differentierbar, $P \in \mathbb{R}^n$ eine kritische Stelle von f und $H = \nabla^2 f(P)$ mit Spektrum $\text{spec}(H) = \{\lambda_1, \dots, \lambda_n\} \subset \mathbb{R}$, dann gilt folgendes.

- (a) Falls $\lambda_1, \dots, \lambda_n < 0$, dann hat f bei P einen Hoch-Punkt.
- (b) Falls $\lambda_1, \dots, \lambda_n > 0$, dann hat f bei P einen Tief-Punkt.
- (c) Falls $\lambda_1, \dots, \lambda_n \neq 0$ mit unterschiedlichen Vorzeichen, dann hat f bei P einen Sattel-Punkt.

Wir betrachten den folgenden Satz.

Satz 2.28 *Lokale Extrema in 2D*

Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zweifach differentierbar, $(x_0; y_0) \in \mathbb{R}^2$ eine kritische Stelle von f und $H = \nabla^2 f(x_0; y_0)$, dann gilt folgendes.

- (a) Falls $\det(H) > 0$ und $H_{11}, H_{22} < 0$, dann hat f bei $(x_0; y_0)$ einen Hoch-Punkt.
- (b) Falls $\det(H) > 0$ und $H_{11}, H_{22} > 0$, dann hat f bei $(x_0; y_0)$ einen Tief-Punkt.
- (c) Falls $\det(H) < 0$, dann hat f bei $(x_0; y_0)$ einen Sattel-Punkt.

Beweis: Wir betrachten die Richtungsvektoren

$$\hat{\mathbf{e}}(t) = \frac{1}{\sqrt{1+t^2}} \cdot \begin{bmatrix} \pm 1 \\ t \end{bmatrix} \quad \text{für } t \in \mathbb{R}. \quad (2.205)$$

Durch die Richtungen gemäss (2.205) in Kombination mit den Richtungen die aus (2.205) durch einen Tausch der Komponenten hervorgehen, können offensichtlich alle möglichen Richtungen in 2D überlappend beschrieben werden. Für die zweite Richtungsableitung von f in die Richtungen $\hat{\mathbf{e}}(t)$ erhalten wir

$$\begin{aligned} \nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f &= \langle \hat{\mathbf{e}}, H \cdot \hat{\mathbf{e}} \rangle = \hat{\mathbf{e}}^T \cdot H \cdot \hat{\mathbf{e}} = \frac{1}{\sqrt{1+t^2}} \cdot \begin{bmatrix} \pm 1 \\ t \end{bmatrix}^T \cdot \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \cdot \frac{1}{\sqrt{1+t^2}} \cdot \begin{bmatrix} \pm 1 \\ t \end{bmatrix} \\ &= \frac{1}{1+t^2} \cdot [\pm 1 \ t] \cdot \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \cdot \begin{bmatrix} \pm 1 \\ t \end{bmatrix} = \frac{1}{1+t^2} \cdot [\pm 1 \ t] \cdot \begin{bmatrix} H_{11} \cdot (\pm 1) + H_{12} \cdot t \\ H_{21} \cdot (\pm 1) + H_{22} \cdot t \end{bmatrix} \\ &= \frac{1}{1+t^2} \cdot (\pm 1 \cdot (H_{11} \cdot (\pm 1) + H_{12} \cdot t) + t \cdot (H_{21} \cdot (\pm 1) + H_{22} \cdot t)) \\ &= \frac{1}{1+t^2} \cdot (H_{11} \pm H_{12} \cdot t \pm H_{21} \cdot t + H_{22} \cdot t^2) = \frac{1}{1+t^2} \cdot (H_{22} \cdot t^2 \pm 2 \cdot H_{12} \cdot t + H_{11}) \\ &=: \frac{1}{1+t^2} \cdot g(t) \end{aligned} \quad (2.206)$$

mit der quadratischen Funktion

$$g(t) = H_{22} \cdot t^2 \pm 2 \cdot H_{12} \cdot t + H_{11}. \quad (2.207)$$

Die Diskriminante von g ist

$$D = (\pm 2 \cdot H_{12})^2 - 4 \cdot H_{22} \cdot H_{11} = 4 \cdot H_{12}^2 - 4 \cdot H_{11} \cdot H_{22} = -4 \cdot (H_{11} \cdot H_{22} - H_{21} \cdot H_{12})$$

$$= -4 \cdot \det(H). \quad (2.208)$$

Dieses Ergebnis erhält man offensichtlich auch dann, wenn man in (2.205) die *Komponenten* von $\hat{\mathbf{e}}(t)$ vertauscht. Wir betrachten die Fälle der verschiedenen *Vorzeichen* von $\det(H)$ getrennt.

Fall 1: $\det(H) < 0$: In diesem Fall gilt $D > 0$, g und somit $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f$ haben zwei *reelle Nullstellen* und nehmen deshalb sowohl *strikt positive* als auch *strikt negative* Werte an. Demnach hat f bei $(x_0; y_0)$ einen Sattel-Punkt.

Fall 2: $\det(H) = 0$: In diesem Fall lassen sich aus H keine Schlüsse über den Typ der *kritischen Stelle* von f ziehen.

Fall 3: $\det(H) > 0$: In diesem Fall gilt $D < 0$, g und somit $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f$ haben keine *reelle Nullstellen* und nehmen deshalb entweder nur *strikt positive* oder nur *strikt negative* Werte an. Es muss also gelten

$$\operatorname{sgn}(\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f) = \operatorname{sgn}(\nabla_{\hat{\mathbf{e}}_1\hat{\mathbf{e}}_1}^2 f) = \operatorname{sgn}(\nabla_{\hat{\mathbf{e}}_2\hat{\mathbf{e}}_2}^2 f) = \operatorname{sgn}(H_{11}) = \operatorname{sgn}(H_{22}). \quad (2.209)$$

Wir betrachten die Fälle der verschiedenen *Vorzeichen* von H_{11} bzw. H_{22} getrennt.

Fall 3.1: $H_{11}, H_{22} < 0$: In diesem Fall gilt für alle *Richtungen* $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f < 0$. Demnach hat f bei $(x_0; y_0)$ einen Hoch-Punkt.

Fall 3.2: $H_{11}, H_{22} > 0$: In diesem Fall gilt für alle *Richtungen* $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f > 0$. Demnach hat f bei $(x_0; y_0)$ einen Tief-Punkt.

Damit haben wir alle Aussagen und den Satz bewiesen. □

Bemerkungen:

- i) Ob an einer *kritischen Stelle* ein *Hoch-Punkt*, *Tief-Punkt* oder *Sattel-Punkt* vorliegt, lässt sich also genau dann mit Hilfe von H entscheiden, wenn H regulär ist. Ist H singulär, dann lassen sich aus H keine Schlüsse über den Typ der *kritischen Stelle* ziehen.
- ii) In nD lässt sich anhand der *Vorzeichen* der *Eigenwerte* von H ablesen, ob $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f$ für alle *Richtungen* $\hat{\mathbf{e}}$ das gleiche *Vorzeichen* hat.
- iii) In 2D lässt sich anhand des *Vorzeichens* von $\det(H)$ ablesen, ob $\nabla_{\hat{\mathbf{e}}\hat{\mathbf{e}}}^2 f$ für alle *Richtungen* $\hat{\mathbf{e}}$ das gleiche *Vorzeichen* hat, denn H ist *symmetrisch* und somit gilt

$$\det(H) = \lambda_1 \cdot \lambda_2. \quad (2.210)$$

- iv) In 2D müssen im Fall $\det(H) > 0$ die *Komponenten* H_{11} und H_{22} das gleiche *Vorzeichen* haben. Anhand dieses *Vorzeichens* kann dann zwischen *Hoch-Punkt* und *Tief-Punkt* unterschieden werden.

Um die *lokalen Extrema* einer *Funktion* in 2D zu bestimmen, kann nach den folgenden Schritten vorgegangen werden.

S1 Die *kritische Stellen* von f sind zu bestimmen als *Lösungen* des *Gleichungssystems*

$$\nabla f(x; y) = 0. \quad (2.211)$$

S2 Für jede gefundene *kritische Stelle* $(x_k; y_k)$ sind zu berechnen

$$z_k = f(x_k; y_k) \quad (2.212)$$

$$H_k = \nabla^2 f(x_k; y_k). \quad (2.213)$$

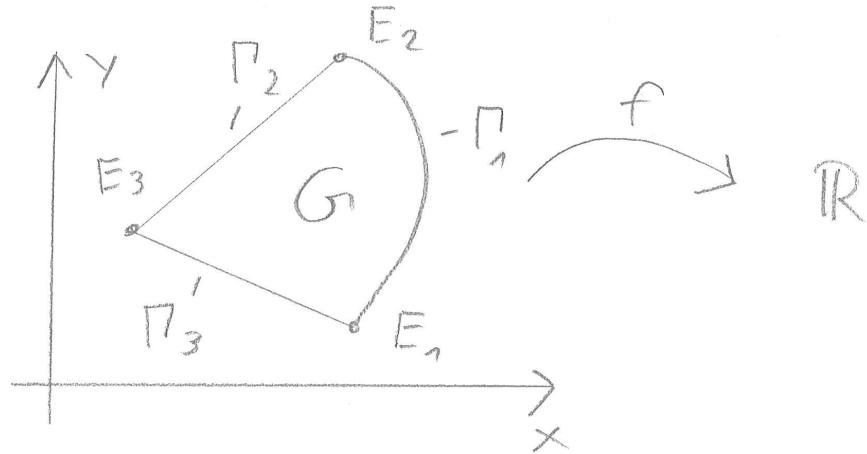
S3 Ergebnistabelle:

k	x_k	y_k	z_k	H_{k11}	H_{k22}	H_{k12}	$\det(H_k)$	Typ:
1	x_1	y_1	z_1
\vdots	\vdots	\vdots	\vdots					
n	x_n	y_n	z_n

(2.214)

2.7.4 Globale Extrema

Wir betrachten ein *Gebiet* $G \subset \mathbb{R}^2$ und eine *Funktion* $f : \mathbb{R}^2 \rightarrow \mathbb{R}$. Die Situation ist in der folgenden Skizze dargestellt.



Um die *globalen Extrema* einer *Funktion* in 2D zu bestimmen, kann nach den folgenden Schritten vorgegangen werden.

S1 Inneres: Kritische Stellen von f im Innern von G bestimmen als *Lösungen des Gleichungssystems*

$$\nabla f(x; y) = 0. \quad (2.215)$$

S2 Randstücke: Um die *Kandidaten* auf den *Randstücken* zu finden, kann nach den folgenden Schritten vorgegangen werden.

S 2.1: Die *Randstücke* Γ_k als *parametrisierte Kurven* beschreiben.

S 2.2: Die *Funktionen* $u_k(\tau) := f(x(\tau); y(\tau))$ für $\tau \in [\tau_0, \tau_E]$ berechnen.

S 2.3: Die *Lösungen* der *Gleichungen* $0 = \dot{u}_k(\tau)$ sind *Kandidaten* für *globale Extrema*.

S3 Eckpunkte: Die *Eckpunkte* des *Gebietes* G sind in jedem Fall *Kandidaten* für *globale Extrema*.

S4 Kandidatenvergleich: Die *Funktionswerte* der *Kandidaten* werden in einer Ergebnistabelle verglichen.

k	x_k	y_k	z_k	Typ:	
1	x_1	y_1	z_1	\dots	
\vdots	\vdots	\vdots	\vdots		
n	x_n	y_n	z_n	\dots	

(2.216)

2.7.5 Extrema unter Nebenbedingungen in 2D

Wir betrachten den folgenden Satz.

Satz 2.29 *Extrema unter Nebenbedingungen in 2D*

Seien $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}$ zwei differentierbare Funktionen und $(x; y) \in \mathbb{R}^2$ ein Extremum von f unter der Nebenbedingung

$$g(x; y) = 0. \quad (2.217)$$

Dann gibt es einen LAGRANGE-Multiplikator $\lambda \in \mathbb{R}$, so dass an der Stelle $(x; y)$ gilt

$$\boxed{\nabla f = \lambda \cdot \nabla g.} \quad (2.218)$$

Beweis: Wir betrachten die *Level-Linie* von g zum *Level* $L = 0$. Liegt ein *Extrema* von f entlang dieser *Level-Linie* am Punkt $(x; y)$ und zeigt $\hat{\mathbf{e}}$ tangential zur *Level-Linie*, dann muss gelten

$$0 = \nabla_{\hat{\mathbf{e}}} f = \langle \hat{\mathbf{e}}, \nabla f \rangle. \quad (2.219)$$

Bei $(x; y)$ steht der *Gradient* von f also *senkrecht* auf der *Level-Linie* von g und somit *parallel* oder *antiparallel* zum *Gradienten* von g . Es gibt daher einen LAGRANGE-Multiplikator $\lambda \in \mathbb{R}$, so dass (2.218) erfüllt ist. Damit haben wir den Satz bewiesen. \square

Um die *Extrema* einer *Funktion* in 2D unter *Nebenbedingungen* zu bestimmen, kann nach den folgenden Schritten vorgegangen werden.

S1 Es sind die *Kandidaten* als *Lösungen* des *Gleichungssystems* zu bestimmen gemäss

$$\begin{cases} \nabla f = \lambda \cdot \nabla g \\ g(x; y) = 0. \end{cases} \quad (2.220)$$

S2 Die *Funktionswerte* der *Kandidaten* sind in einer Ergebnistabelle zu vergleichen.

Bemerkungen:

- i) Das *Gleichungssystem* ist äquivalent zu den Bedingungen für die *kritischen Punkte* in 3D der LAGRANGE-Erweiterung

$$w(x; y; \lambda) = f(x; y) - \lambda \cdot g(x; y). \quad (2.221)$$

- ii) Das Verfahren lässt sich durch mehrere LAGRANGE-Multiplikatoren auf nD erweitern.

Kapitel 3

Integralrechnung

3.1 Integrationsmethoden

3.1.1 Substitution

Wir betrachten den folgenden Satz.

Satz 3.1 Integration durch Substitution

Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ eine integrierbare Funktion mit Stammfunktion $F : \mathbb{R} \rightarrow \mathbb{R}$, $u : \mathbb{R} \rightarrow \mathbb{R}$ eine differentierbare Funktion und $x_0, x_E \in \mathbb{R}$ mit $x_0 < x_E$, dann gilt folgendes.

$$(a) \quad \int f(u(x)) \cdot u'(x) dx = \int f(u) du = F(u(x)) + c$$

$$(b) \quad \int_{x_0}^{x_E} f(u(x)) \cdot u'(x) dx = \int_{u(x_0)}^{u(x_E)} f(u) du = F(u(x_E)) - F(u(x_0))$$

Beweis: Durch Anwenden der *Summen-Regel* und *Ketten-Regel* und weil $F' = f$ erhalten wir

$$(F(u(x)) + c)' = F'(u(x)) \cdot u'(x) + 0 = f(u(x)) \cdot u'(x). \quad (3.1)$$

Aus der NEWTON-LEIBNIZ-Formel folgt sofort (a) und durch Einsetzen der *Integrationsgrenzen* erhalten wir (b). Damit haben wir den Satz bewiesen. \square

Beispiele:

- Wir betrachten das *unbestimmte Integral*

$$F(x) = \int x \cdot \cos(x^2) dx. \quad (3.2)$$

Als *Substitution* wählen wir

$$u(x) := x^2 \Rightarrow u'(x) = 2x. \quad (3.3)$$

Wir zeigen mehrere Varianten, um das *unbestimmte Integral* zu berechnen.

Variante 1: Durch *strukturelle Ergänzung* erhalten wir

$$\begin{aligned}\underline{\underline{F(x)}} &= \int x \cdot \cos(x^2) dx = \int \frac{1}{2} \cdot 2x \cdot \cos(x^2) dx = \frac{1}{2} \int 2x \cdot \cos(x^2) dx \\ &= \frac{1}{2} \int u'(x) \cdot \cos(u(x)) dx = \frac{1}{2} \int \cos(u) du = \frac{1}{2} \cdot \sin(u) + c = \frac{1}{2} \cdot \sin(x^2) + c.\end{aligned}\tag{3.4}$$

Variante 2: Durch *Kalkulieren mit den Differentialsymbolen* erhalten wir

$$\frac{du}{dx} = 2x \Leftrightarrow du = 2x dx \Leftrightarrow dx = \frac{1}{2x} du\tag{3.5}$$

und somit

$$\begin{aligned}\underline{\underline{F(x)}} &= \int x \cdot \cos(x^2) dx = \int x \cdot \cos(u) \cdot \frac{1}{2x} du = \frac{1}{2} \int \cos(u) du = \frac{1}{2} \cdot \sin(u) + c \\ &= \frac{1}{2} \cdot \sin(x^2) + c.\end{aligned}\tag{3.6}$$

- Wir betrachten das *unbestimmte Integral*

$$F(x) = \int \tan(x) dx = \int \frac{\sin(x)}{\cos(x)} dx.\tag{3.7}$$

Als *Substitution* wählen wir

$$u(x) := \cos(x) \Rightarrow u'(x) = -\sin(x).\tag{3.8}$$

Durch *Kalkulieren mit den Differentialsymbolen* erhalten wir

$$\frac{du}{dx} = -\sin(x) \Leftrightarrow du = -\sin(x) dx \Leftrightarrow dx = -\frac{1}{\sin(x)} du\tag{3.9}$$

und somit

$$\begin{aligned}\underline{\underline{F(x)}} &= \int \tan(x) dx = \int \frac{\sin(x)}{\cos(x)} \cdot (-1) \cdot \frac{1}{\sin(x)} du = - \int \frac{1}{u} du = -\ln(|u|) + c \\ &= -\ln(|\cos(x)|) + c.\end{aligned}\tag{3.10}$$

- Wir betrachten das *bestimmte Integral*

$$I = \int_0^1 \sqrt{1-x^2} dx = \int_0^1 \sqrt{1-u^2} du.\tag{3.11}$$

Als *Substitution* wählen wir

$$u(\varphi) := \sin(\varphi) \Rightarrow u'(\varphi) = \cos(\varphi).\tag{3.12}$$

Durch *Kalkulieren mit den Differentialsymbolen* erhalten wir

$$\frac{du}{d\varphi} = \cos(\varphi) \Leftrightarrow du = \cos(\varphi) d\varphi\tag{3.13}$$

und somit

$$\begin{aligned} \underline{\underline{I}} &= \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2(\varphi)} \cdot \cos(\varphi) d\varphi = \int_0^{\frac{\pi}{2}} \sqrt{\cos^2(\varphi)} \cdot \cos(\varphi) d\varphi = \int_0^{\frac{\pi}{2}} |\cos(\varphi)| \cdot \cos(\varphi) d\varphi \\ &= \int_0^{\frac{\pi}{2}} \cos(\varphi) \cdot \cos(\varphi) d\varphi = \int_0^{\frac{\pi}{2}} \cos^2(\varphi) d\varphi = \dots = \underline{\underline{\frac{\pi}{4}}} . \end{aligned} \quad (3.14)$$

Das verbleibende *Integral* ist ein *Standard-Integral*, dass durch *partielle Integration* berechnet werden kann.

Bemerkungen:

- i) Die Idee hinter der *Integration* durch *Substitution* ist die Umkehrung der *Ketten-Regel* aus der *Differentialrechnung*.
- ii) Durch Anwenden der *Substitution* kann eine schwierige *Integration* auf eine einfachere *Integration* zurückgeführt werden.
- iii) Der Begriff *Substitution* bedeutet “Ersetzung”.
- iv) Bei der *Substitution* wird die *unabhängige Variable* gewechselt gemäss

$$x \mapsto u \quad \text{oder} \quad u \mapsto x. \quad (3.15)$$

- v) Das *Kalkulieren* mit den *Differentialsymbolen* ist analog zu den Berechnungsschritten in einem **ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess** gemäss

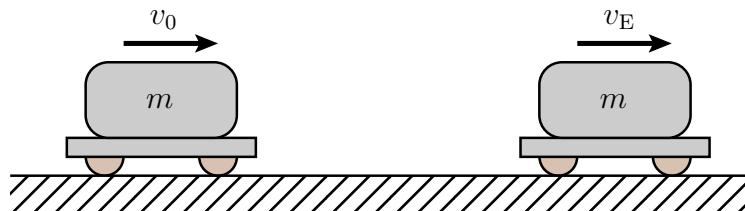
$$\delta I \approx f(u(x)) \cdot u'(x) \cdot \delta x \approx f(u(x)) \cdot \frac{\delta u}{\delta x} \cdot \delta x = f(u) \cdot \delta u. \quad (3.16)$$

- vi) Eine häufige Fehlerquelle bei der Berechnung von *bestimmten Integralen* durch *Substitution* ist das Anpassen der *Integrationsgrenzen*.
- vii) In den meisten Fällen aus der Praxis (ca. 90%) ist die *Substitution* u eine *lineare Funktion* der Form

$$u(x) = m \cdot x + q \Rightarrow \frac{du}{dx} = u'(x) = m \Leftrightarrow dx = \frac{1}{m} du. \quad (3.17)$$

Erkennt man dies vorweg, dann kann das einfachere Verfahren der *linearen Modifikation* durchgeführt werden.

Die Methode der *Substitution* hat nicht nur praktische Bedeutung bei der Berechnung von konkreten *Integralen* sondern auch bei der Herleitung von *Integralformeln*. Ein Standard-Beispiel ist die Formel für die *kinetische Energie* aus der *Physik*. Dazu betrachten wir einen Körper der *Masse* m , welcher von der *Anfangsgeschwindigkeit* v_0 auf die *Endgeschwindigkeit* v_E beschleunigt wird. Die Situation ist in der folgenden Skizze dargestellt.



Wir betrachten dazu den folgenden Satz.

Satz 3.2 Beschleunigungsarbeit

Um einen Körper der Masse m von der Anfangsgeschwindigkeit v_0 auf die Endgeschwindigkeit v_E zu beschleunigen, muss am Körper eine *Beschleunigungsarbeit* geleistet werden von

$$\Delta W = \frac{m}{2} \cdot v_E^2 - \frac{m}{2} \cdot v_0^2. \quad (3.18)$$

Beweis: Wir zeigen mehrere Varianten um die *Beschleunigungsarbeit* zu berechnen.

Variante 1: Mit Hilfe des *Arbeitsintegrals* und mehrfacher *Substitution* erhalten wir

$$\begin{aligned} \underline{\Delta W} &= \int_{s_0}^{s_E} F(s) \, ds = \int_{t_0}^{t_E} F(s(t)) \cdot \dot{s}(t) \, dt = \int_{t_0}^{t_E} m \cdot a(t) \cdot v(t) \, dt = m \int_{t_0}^{t_E} v(t) \cdot \dot{v}(t) \, dt \\ &= m \int_{v_0}^{v_E} v \, dv = m \cdot \frac{1}{2} \cdot [v^2] \Big|_{v_0}^{v_E} = \frac{m}{2} \cdot (v_E^2 - v_0^2) = \underline{\underline{\frac{m}{2} \cdot v_E^2 - \frac{m}{2} \cdot v_0^2}}. \end{aligned} \quad (3.19)$$

Variante 2: Wir verwenden einen ARCHIMEDES-CAUCHY-RIEMANN-Approximationsprozess.

Dabei gehen wir nach folgenden Schritten vor.

S1 Lokal: Wir betrachten ein kleines *Wegstück* δs . Die *Beschleunigungsarbeit* entlang δs ist

$$\underline{\delta W} \approx F(s) \cdot \delta s \approx m \cdot a(t) \cdot \frac{\delta s}{\delta t} \cdot \delta t \approx m \cdot \frac{\delta v}{\delta t} \cdot v \cdot \delta t = \underline{m \cdot v \cdot \delta v}. \quad (3.20)$$

S2 Global: Durch *Integration* über v erhalten wir die *Beschleunigungsarbeit*

$$\underline{\Delta W} = m \int_{v_0}^{v_E} v \, dv = m \cdot \frac{1}{2} \cdot [v^2] \Big|_{v_0}^{v_E} = \frac{m}{2} \cdot (v_E^2 - v_0^2) = \underline{\underline{\frac{m}{2} \cdot v_E^2 - \frac{m}{2} \cdot v_0^2}}. \quad (3.21)$$

Damit haben wir den Satz bewiesen. □

3.1.2 Partielle Integration

Wir betrachten den folgenden Satz.

Satz 3.3 Partielle Integration

Seien $g, h : \mathbb{R} \rightarrow \mathbb{R}$ differentierbare und integrierbare Funktionen und $x_0, x_E \in \mathbb{R}$ mit $x_0 < x_E$, dann gilt folgendes.

(a)
$$\int g(x) \cdot h'(x) \, dx = \downarrow g(x) \cdot h(x) \uparrow - \int g'(x) \cdot h(x) \, dx$$

(b)
$$\int_{x_0}^{x_E} g(x) \cdot h'(x) \, dx = \left[g(x) \cdot h(x) \right] \Big|_{x_0}^{x_E} - \int_{x_0}^{x_E} g'(x) \cdot h(x) \, dx$$

Beweis: Übung. □

Beispiele:

- Durch *partielle Integration* erhalten wir

$$\underline{\underline{F(x)}} = \int \overset{\downarrow}{x} \cdot \overset{\uparrow}{e^x} dx = x e^x - \int 1 \cdot e^x dx = \underline{\underline{x e^x - e^x + c}} = \underline{\underline{(x-1) e^x + c}} \quad (3.22)$$

- Durch *partielle Integration* und mit Hilfe des PYTHAGORAS-Satzes für *trigonometrische Funktionen* finden wir die Gleichung

$$\begin{aligned} F(x) &= \int \sin^2(x) dx = \int \overset{\downarrow}{\sin(x)} \cdot \overset{\uparrow}{\sin(x)} dx \\ &= \sin(x) \cdot (-\cos(x)) - \int \cos(x) \cdot (-\cos(x)) dx = -\sin(x) \cos(x) + \int \cos^2(x) dx \\ &= -\sin(x) \cos(x) + \int (1 - \sin^2(x)) dx = -\sin(x) \cos(x) + \int 1 dx - \int \sin^2(x) dx \\ &= -\sin(x) \cos(x) + x + b - F(x). \end{aligned} \quad (3.23)$$

Es gilt also

$$F(x) = -\sin(x) \cos(x) + x + b - F(x) \quad | + F(x) \quad (3.24)$$

$$2 \cdot F(x) = -\sin(x) \cos(x) + x + b \quad | : 2. \quad (3.25)$$

Daraus erhalten wir

$$\underline{\underline{F(x)}} = \frac{-\sin(x) \cos(x) + x + b}{2} = \frac{x - \sin(x) \cos(x)}{2} + c. \quad (3.26)$$

- Durch *partielle Integration* erhalten wir

$$\begin{aligned} \underline{\underline{F(x)}} &= \int \ln(x) dx = \int \overset{\downarrow}{\ln(x)} \cdot \overset{\uparrow}{1} dx = \ln(x) \cdot x - \int \ln'(x) \cdot x dx = x \cdot \ln(x) - \int \frac{1}{x} \cdot x dx \\ &= x \cdot \ln(x) - \int 1 dx = x \cdot \ln(x) - x + c = \underline{\underline{x \cdot (\ln(x) - 1) + c}} \end{aligned} \quad (3.27)$$

Bemerkungen:

- Die Idee hinter der *partiellen Integration* ist die Umkehrung der *Produkt-Regel* aus der *Differentialrechnung*.
- Durch Anwenden der *partiellen Integration* kann eine schwierige *Integration* auf eine einfache *Integration* zurückgeführt werden.
- Der Begriff *partielle Integration* bedeutet “teilweise *Integration*”.
- Eine häufige Fehlerquelle bei der Anwendung der *partiellen Integration* ist das *negative Vorzeichen* vor dem *Integral* auf der rechten Seite.

- v)** Es ist üblich und sinnvoll, vor Anwendung der *partiellen Integration* die Rollen der *Faktoren* im *Integranden* mit Pfeilen zu markieren.
- vi)** Beim *Aufleiten* von $h' \mapsto h$ darf eine beliebige *Stammfunktion* h von h' verwendet werden. Es ist jedoch zwingend nötig, dass auf der rechten Seite in beiden Termen die gleiche *Stammfunktion* h eingesetzt wird!
- vii)** Allgemein gilt

$$\int f(x) dx = \int f(x) \cdot 1 dx = f(x) \cdot x - \int f'(x) \cdot x dx. \quad (3.28)$$

Mit Hilfe dieses 1er-Tricks lassen sich u.a. die *Stammfunktionen* finden von

$$f \in \{\ln, \log_a, \arcsin, \arccos, \arctan, \operatorname{arccot}, \operatorname{arsinh}, \operatorname{arcosh}, \operatorname{artanh}, \operatorname{arcoth}\}. \quad (3.29)$$

Anwendungen:

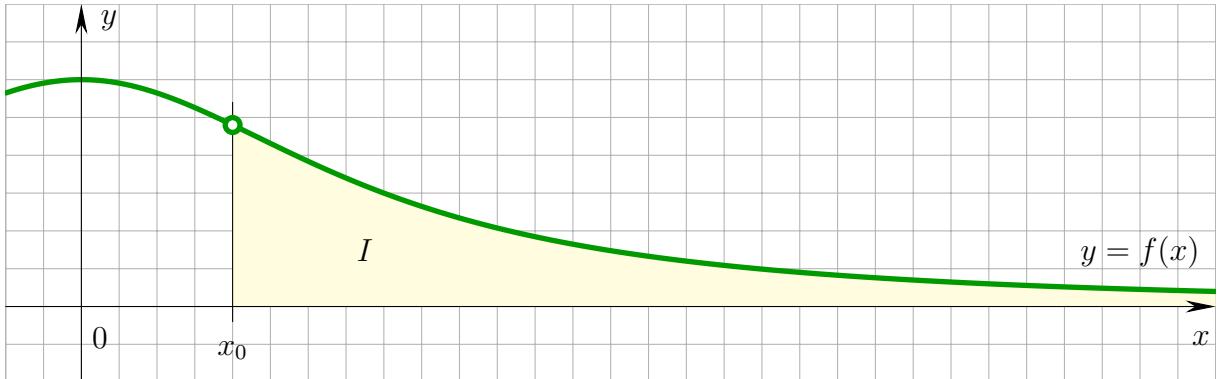
- Berechnung von *Integralen*
- TAYLOR-*Entwicklungen*
- Differentialgleichungen
- Variationsrechnung
- FOURIER- und LAPLACE-*Transformation*
- FEM-Simulationen

3.2 Uneigentliche Integrale

3.2.1 Integration über unendliche Intervalle

3.2.1.1 Integration über ein einseitig unendliches Intervall

Eine *Fläche* kann durch eine *unendlich lange Randkurve* begrenzt sein und trotzdem einen *endlichen Flächeninhalt* haben. Wir betrachten den *Graphen* einer *Funktion* $f : \mathbb{R} \rightarrow \mathbb{R}$, welche sich für $x \rightarrow \infty$ *asymptotisch* der x -Achse annähert. Eine solche Situation ist im folgenden x - y -Diagramm dargestellt.



Wir betrachten die folgende Definition.

Definition 3.1 *Uneigentliches Integral* mit einer *unendlichen Grenze*.

Seien $x_0, x_E \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ eine *integrierbare Funktion*. Die *uneigentlichen Integrale* von f Richtung $\pm\infty$ sind

$$(a) \quad \int_{x_0}^{\infty} f(x) dx := \lim_{s \rightarrow \infty} \int_{x_0}^s f(x) dx,$$

$$(b) \quad \int_{-\infty}^{x_E} f(x) dx := \lim_{s \rightarrow \infty} \int_{-s}^{x_E} f(x) dx,$$

falls die betreffenden *Grenzwerte* jeweils *konvergieren*.

Beispiele:

- Wir betrachten das *uneigentliche Integral*

$$\underline{I} = \int_2^{\infty} \frac{1}{x^2} dx = \lim_{s \rightarrow \infty} \int_2^s \frac{1}{x^2} dx = \lim_{s \rightarrow \infty} \left[-\frac{1}{x} \right]_2^s = \lim_{s \rightarrow \infty} \left(-\frac{1}{s} + \frac{1}{2} \right) = 0 + \frac{1}{2} = \underline{\underline{\frac{1}{2}}}. \quad (3.30)$$

- Wir betrachten das *uneigentliche Integral*

$$I = \int_2^{\infty} \frac{1}{x} dx = \lim_{s \rightarrow \infty} \int_2^s \frac{1}{x} dx = \lim_{s \rightarrow \infty} \ln\left(\frac{s}{2}\right) = \infty. \quad (3.31)$$

Dieses *uneigentliche Integral* ist divergent und existiert daher nicht.

- Wir betrachten das *uneigentliche Integral*

$$\begin{aligned} \underline{I} &= \int_{\ln(2)}^{\infty} e^{-x} dx = \lim_{s \rightarrow \infty} \int_{\ln(2)}^s e^{-x} dx = \lim_{s \rightarrow \infty} \left[-e^{-x} \right]_{\ln(2)}^s = \lim_{s \rightarrow \infty} \left(-e^{-s} + e^{-\ln(2)} \right) = 0 + e^{-\ln(2)} \\ &= e^{-\ln(2)} = \frac{1}{e^{\ln(2)}} = \underline{\underline{\frac{1}{2}}}. \end{aligned} \quad (3.32)$$

Bemerkungen:

- i) Uneigentliche Integrale sind als Grenzwerte definiert und existieren nur, wenn die betreffenden Grenzwerte konvergieren.
- ii) Bei der Berechnung eines uneigentlichen Integrals ist die Reihenfolge der Rechenschritte sehr wichtig. Zuerst muss eine Stammfunktion F von f gefunden und an den Grenzen x_0 und s bzw. $-s$ und x_E ausgewertet werden. Erst dann wird der Grenzwert $s \rightarrow \infty$ betrachtet. Dieses Vorgehen führt zu Rechnungen der Form

$$I = \int_{x_0}^{\infty} f(x) dx = \lim_{s \rightarrow \infty} \int_{x_0}^s f(x) dx = \lim_{s \rightarrow \infty} [F(x)] \Big|_{x_0}^s = \lim_{s \rightarrow \infty} (F(s) - F(x_0)) = \dots \quad (3.33)$$

- iii) Das uneigentliche Integral kann nur existieren, wenn sich der Integrand der x -Achse asymptotisch annähert, d.h. es muss gelten

$$\lim_{x \rightarrow \infty} f(x) = 0 \quad \text{bzw.} \quad \lim_{x \rightarrow -\infty} f(x) = 0. \quad (3.34)$$

- iv) Allein die Tatsache, dass sich der Integrand der x -Achse asymptotisch annähert reicht für die Existenz des uneigentlichen Integrals jedoch nicht aus! Es kommt vielmehr darauf an, wie "schnell" diese asymptotische Annäherung stattfindet.
- v) Beispiel-Codes zur Berechnung von uneigentlichen Integralen mit gängiger Software.

Mathematica/WolframAlpha	<code>Integrate[1/x^2, {x, 2, Infinity}]</code>
Python/Sympy	<code>import sympy as sp; sp.integrate(1/x**2, (x, 2, sp.oo));</code>

Wir betrachten den folgenden Satz.

Satz 3.4 Uneigentliches Integral von reziproken Potenzen

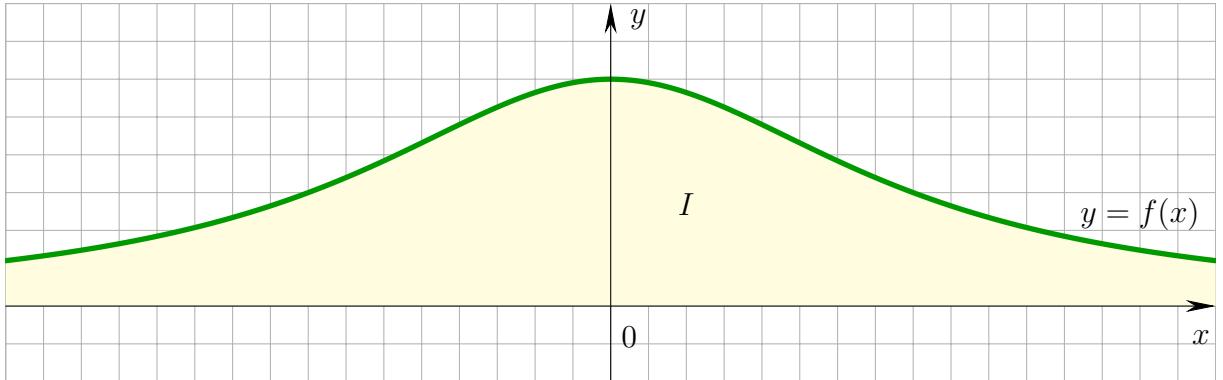
Seien $p, x_0 \in \mathbb{R}$ mit $x_0 > 0$, dann gilt

$$\int_{x_0}^{\infty} \frac{1}{x^p} dx = \begin{cases} \frac{1}{p-1} \cdot \frac{1}{x_0^{p-1}} & p > 1 \\ \text{divergent} & p \leq 1. \end{cases} \quad (3.35)$$

Beweis: Übung. □

3.2.1.2 Integration über ein beidseitig unendliches Intervall

Wir betrachten die gesamte Fläche zwischen dem Graphen einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ und der x -Achse. Eine solche Situation ist im folgenden x - y -Diagramm dargestellt.



Wir betrachten die folgende Definition.

Definition 3.2 Uneigentliches Integral mit zwei unendlichen Grenzen.

Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ eine integrierbare Funktion und $x_0 \in \mathbb{R}$. Das uneigentliche Integral von f über die reellen Zahlen ist

$$\int_{-\infty}^{\infty} f(x) dx := \lim_{r \rightarrow \infty} \int_{-r}^{x_0} f(x) dx + \lim_{s \rightarrow \infty} \int_{x_0}^s f(x) dx, \quad (3.36)$$

falls beide Grenzwerte konvergieren.

Bemerkungen:

- i) Das uneigentliche Integral wird als Summe von zwei Grenzwerten berechnet und existiert somit nur dann, wenn beide einzeln konvergieren.
- ii) Für die Aufteilung kann ein beliebiges $x_0 \in \mathbb{R}$ gewählt werden.
- iii) Falls das uneigentliche Integral existiert, dann konvergiert auch der zweiseitige Grenzwert und es gilt

$$I = \int_{-\infty}^{\infty} f(x) dx = \lim_{s \rightarrow \infty} \int_{-s}^s f(x) dx. \quad (3.37)$$

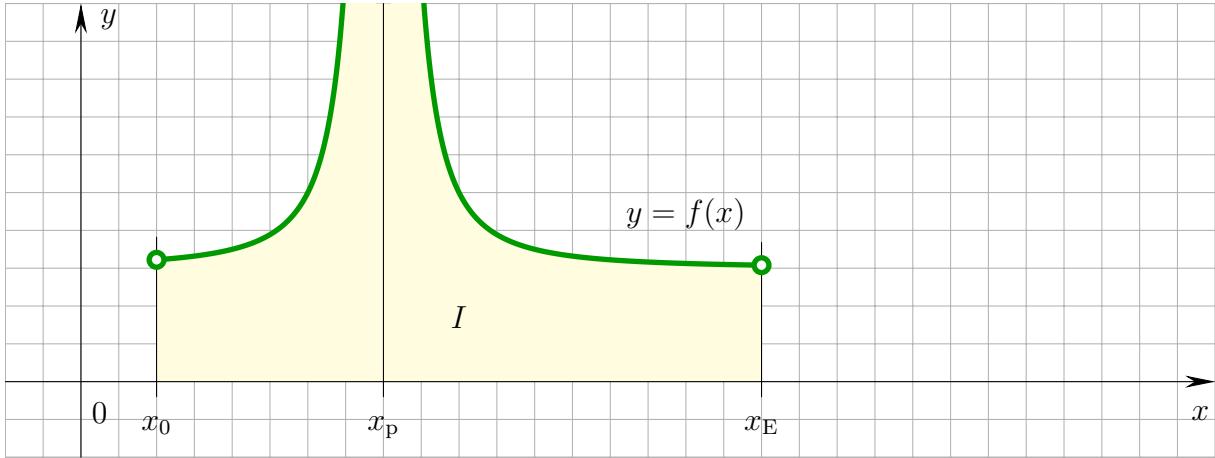
- iv) Aus der Konvergenz des zweiseitigen Grenzwertes folgt jedoch nicht die Existenz des uneigentlichen Integrals! In solchen Fällen führt das Verwenden des zweiseitigen Grenzwertes zu Fehlschlüssen und falschen Ergebnissen, wie die folgenden Beispiele zeigen.

$$\begin{aligned} I &= \int_{-\infty}^{\infty} x dx = \lim_{s \rightarrow \infty} \int_{-s}^s x dx = \lim_{s \rightarrow \infty} \frac{1}{2} \cdot [x^2] \Big|_{-s}^s = \lim_{s \rightarrow \infty} \frac{1}{2} \cdot (s^2 - (-s)^2) \\ &= \frac{1}{2} \cdot \lim_{s \rightarrow \infty} 0 = 0 \quad \text{konvergent!} \end{aligned} \quad (3.38)$$

$$\begin{aligned} I &= \int_{-\infty}^{\infty} x dx = \lim_{s \rightarrow \infty} \int_{-s}^{s+1} x dx = \lim_{s \rightarrow \infty} \frac{1}{2} \cdot [x^2] \Big|_{-s}^{s+1} = \lim_{s \rightarrow \infty} \frac{1}{2} \cdot ((s+1)^2 - (-s)^2) \\ &= \frac{1}{2} \cdot \lim_{s \rightarrow \infty} (s^2 + 2s + 1 - s^2) = \frac{1}{2} \cdot \lim_{s \rightarrow \infty} (2s + 1) = \infty \quad \text{divergent!} \end{aligned} \quad (3.39)$$

3.2.2 Integration über Polstellen

Wir betrachten die *Fläche* zwischen dem *Graphen* einer *Funktion* $f : \mathbb{R} \setminus \{x_p\} \rightarrow \mathbb{R}$ und einer *vertikalen Asymptote* bei $x_p \in \mathbb{R}$. Eine solche Situation ist im folgenden *x-y-Diagramm* dargestellt.



Wir betrachten die folgende Definition.

Definition 3.3 *Uneigentliches Integral* über eine *Polstelle*.

Seien $x_0, x_p, x_E \in \mathbb{R}$ mit $x_0 \leq x_p \leq x_E$ und $f : \mathbb{R} \setminus \{x_p\} \rightarrow \mathbb{R}$ eine *integrierbare Funktion*. Das *uneigentliche Integral* von f über die *Polstelle* x_p ist

$$\int_{x_0}^{x_E} f(x) dx := \lim_{r \nearrow x_p} \int_{x_0}^r f(x) dx + \lim_{s \searrow x_p} \int_s^{x_E} f(x) dx, \quad (3.40)$$

falls beide *Grenzwerte konvergieren*.

Bemerkungen:

- i) Das *uneigentliche Integral* wird als *Summe* von zwei *Grenzwerten* berechnet und existiert somit nur dann, wenn beide einzeln *konvergieren*.
- ii) Falls $x_p \in \{x_0, x_E\}$, dann fällt einer der beiden *Grenzwerte* weg. Es gilt

$$x_p = x_0 \Rightarrow \int_{x_0}^{x_E} f(x) dx = \int_{x_p}^{x_E} f(x) dx = \lim_{s \searrow x_p} \int_s^{x_E} f(x) dx \quad (3.41)$$

$$x_p = x_E \Rightarrow \int_{x_0}^{x_E} f(x) dx = \int_{x_0}^{x_p} f(x) dx = \lim_{r \nearrow x_p} \int_{x_0}^r f(x) dx. \quad (3.42)$$

- iii) Befinden sich mehr als eine *Polstelle* im *Intervall* $[x_0, x_E]$, dann muss das *Integral* in *Teilabschnitte* unterteilt und weitere *Grenzwerte* eingeführt werden.
- iv) In der Praxis ist aus *Integrand* und *Integrationsgrenzen* nicht immer auf den ersten Blick ersichtlich, ob es sich um ein *uneigentliches Integral* mit *Polstellen* handelt, wie z.B. bei

$$I = \int_0^3 \frac{2x - 4}{1 - e^{\frac{2-x}{3}}} dx \quad (3.43)$$

mit einem *Pol* bei $x = 2$.

Kapitel 4

Taylor-Entwicklungen

4.1 Maclaurin-Entwicklungen

4.1.1 Maclaurin-Formel

Wir betrachten den folgenden Satz.

Satz 4.1 MACLAURIN-Entwicklung

Seien $n \in \mathbb{N}$, $x \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ unendlich oft differentierbar. Dann gilt

$$f(x) = T_n(x) + R_n(x), \quad (4.1)$$

mit dem MACLAURIN-Polynom $T_n(x)$ und Restglied $R_n(x)$ gemäss

$$\boxed{T_n(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} \cdot x^k = f(0) + f'(0) \cdot x + \frac{f''(0)}{2!} \cdot x^2 + \dots + \frac{f^{(n)}(0)}{n!} \cdot x^n}$$
$$R_n(x) = \frac{(-1)^n}{n!} \int_0^x f^{(n+1)}(s) \cdot (s-x)^n \, ds. \quad (4.2)$$

Beweis: Zunächst bemerken wir, dass gilt

$$f(x) - f(0) = [f(s)] \Big|_0^x = \int_0^x f'(s) \, ds \quad | + f(0). \quad (4.3)$$

Daraus und durch n -fache partielle Integration folgt

$$\begin{aligned} \underline{\underline{f(x)}} &= f(0) + \int_0^x f'(s) \, ds = f(0) + \int_0^x f'(s) \cdot 1 \, ds \\ &= f(0) + [f'(s) \cdot (s-x)] \Big|_0^x - \int_0^x f''(s) \cdot (s-x) \, ds \\ &= f(0) + f'(0) \cdot x - \left[f''(s) \cdot \frac{(s-x)^2}{2} \right] \Big|_0^x + \int_0^x f'''(s) \cdot \frac{(s-x)^2}{2} \, ds \\ &= f(0) + f'(0) \cdot x + \frac{f''(0)}{2} \cdot x^2 + \left[f'''(s) \cdot \frac{(s-x)^3}{2 \cdot 3} \right] \Big|_0^x - \int_0^x f^{(4)}(s) \cdot \frac{(s-x)^3}{2 \cdot 3} \, ds \end{aligned}$$

$$\begin{aligned}
&= \frac{f(0)}{0!} + \frac{f'(0)}{1!} \cdot x + \frac{f''(0)}{2!} \cdot x^2 + \dots + \frac{f^{(n)}(0)}{n!} \cdot x^n + \frac{(-1)^n}{n!} \int_0^x f^{(n+1)}(s) \cdot (s-x)^n \, ds \\
&= \underline{\underline{T_n(x)}} + \underline{\underline{R_n(x)}}.
\end{aligned} \tag{4.4}$$

Damit haben den Satz bewiesen. \square

Besonders interessant ist die Situation, wenn das *Restglied* für grosse n immer kleiner wird. Dazu betrachten wir die folgende Definition.

Definition 4.1 *Analytische Funktion*

Seien $I \subseteq \mathbb{R}$ ein Intervall mit $0 \in I$ und $f : I \rightarrow \mathbb{R}$ eine unendlich oft differentierbare Funktion. Die Funktion f heisst *analytisch* auf I , falls für alle $x \in I$ gilt

$$\lim_{n \rightarrow \infty} R_n(x) = 0. \tag{4.5}$$

Bemerkungen:

- i) Das MACLAURIN-Polynom $T_n(x)$ ist ein Polynom vom Grad n .
- ii) Um das MACLAURIN-Polynom $T_n(x)$ aufzustellen, müssen die Funktionswerte von f und ihren Ableitungen nur an der Stelle $x = 0$ bekannt sein.
- iii) Eine *analytische Funktion* lässt sich durch ihre MACLAURIN-Reihe vollständig darstellen, d.h. für alle $x \in I$ gilt

$$\underline{\underline{f(x)}} = \lim_{n \rightarrow \infty} T_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} \cdot x^k. \tag{4.6}$$

- iv) In jedem Fall, d.h. auch für *nicht analytische Funktionen* kann das MACLAURIN-Polynom in der Nähe von $x = 0$ als Näherung für f verwendet werden, d.h. für x nahe genug bei 0 und n gross genug gilt zumindest

$$f(x) = T_n(x) + R_n(x) \approx T_n(x). \tag{4.7}$$

- v) Hat f eine Parität, dann hat das MACLAURIN-Polynom die gleiche Parität. In jedem Fall gilt

$$f \text{ hat positive Parität} \Leftrightarrow T_n(x) \text{ enthält nur gerade Potenzen von } x, \tag{4.8}$$

$$f \text{ hat negative Parität} \Leftrightarrow T_n(x) \text{ enthält nur ungerade Potenzen von } x. \tag{4.9}$$

- vi) Ist f selbst ein Polynom vom Grad $p \in \mathbb{N}$, dann ist f auf ganz \mathbb{R} *analytisch* und es gilt

$$T_n(x) = f(x) \quad \text{für alle } n \geq p. \tag{4.10}$$

- vii) Beispiel-Codes zur Berechnung von MACLAURIN-Entwicklungen mit gängiger Software.

Mathematica/WolframAlpha	<code>Series[Exp[x], {x, 0, n}]</code>
Python/Sympy	<code>import sympy as sp;</code> <code>sp.series(sp.exp(x), x, 0, n+1);</code>

4.1.2 Maclaurin-Reihen der Elementarfunktionen

Die meisten bekannten *Elementarfunktionen* können durch MACLAURIN-*Reihen* dargestellt werden. Wir betrachten dazu den folgenden Satz.

Satz 4.2 MACLAURIN-Reihen der Elementarfunktionen

Die folgenden *Elementarfunktionen* sind auf ganz \mathbb{R} *analytisch* und für alle $x \in \mathbb{R}$ gelten die folgenden Darstellungen durch MACLAURIN-*Reihen*.

$$\begin{aligned}
 \exp(x) &= \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots \\
 \sin(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{(2k+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{x^9}{9!} - \dots \\
 \cos(x) &= \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k}}{(2k)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \frac{x^8}{8!} - \dots \\
 \sinh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \frac{x^9}{9!} + \dots \\
 \cosh(x) &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \frac{x^8}{8!} + \dots
 \end{aligned} \tag{4.11}$$

Beweis: Wir berechnen die *Ableitungen* von $f(x) := \exp(x) = e^x$ sowie deren Werte an der Stelle $x_0 = 0$ und stellen die Resultate in der folgenden Tabelle zusammen.

k	0	1	2	3	4	5	6	...
$f^{(k)}(x)$	$\exp(x)$...						
$f^{(k)}(0)$	1	1	1	1	1	1	1	...

(4.12)

Durch Einsetzen in die Formel der MACLAURIN-*Entwicklung* für ein $n \in \mathbb{N}$ und ein $x \in \mathbb{R}$ erhalten wir

$$\begin{aligned}
 f(x) &= f(0) + \frac{f'(0)}{1!} \cdot x + \frac{f''(0)}{2!} \cdot x^2 + \frac{f'''(0)}{3!} \cdot x^3 + \dots + \frac{f^{(n)}(0)}{n!} \cdot x^n + R_n(x) \\
 &= 1 + \frac{1}{1!} \cdot x + \frac{1}{2!} \cdot x^2 + \frac{1}{3!} \cdot x^3 + \frac{1}{4!} \cdot x^4 + \frac{1}{5!} \cdot x^5 + \frac{1}{6!} \cdot x^6 + \frac{1}{7!} \cdot x^7 + \dots + R_n(x) \\
 &= \sum_{k=0}^n \frac{x^k}{k!} + R_n(x).
 \end{aligned} \tag{4.13}$$

Es sei $p := \operatorname{sgn}(x)$, dann finden wir für das *Restglied* die *Abschätzung*

$$\begin{aligned}
 |R_n(x)| &= \left| \frac{(-1)^n}{n!} \int_0^x f^{(n+1)}(s) \cdot (s-x)^n ds \right| = \left| \frac{(-1)^n}{n!} \int_0^x e^s \cdot (s-x)^n ds \right| \\
 &= \frac{1}{n!} \cdot \left| \int_0^x e^s \cdot (s-x)^n ds \right| \leq \frac{p}{n!} \int_0^x |e^s \cdot (s-x)^n| ds = \frac{p}{n!} \int_0^x |e^s| \cdot |(s-x)^n| ds
 \end{aligned}$$

$$\begin{aligned}
&= \frac{p}{n!} \int_0^x e^s \cdot |s - x|^n ds \leq \frac{p}{n!} \int_0^x e^{|x|} \cdot |0 - x|^n ds = \frac{p}{n!} \int_0^x e^{|x|} \cdot |x|^n ds \\
&= e^{|x|} \cdot |x|^n \cdot \frac{p}{n!} \int_0^x 1 ds = e^{|x|} \cdot \frac{|x|^n}{n!} \cdot p \cdot [s] \Big|_0^x = e^{|x|} \cdot \frac{|x|^n}{n!} \cdot p \cdot (x - 0) \\
&= e^{|x|} \cdot \frac{|x|^n}{n!} \cdot p \cdot x = e^{|x|} \cdot |x| \cdot \frac{|x|^n}{n!} \xrightarrow{n \rightarrow \infty} e^{|x|} \cdot |x| \cdot 0 = 0.
\end{aligned} \tag{4.14}$$

Somit ist \exp auf ganz \mathbb{R} analytisch und kann dargestellt werden durch die MACLAURIN-Reihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \dots \tag{4.15}$$

Siehe Übungen für die andern *Funktionen*. □

Bemerkungen:

- i) Es ist heute allgemein üblich, diese *Elementarfunktionen* durch ihre MACLAURIN-Reihen zu definieren.
- ii) Die MACLAURIN-Entwicklungen dieser *Elementarfunktionen* sind in modernen Taschenrechnern, PCs und auch Grossrechnern hardwareseitig implementiert.
- iii) Die MACLAURIN-Reihe der *natürlichen Exponentialfunktion* wird *Exponentialreihe* genannt. Durch Einsetzen von $x = 1$ erhält man für die EULER-Zahl die bekannte *Reihendarstellung*

$$e = \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \frac{1}{5!} + \dots \tag{4.16}$$

- iv) Die MACLAURIN-Entwicklungen dürfen Termweise *abgeleitet* werden, wordurch die bekannten *Ableitungsregeln* offensichtlich werden.
- v) Die MACLAURIN-Entwicklungen zeigen deutlich die Verwandtschaft zwischen *trigonometrischen* und *hyperbolischen Funktionen*.

4.1.3 Anwendungen

MACLAURIN-Entwicklungen haben zahlreiche Anwendungen.

- Definition von *Funktionen*
- Numerische Berechnung von *Funktionswerten*
- Numerische Näherungen von *Funktionswerten*
- Vereinfachungen von *Funktionen* in der Nähe von 0 (z.B. *Kleinwinkelnäherungen*)
- Linearisierungen von *Differentialgleichungen* (z.B. in der *Regelungstechnik*)
- Untersuchung einer *Funktion* in der Nähe von 0
- Untersuchung von Beziehungen zwischen *Funktionen* (z.B. *EULER-Formel*)
- Übergänge zwischen Modellen (z.B. *Relativitätstheorie* und *NEWTON-Mechanik*)
- Informationstheorie

4.2 Taylor-Entwicklungen allgemein

4.2.1 Taylor-Formel

Wir betrachten den folgenden Satz.

Satz 4.3 TAYLOR-Entwicklung

Seien $n \in \mathbb{N}$, $x, x_0 \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ unendlich oft differentierbar. Dann gilt

$$f(x) = T_n(x) + R_n(x), \quad (4.17)$$

mit dem TAYLOR-Polynom $T_n(x)$ und Restglied $R_n(x)$ gemäss

$$\begin{aligned} T_n(x) &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k \\ &= f(x_0) + f'(x_0) \cdot (x - x_0) + \frac{f''(x_0)}{2!} \cdot (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} \cdot (x - x_0)^n \\ R_n(x) &= \frac{(-1)^n}{n!} \int_{x_0}^x f^{(n+1)}(s) \cdot (s - x)^n \, ds. \end{aligned} \quad (4.18)$$

Beweis: Analog zum Beweis der MACLAURIN-Entwicklung. \square

Besonders interessant ist die Situation, wenn das Restglied für grosse n immer kleiner wird. Dazu betrachten wir die folgende Definition.

Definition 4.2 Analytische Funktion

Seien $I \subseteq \mathbb{R}$ ein Intervall mit $x_0 \in I$ und $f : I \rightarrow \mathbb{R}$ eine unendlich oft differentierbare Funktion. Die Funktion f heisst *analytisch* auf I , falls für alle $x \in I$ gilt

$$\lim_{n \rightarrow \infty} R_n(x) = 0. \quad (4.19)$$

Bemerkungen:

- i) Die TAYLOR-Entwicklung für $x_0 = 0$ ist gerade die MACLAURIN-Entwicklung.
- ii) Das TAYLOR-Polynom $T_n(x)$ ist ein Polynom vom Grad n .
- iii) Um das TAYLOR-Polynom $T_n(x)$ aufzustellen, müssen die Funktionswerte von f und ihren Ableitungen nur an der Stelle x_0 bekannt sein.
- iv) Eine analytische Funktion lässt sich durch ihre TAYLOR-Reihe vollständig darstellen, d.h. für alle $x \in I$ gilt

$$\underline{\underline{f(x)}} = \lim_{n \rightarrow \infty} T_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k. \quad (4.20)$$

- v) In jedem Fall, d.h. auch für nicht analytische Funktionen kann das TAYLOR-Polynom in der Nähe von $x = x_0$ als Näherung für f verwendet werden, d.h. für x nahe genug bei x_0 und n gross genug gilt zumindest

$$f(x) = T_n(x) + R_n(x) \approx T_n(x). \quad (4.21)$$

vi) Das TAYLOR-Polynom für $n = 1$ heisst auch *Linearisierung* von f an der Stelle x_0 . Es gilt

$$T_1(x) = f(x_0) + f'(x_0) \cdot (x - x_0). \quad (4.22)$$

Der Graph von T_1 ist offensichtlich die *Tangente* an den *Funktionsgraphen* von f an der Stelle x_0 .

vii) Ist f selbst ein *Polynom* vom *Grad* $p \in \mathbb{N}$, dann ist f auf ganz \mathbb{R} *analytisch* und es gilt

$$T_n(x) = f(x) \quad \text{für alle } n \geq p. \quad (4.23)$$

viii) Beispiel-Codes zur Berechnung von TAYLOR-Entwicklungen mit gängiger Software.

Mathematica/WolframAlpha	<code>Series[Exp[x], {x, x_0, n}]</code>
Python/Sympy	<code>import sympy as sp;</code> <code>sp.series(sp.exp(x), x, x_0, n+1);</code>

4.2.2 Anwendung lokale Extrema

Wir betrachten den folgenden Satz aus der *Kurvendiskussion*.

Satz 4.4 *Lokale Extrema* - erweiterte Kriterien

Seien $f : \mathbb{R} \rightarrow \mathbb{R}$ eine genügend oft *differentierbare, reelle Funktion*, $x_k \in \mathbb{R}$ und $m \in \mathbb{N}^+ \setminus \{1\}$, so dass

$$0 = f'(x_k) = f''(x_k) = \dots = f^{(m-1)}(x_k) \quad (4.24)$$

$$0 \neq f^{(m)}(x_k). \quad (4.25)$$

Dann gilt folgendes.

(a) Falls m gerade und $f^{(m)}(x_k) < 0$, dann hat f bei x_k einen *Hoch-Punkt*.

(b) Falls m gerade und $f^{(m)}(x_k) > 0$, dann hat f bei x_k einen *Tief-Punkt*.

(c) Falls m ungerade, dann hat f bei x_k einen *Sattel-Punkt*.

Beweis: Wir betrachten die TAYLOR-Entwicklung von f der *Ordnung* m an der Stelle x_k . Für alle $x \in \mathbb{R}$ nahe genug bei x_k gilt

$$\underline{f(x) = T_m(x) + R_m(x)} \approx T_m(x) = f(x_k) + \frac{f^{(m)}(x_k)}{m!} \cdot (x - x_k)^m. \quad (4.26)$$

Wir betrachten die Fälle m gerade und m ungerade getrennt.

Fall 1: m gerade. In diesem Fall hat T_m und somit auch f bei x_k ein *lokales Extremum*. Aus dem Vorzeichen des Faktors vor der Potenz $(x - x_k)^m$ lässt sich der Typ des *lokalen Extremums* ablesen. Es gilt

$$\underline{f^{(m)}(x_k) < 0} \Rightarrow T_m \text{ und somit auch } f \text{ hat bei } x_k \text{ einen } \underline{\text{Hoch-Punkt}}, \quad (4.27)$$

$$\underline{f^{(m)}(x_k) > 0} \Rightarrow T_m \text{ und somit auch } f \text{ hat bei } x_k \text{ einen } \underline{\text{Tief-Punkt}}. \quad (4.28)$$

Fall 2: m ungerade. In diesem Fall hat T_m und somit auch f bei x_k einen *Sattel-Punkt*.

Damit haben wir den Satz bewiesen. □

Teil II

Lineare Algebra

Kapitel 5

Komplexe Zahlen

5.1 Grundlagen

5.1.1 Einleitung

Bekanntlich gibt es kein *reelles* $x \in \mathbb{R}$ mit $x^2 < 0$. Das bedeutet, für $a < 0$ hat die *Gleichung*

$$x^2 = a \tag{5.1}$$

keine *Lösungen*. Die Idee ist nun, eine noch grössere *Zahlenmenge* zu konstruieren, so dass auch alle *negative reellen Zahlen* als *Quadrat*e geschrieben werden können.

5.1.2 Konstruktion

Wir betrachten folgende Definition.

Definition 5.1 *Komplexe Zahlen*

Die *komplexen Zahlen* \mathbb{C} sind die kleinste *Menge*, welche die folgenden Eigenschaften erfüllt.

A1 $\mathbb{R} \subseteq \mathbb{C}$.

A2 $(\mathbb{C}; +; \cdot)$ bildet einen *Zahlenkörper*.

A3 Es gibt ein $i \in \mathbb{C}$ mit $i^2 = -1$.

Bemerkungen:

- i)** Die *Zahl* i heisst *imaginäre Einheit*.
- ii)** Weil in der *Elektrotechnik* i schon die *elektrische Stromstärke* bezeichnet, wird in der Literatur oft auch ein j verwendet.
- iii)** Alle *Elemente* von \mathbb{C} lassen sich durch *reelle Zahlen* und i beschreiben.
- iv)** Man kann zeigen, dass \mathbb{C} der grösstmögliche *Zahlenkörper* ist.
- v)** Man kann zeigen, dass \mathbb{C} *algebraisch abgeschlossen* ist und somit kein Bedarf für eine noch grössere *Zahlenmenge* besteht.

5.1.3 Zahenumfang

Nach Konstruktion enthält \mathbb{C} offensichtlich die folgenden *Zahlen*.

- 1)** alle $x \in \mathbb{R}$
- 2)** $i \in \mathbb{C} \setminus \mathbb{R}$
- 3)** alle *reellen Vielfachen* von i , d.h. $y \cdot i$ mit $y \in \mathbb{R}$
- 4)** alle *Linearkombinationen* der Form $z = x + y \cdot i$ mit $x, y \in \mathbb{R}$

Tatsache ist, dass diese Aufzählung bereits alle *komplexen Zahlen* umfasst. Wir betrachten dazu den folgenden Satz.

Satz 5.1 Arithmetische Form

Für jedes $z \in \mathbb{C}$ gibt es eindeutige $x, y \in \mathbb{R}$, so dass gilt

$$z = x + y \cdot i. \quad (5.2)$$

Bemerkungen:

- i)** Jede *komplexe Zahl* lässt sich darstellen durch 2 *reelle Zahlen*. Somit muss \mathbb{C} eine ähnliche Struktur haben wie \mathbb{R}^2 .
- ii)** Für $y = 0$ erhält man gerade alle *reellen Zahlen*.
- iii)** Die *reellen Vielfachen* von i sind die gesuchten neuen *Zahlen*, deren *Quadrat* einen *negativen reellen Wert* haben. Beispiele:

$$z = 2 \cdot i \Rightarrow z^2 = (2 \cdot i)^2 = 2^2 \cdot i^2 = 4 \cdot (-1) = -4 \quad (5.3)$$

$$z = \sqrt{3} \cdot i \Rightarrow z^2 = (\sqrt{3} \cdot i)^2 = 3 \cdot (-1) = -3 \quad (5.4)$$

$$z = y \cdot i \Rightarrow z^2 = (y \cdot i)^2 = y^2 \cdot (-1) = -y^2. \quad (5.5)$$

Wir betrachten folgende Definition.

Definition 5.2 Weitere Bezeichnungen

Seien $z \in \mathbb{C}$ und $x, y \in \mathbb{R}$ mit

$$z = x + y \cdot i. \quad (5.6)$$

Es werden die folgenden Bezeichnungen verwendet.

(a) *Realteil* von z

(c) *Betrag* von z

$$\operatorname{Re}(z) := x.$$

$$|z| := \sqrt{x^2 + y^2}.$$

(b) *Imaginärteil* von z

(d) *Komplex-Konjugierte* von z

$$\operatorname{Im}(z) := y.$$

$$z^* := x - y \cdot i.$$

5.1.4 Grundoperationen

Wir betrachten $x_1, y_1, x_2, y_2 \in \mathbb{R}$ und die beiden *komplexen Zahlen*

$$z_1 = x_1 + y_1 \cdot i \quad \text{und} \quad z_2 = x_2 + y_2 \cdot i. \quad (5.7)$$

Beim Durchführen der *algebraischen Grundoperationen* erhalten wir die folgenden Ergebnisse:

Addition: Es gilt

$$\underline{\underline{z_1 + z_2}} = x_1 + y_1 \cdot i + x_2 + y_2 \cdot i = x_1 + x_2 + y_1 \cdot i + y_2 \cdot i = \underline{\underline{x_1 + x_2 + (y_1 + y_2) \cdot i}}. \quad (5.8)$$

Subtraktion: Es gilt

$$\underline{\underline{z_1 - z_2}} = x_1 + y_1 \cdot i - x_2 - y_2 \cdot i = x_1 - x_2 + y_1 \cdot i - y_2 \cdot i = \underline{\underline{x_1 - x_2 + (y_1 - y_2) \cdot i}}. \quad (5.9)$$

Multiplikation: Es gilt

$$\begin{aligned} \underline{\underline{z_1 \cdot z_2}} &= (x_1 + y_1 \cdot i) \cdot (x_2 + y_2 \cdot i) = x_1 \cdot x_2 + y_1 \cdot i \cdot x_2 + x_1 \cdot y_2 \cdot i + y_1 \cdot i \cdot y_2 \cdot i \\ &= x_1 \cdot x_2 + x_2 \cdot y_1 \cdot i + x_1 \cdot y_2 \cdot i + y_1 \cdot y_2 \cdot i^2 \\ &= x_1 \cdot x_2 + x_1 \cdot y_2 \cdot i + x_2 \cdot y_1 \cdot i + y_1 \cdot y_2 \cdot (-1) \\ &= \underline{\underline{x_1 \cdot x_2 - y_1 \cdot y_2 + (x_1 \cdot y_2 + x_2 \cdot y_1) \cdot i}}. \end{aligned} \quad (5.10)$$

Division: Für $z_2 \neq 0$ gilt

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{x_1 + y_1 \cdot i}{x_2 + y_2 \cdot i} = \frac{(x_1 + y_1 \cdot i) \cdot (x_2 - y_2 \cdot i)}{(x_2 + y_2 \cdot i) \cdot (x_2 - y_2 \cdot i)} = \frac{x_1 \cdot x_2 + y_1 \cdot y_2 + (-x_1 \cdot y_2 + x_2 \cdot y_1) \cdot i}{x_2 \cdot x_2 + y_2 \cdot y_2 + (x_2 \cdot y_2 - x_2 \cdot y_2) \cdot i} \\ &= \frac{x_1 \cdot x_2 + y_1 \cdot y_2 + (x_2 \cdot y_1 - x_1 \cdot y_2) \cdot i}{x_2^2 + y_2^2 + 0} = \underline{\underline{\frac{x_1 \cdot x_2 + y_1 \cdot y_2}{x_2^2 + y_2^2} + \frac{x_2 \cdot y_1 - x_1 \cdot y_2}{x_2^2 + y_2^2} \cdot i}}. \end{aligned} \quad (5.11)$$

Die Ergebnisse lassen sich also auf einfache Weise wieder in *arithmetischer Form* schreiben. Wir betrachten den folgenden Satz.

Satz 5.2 *Betrag & Konjugation*

Sei $z \in \mathbb{C}$, dann gilt

$$\boxed{z \cdot z^* = |z|^2}. \quad (5.12)$$

Beweis: Es gibt $x, y \in \mathbb{R}$, so dass gilt

$$z = x + y \cdot i \quad \text{und} \quad z^* = x - y \cdot i. \quad (5.13)$$

Daraus erhalten wir

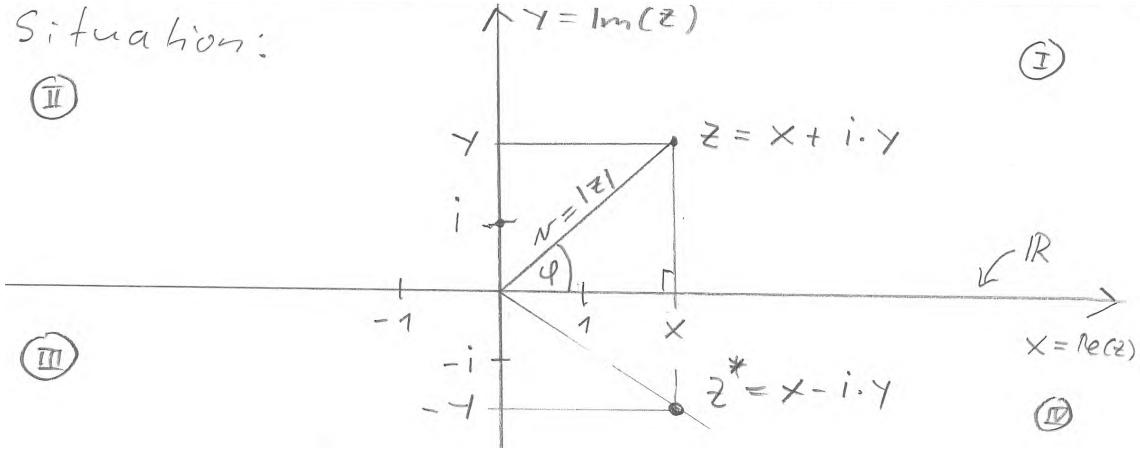
$$\underline{\underline{z \cdot z^*}} = (x + y \cdot i) \cdot (x - y \cdot i) = x^2 + y \cdot i \cdot x - x \cdot y \cdot i - y^2 \cdot i^2 = x^2 + y^2 = \underline{\underline{|z|^2}}. \quad (5.14)$$

Damit haben wir den Satz bewiesen. \square

5.2 Gauss-Ebene

5.2.1 Elementare Entsprechungen

Wir betrachten das folgende Diagramm.



Bemerkungen:

- i) Die Re-Achse entspricht den *reellen Zahlen*.
- ii) Die GAUSS-Ebene entspricht den *komplexen Zahlen* gemäss

$$x + y \cdot i \longleftrightarrow \begin{bmatrix} x \\ y \end{bmatrix}. \quad (5.15)$$

- iii) Es gibt die folgenden weiteren Entsprechungen zwischen \mathbb{C} und \mathbb{R}^2 :

Addition in \mathbb{C}	Addition in \mathbb{R}^2
Betrag in \mathbb{C}	Betrag in \mathbb{R}^2
Konjugation in \mathbb{C}	Spiegelung an Re-Achse in \mathbb{R}^2

5.2.2 Trigonometrische Form

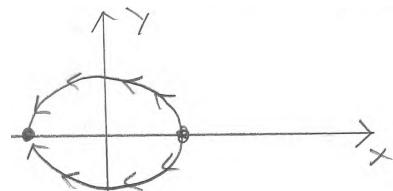
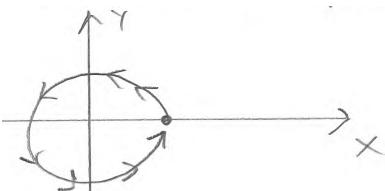
Die *Punkte* in der GAUSS-Ebene lassen sich sowohl durch *kartesische Koordinaten* als auch durch *Polarkoordinaten* beschreiben. Wir betrachten dazu folgende Definition.

Definition 5.3 Argument

Sei $z \in \mathbb{C} \setminus \{0\}$. Das *Argument* von $\arg(z)$ ist der *Winkel* in der GAUSS-Ebene zwischen der x -Achse und der *Verbindungsstrecke* vom Ursprung zu z nach folgenden Varianten.

(a) *Basler-Variante*: $\varphi \in [0, 2\pi[$.

(b) *Zürcher-Variante*: $\varphi \in]-\pi, \pi]$.



Bemerkungen:

- i)** Für $z = 0$ existiert in beiden Varianten kein eindeutiges *Argument*. Meistens wählt man $\arg(0) = 0$.
- ii)** Für fast alle Anwendungen kann sowohl die *Basler-Variante* als auch die *Zürcher-Variante* verwendet werden.
- iii)** Die Formeln zur Berechnung von $\arg(z)$ aus *Real-* und *Imaginärteil* sind bei der *Zürcher-Variante* etwas einfacher.
- iv)** Beispiel-Codes zur Berechnung des *Argumentes* mit gängiger Software.

MATLAB/Octave	<code>phi=angle(z)</code>
Mathematica/WolframAlpha	<code>phi=Arg[z]</code>
Python/Numpy	<code>import numpy as np;</code> <code>phi=np.angle(z)</code>
Python/Sympy	<code>import sympy as sp;</code> <code>phi=sp.arg(z)</code>

Wir betrachten den folgenden Satz.

Satz 5.3 Trigonometrische Form

Seien $x, y \in \mathbb{R}$, $r \in \mathbb{R}^+$ und $\varphi \in [0, 2\pi[$ oder $\varphi \in]-\pi, \pi]$ sowie

$$\text{cis}(\varphi) := \cos(\varphi) + i \cdot \sin(\varphi), \quad (5.17)$$

dann gibt es ein eindeutiges $z \in \mathbb{C} \setminus \{0\}$ mit

$$z = x + y \cdot i = r \cdot \text{cis}(\varphi). \quad (5.18)$$

Ferner gelten die folgenden Umrechnungsformeln.

(a) $x = r \cdot \cos(\varphi) \wedge y = r \cdot \sin(\varphi)$	(b) $r = z = \sqrt{x^2 + y^2} \wedge \varphi = \arg(z)$
---	--

Beweis: Die Aussage folgt sofort durch Anwenden der *Trigonometrie* in der *GAUSS-Ebene*. \square

Bemerkungen:

- i)** Jede *komplexe Zahl* $z \in \mathbb{C}$ lässt sich auf zwei Arten darstellen, nämlich

$$z = \underbrace{x + y \cdot i}_{\text{arithmetische Form}} = \underbrace{r \cdot \text{cis}(\varphi)}_{\text{trigonometrische Form}}. \quad (5.19)$$

- ii)** Man bezeichnet die Schreibweise $z = r \cdot \text{cis}(\varphi)$ nur dann als *trigonometrische Form*, wenn gilt $r = |z| \geq 0$. Der *Winkel* $\varphi \in \mathbb{R}$ darf jedoch beliebig gewählt werden, d.h. φ muss nicht unbedingt das *Argument* gemäss einer der beiden Varianten sein.

5.3 Quadratische Gleichungen

Wir betrachten eine *quadratische Gleichung* mit *reellen Koeffizienten* $a, b, c \in \mathbb{R}$ mit $a \neq 0$ der Form

$$a z^2 + b z + c = 0. \quad (5.20)$$

Dabei suchen wir alle *komplexen Lösungen* $z \in \mathbb{C}$ dieser *Gleichung*. Zunächst betrachten die *Diskriminante*

$$D := b^2 - 4 \cdot a \cdot c. \quad (5.21)$$

Je nach *Vorzeichen* der *Diskriminante* ergeben sich unterschiedliche *Lösungsmengen*. Wir betrachten diese Fälle getrennt.

Fall 1: Fall $D > 0$. In diesem Fall erhalten wir zwei *reelle Lösungen*

$$z_{1,2} = \frac{-b \pm \sqrt{D}}{2 \cdot a}. \quad (5.22)$$

Fall 2: Fall $D = 0$. In diesem Fall erhalten wir eine *reelle Lösung*

$$z = x_s = -\frac{b}{2 \cdot a}. \quad (5.23)$$

Fall 3: Fall $D < 0$. In diesem Fall erhalten wir zwei *komplexe Lösungen*

$$z_{1,2} = \frac{-b \pm \sqrt{D}}{2 \cdot a} = \frac{-b \pm i \cdot \sqrt{|D|}}{2 \cdot a} = -\frac{b}{2 \cdot a} \pm i \cdot \frac{\sqrt{|D|}}{2 \cdot a}. \quad (5.24)$$

Bemerkungen:

- i) Jede *quadratische Gleichung* mit *reellen Koeffizienten* hat demnach in \mathbb{C} **mindestens 1 Lösung** und **höchstens 2 Lösungen**.
- ii) Im Fall $D < 0$ gilt

$$\operatorname{Re}(z_1) = \operatorname{Re}(z_2) = x_s = -\frac{b}{2 \cdot a} \quad (5.25)$$

$$\operatorname{Im}(z_1) = -\operatorname{Im}(z_2). \quad (5.26)$$

Daraus folgt, dass die beiden *Lösungen* zueinander *komplex konjugiert* sind, d.h.

$$z_2 = z_1^*. \quad (5.27)$$

- iii) *Quadratische Gleichungen* mit *nichtreellen Koeffizienten* haben so gut wie keine Anwendungen in der Praxis.

5.4 Potenzen

5.4.1 Euler-Formel

Wir betrachten den folgenden Satz.

Satz 5.4 EULER-Formel

Sei $\varphi \in \mathbb{R}$, dann gilt

$$e^{i\varphi} = \text{cis}(\varphi) = \cos(\varphi) + i \cdot \sin(\varphi). \quad (5.28)$$

Beweis: Übung mit Hilfe von MACLAURIN-Entwicklungen. □

Wir betrachten den folgenden Satz.

Satz 5.5 Umkehrungen der EULER-Formel

Sei $\varphi \in \mathbb{R}$, dann gilt folgendes.

$$(a) \quad \sin(\varphi) = \frac{e^{i\varphi} - e^{-i\varphi}}{2 \cdot i} = -i \cdot \sinh(i \cdot \varphi)$$

$$(b) \quad \cos(\varphi) = \frac{e^{i\varphi} + e^{-i\varphi}}{2} = \cosh(i \cdot \varphi)$$

Beweis: Übung. □

Bemerkungen:

- i) Die EULER-Formel beschreibt die fundamentalen Zusammenhänge zwischen der *natürlichen Exponentialfunktion* und den *trigonometrischen bzw. hyperbolischen Funktionen*.
- ii) Aus der EULER-Formel lässt sich eine einfache algebraische Beziehung zwischen den fundamentalen Zahlen in $\{0, 1, e, \pi, i\}$ herleiten. Es gilt

$$e^{i\pi} + 1 = 0. \quad (5.29)$$

5.4.2 Exponentielle Form

Wir betrachten den folgenden Satz.

Satz 5.6 Exponentielle Form

Seien $x, y \in \mathbb{R}$, $r \in \mathbb{R}^+$ und $\varphi \in [0, 2\pi[$ oder $\varphi \in]-\pi, \pi]$ dann gibt es ein eindeutiges $z \in \mathbb{C} \setminus \{0\}$ mit

$$z = x + y \cdot i = r \cdot e^{i\varphi}. \quad (5.30)$$

Ferner gelten die folgenden Umrechnungsformeln.

$$(a) \quad x = r \cdot \cos(\varphi) \wedge y = r \cdot \sin(\varphi)$$

$$(b) \quad r = |z| = \sqrt{x^2 + y^2} \wedge \varphi = \arg(z)$$

Beweis: Die Aussage folgt sofort aus der EULER-Formel. □

Bemerkungen:

- i) Jede komplexe Zahl $z \in \mathbb{C}$ lässt sich auf drei Arten darstellen, nämlich

$$z = \underbrace{x + y \cdot i}_{\text{arithmetische Form}} = \underbrace{r \cdot \text{cis}(\varphi)}_{\text{trigonometrische Form}} = \underbrace{r \cdot e^{i\varphi}}_{\text{exponentielle Form}}. \quad (5.31)$$

- ii) Man bezeichnet die Schreibweise $z = r \cdot e^{i\varphi}$ nur dann als *exponentielle Form*, wenn gilt $r = |z| \geq 0$. Der Winkel $\varphi \in \mathbb{R}$ darf jedoch beliebig gewählt werden, d.h. φ muss nicht unbedingt das Argument gemäß einer der beiden Varianten sein.

- iii) In *exponentieller Form* lässt sich die *Multiplikation* von zwei komplexen Zahlen auf einfache Weise in der GAUSS-Ebene *geometrisch* interpretieren. Es gilt

$$\underline{\underline{z_1 \cdot z_2}} = r_1 \cdot e^{i\varphi_1} \cdot r_2 \cdot e^{i\varphi_2} = r_1 \cdot r_2 \cdot e^{i\varphi_1 + i\varphi_2} = \underline{\underline{r_1 \cdot r_2 \cdot e^{i(\varphi_1 + \varphi_2)}}}. \quad (5.32)$$

- iv) In *exponentieller Form* lässt sich das *Potenzieren* einer komplexen Zahl mit einem *strikt positiven Exponenten* $p > 0$ auf einfache Weise in der GAUSS-Ebene *geometrisch* interpretieren. Es gilt

$$\underline{\underline{z^p}} = (r \cdot e^{i\varphi})^p = r^p \cdot (e^{i\varphi})^p = \underline{\underline{r^p \cdot e^{ip\varphi}}}. \quad (5.33)$$

5.4.3 Potenz-Gleichungen

Seien $w \in \mathbb{C}$ und $n \in \mathbb{N}^+$. Wir betrachten eine *Potenz-Gleichung* der Form

$$z^n = w. \quad (5.34)$$

Dabei suchen wir alle *komplexen Lösungen* $z \in \mathbb{C}$. Beispiele:

- Wir betrachten die *Potenz-Gleichung*

$$z^2 = 4 \Rightarrow z_{1,2} = \pm \sqrt{4} = \pm 2 \Rightarrow \underline{\underline{\mathbb{L} = \{-2, 2\}}} \quad (5.35)$$

- Wir betrachten die *Potenz-Gleichung*

$$z^2 = -4 \Rightarrow z_{1,2} = \pm i \cdot \sqrt{4} = \pm 2 \cdot i \Rightarrow \underline{\underline{\mathbb{L} = \{-2 \cdot i, 2 \cdot i\}}}. \quad (5.36)$$

- Wir betrachten die *Potenz-Gleichung*

$$z^3 = -27 = |-27| \cdot e^{i \cdot \arg(-27)} = 27 \cdot e^{i \cdot \pi} = 27 \cdot e^{i \cdot (\pi + (k-1) \cdot 2\pi)} \quad \text{für } k \in \mathbb{Z}. \quad (5.37)$$

Wir erhalten die folgenden drei *Lösungen*

$$k=1 : z_1 = \sqrt[3]{27} \cdot e^{i \cdot \frac{\pi+0 \cdot 2\pi}{3}} = 3 \cdot e^{i \cdot \frac{\pi}{3}} \quad (5.38)$$

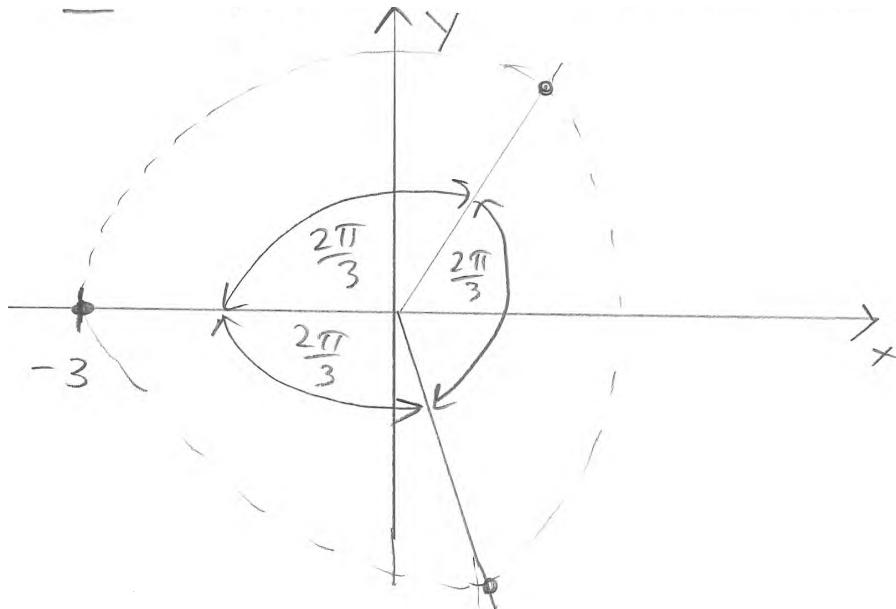
$$k=2 : z_2 = \sqrt[3]{27} \cdot e^{i \cdot \frac{\pi+1 \cdot 2\pi}{3}} = 3 \cdot e^{i \cdot \pi} = -3 \quad (5.39)$$

$$k=3 : z_3 = \sqrt[3]{27} \cdot e^{i \cdot \frac{\pi+2 \cdot 2\pi}{3}} = 3 \cdot e^{i \cdot \frac{5\pi}{3}} = 3 \cdot e^{-i \cdot \frac{\pi}{3}} \quad (5.40)$$

Die *Lösungsmenge* ist daher

$$\underline{\underline{L = \{3 \cdot e^{i \cdot \frac{\pi}{3}}, -3, 3 \cdot e^{-i \cdot \frac{\pi}{3}}\}}} \quad (5.41)$$

Die *Lösungen* bilden ein *gleichseitiges Dreieck* in der GAUSS-Ebene.



Wir betrachten den folgenden Satz.

Satz 5.7 *Lösungen von Potenz-Gleichungen*

Seien $w, z \in \mathbb{C}$, $n \in \mathbb{N}^+$, $r, \varphi \in \mathbb{R}$ mit $r > 0$ und z erfülle die *Potenz-Gleichung*

$$z^n = w = r \cdot e^{i \cdot \varphi}. \quad (5.42)$$

Dann gilt folgendes.

(a) Für $w = 0$ hat (5.42) genau eine *Lösung* für z , nämlich

$$z = 0. \quad (5.43)$$

(b) Für $w \neq 0$ hat (5.42) genau n *Lösungen* für z , nämlich

$$z_k = \sqrt[n]{r} \cdot e^{i \cdot \frac{\varphi + (k-1) \cdot 2\pi}{n}} \quad \text{für } k \in \{1, \dots, n\}. \quad (5.44)$$

Bemerkungen:

- i) Die Tatsache, dass (5.42) ausser für $w = 0$ genau n *Lösungen* hat, wird als *algebraische Vollständigkeit* oder *algebraische Abgeschlossenheit* von \mathbb{C} bezeichnet.
- ii) Für alle *Lösungen* von (5.42) gilt

$$|z_k| = \sqrt[n]{r} = \sqrt[n]{|w|}. \quad (5.45)$$

Die *Lösungen* liegen in der GAUSS-Ebene daher auf dem *Kreis* um den *Ursprung* mit *Radius* $\sqrt[n]{|w|}$.

- iii) Zwischen zwei benachbarten *Lösungen* von (5.42) liegt jeweils ein *Winkel* von

$$\Delta\varphi = \frac{2\pi}{n}. \quad (5.46)$$

- iv) Die *Lösungen* von (5.42) bilden ein *reguläres n -Eck* in der GAUSS-Ebene.

5.4.4 Wurzeln

Es gibt verschiedene Konventionen, was eine *Wurzel* mit *komplexem Radikanden* bedeuten soll.

1. *Alte Konvention:* Die *Wurzel* ist als *Menge* aller *Lösungen* der entsprechenden *Potenz-Gleichung* aufzufassen. Das *Element* mit dem *kleinsten Argument* gemäss *Basler-Variante* wird dabei als *Hauptwert* bezeichnet und die anderen *Elemente* als *Nebenwerte*. Beispiele:

- $\sqrt{4} = \left\{ \underbrace{4}_{\text{Hauptwert}}, \underbrace{-4}_{\text{Nebenwert}} \right\}$
- $\sqrt{-4} = \left\{ \underbrace{2 \cdot i}_{\text{Hauptwert}}, \underbrace{-2 \cdot i}_{\text{Nebenwert}} \right\}$
- $\sqrt{4 \cdot i} = \left\{ \underbrace{2 \cdot e^{i \cdot \frac{\pi}{4}}}_{\text{Hauptwert}}, \underbrace{2 \cdot e^{-i \cdot \frac{\pi}{4}}}_{\text{Nebenwert}} \right\}$
- $\sqrt[3]{-27} = \left\{ \underbrace{3 \cdot e^{i \cdot \frac{\pi}{3}}}_{\text{Hauptwert}}, \underbrace{-3, 3 \cdot e^{-i \cdot \frac{\pi}{3}}}_{\text{Nebenwerte}} \right\}$

2. *Neue Konvention:* Es wird zwischen *reellen* und *komplexen Wurzeln* unterschieden. Beispiele:

- $\sqrt{4} = 2 \quad (\text{reelle Lösung})$
- $\sqrt{-4} \quad \text{existiert nicht}$
- $\sqrt{4 \cdot i} \quad \text{existiert nicht}$
- $\sqrt[3]{-27} = -3 \quad (\text{reelle Lösung})$
- $\sqrt{4^C} = 2 \quad (\text{reelle Lösung, Hauptwert})$
- $\sqrt{-4^C} = 2 \cdot i \quad (\text{Hauptwert})$
- $\sqrt{4 \cdot i^C} = 2 \cdot e^{i \cdot \frac{\pi}{4}} \quad (\text{Hauptwert})$
- $\sqrt[3]{-27^C} = 3 \cdot e^{i \cdot \frac{\pi}{3}} \quad (\text{Hauptwert})$

3. *Allgemeine Praxis:* Es werden nur *reelle Radikanden* zugelassen und *reelle Lösungen* vor *Hauptwerten* bevorzugt. Ansonsten werden keine *Wurzeln* verwendet. Beispiele:

- $\sqrt{4} = 2 \quad (\text{reelle Lösung, Hauptwert})$
- $\sqrt{-4} = 2 \cdot i \quad (\text{Hauptwert})$
- $\sqrt[3]{-27} = -3 \quad (\text{reelle Lösung})$

Generell ist im Umgang mit *Wurzeln* Vorsicht geboten, wie die folgende Rechnung zeigt. Man könnte vermeintlich zu Schluss kommen

$$1 = \sqrt{1} = \sqrt{(-1) \cdot (-1)} = \sqrt{-1} \cdot \sqrt{-1} = i \cdot i = i^2 = -1. \quad (5.47)$$

Das dritte *Gleichheitszeichen* in der obigen Rechnung ist falsch. Allgemein gilt für alle *Wurzel-Konventionen*

$$\sqrt{x \cdot y} = \sqrt{x} \cdot \sqrt{y} \quad \text{für } x, y \in \mathbb{R}^+ \quad (5.48)$$

$$\sqrt{w \cdot z} = \pm \sqrt{w} \cdot \sqrt{z} \quad \text{für } w, z \in \mathbb{C}, \quad (5.49)$$

wobei das *Vorzeichen* von Fall zu Fall entschieden werden muss.

Kapitel 6

Matrizen & Lineare Abbildungen

6.1 Matrix-Algebra

6.1.1 Definition

Wir machen folgende Definition.

Definition 6.1 Matrix

Seien $m, n \in \mathbb{N}^+$. Eine *reelle $m \times n$ -Matrix* A ist eine Zahlentabelle mit m Zeilen und n Spalten der Form

$$A = \begin{bmatrix} A^1_1 & A^1_2 & \dots & A^1_n \\ A^2_1 & A^2_2 & \dots & A^2_n \\ \vdots & \vdots & \vdots & \vdots \\ A^m_1 & A^m_2 & \dots & A^m_n \end{bmatrix}, \quad (6.1)$$

wobei alle $A^i_j \in \mathbb{R}$.

Bemerkungen:

- i) Die Zahlen m und n heissen *Dimensionen* der Matrix A .
- ii) Die *reellen Zahlen* A^i_j heissen *Komponenten* der Matrix A . Eine $m \times n$ -Matrix besteht offensichtlich aus $m \cdot n$ Komponenten.
- iii) Wir werden später sehen, dass es sinnvoll ist, den *Zeilen-Index* i oben und den *Spalten-Index* j unten zu schreiben.
- iv) Für die *Menge* aller *reellen $m \times n$ -Matrizen* gibt es verschiedene Bezeichnungen. Wir verwenden

$$\mathbb{M}(m, n, \mathbb{R}) = \mathbb{R}^{m \times n}. \quad (6.2)$$

- v) Wir betrachten die *Matrix*

$$M = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}. \quad (6.3)$$

Beispiel-Codes zur Definition dieser *Matrix* mit gängiger Software.

MATLAB/Octave	<code>M=[1,2,3;4,5,6]</code>
Mathematica/WolframAlpha	<code>M={{1,2,3},{4,5,6}}</code>
Python/Numpy	<code>import numpy as np; M=np.array([[1,2,3],[4,5,6]])</code>
Python/Sympy	<code>import sympy as sp; M=sp.Matrix([[1,2,3],[4,5,6]])</code>

Beispiele:

- Eine 2×3 -Matrix: $A = \begin{bmatrix} 2 & -1 & 3 \\ 7 & 5 & -4 \end{bmatrix}$
- Eine 2×2 -Matrix: $B = \begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix}$
- Eine 1×3 -Matrix: $C = [2 \ -1 \ 3]$
- Eine 2×1 -Matrix: $D = \begin{bmatrix} 2 \\ 7 \end{bmatrix}$

6.1.2 Operationen

6.1.2.1 Addition & Subtraktion

Zwei reelle Matrizen mit gleichen Dimensionen lassen sich addieren und subtrahieren.

Definition 6.2 Addition & Subtraktion

Seien $m, n \in \mathbb{N}^+$ und $A, B \in \mathbb{M}(m, n, \mathbb{R})$, dann ist

$$A + B := \begin{bmatrix} A^1_1 + B^1_1 & A^1_2 + B^1_2 & \dots & A^1_n + B^1_n \\ A^2_1 + B^2_1 & A^2_2 + B^2_2 & \dots & A^2_n + B^2_n \\ \vdots & \vdots & \vdots & \vdots \\ A^m_1 + B^m_1 & A^m_2 + B^m_2 & \dots & A^m_n + B^m_n \end{bmatrix}. \quad (6.4)$$

und

$$A - B := \begin{bmatrix} A^1_1 - B^1_1 & A^1_2 - B^1_2 & \dots & A^1_n - B^1_n \\ A^2_1 - B^2_1 & A^2_2 - B^2_2 & \dots & A^2_n - B^2_n \\ \vdots & \vdots & \vdots & \vdots \\ A^m_1 - B^m_1 & A^m_2 - B^m_2 & \dots & A^m_n - B^m_n \end{bmatrix}. \quad (6.5)$$

Bemerkungen:

- i) Für alle $A, B \in \mathbb{M}(m, n, \mathbb{R})$ gilt $A \pm B \in \mathbb{M}(m, n, \mathbb{R})$.

- ii) Addiert oder subtrahiert man zwei reelle Matrizen, dann addiert bzw. subtrahiert man ihre Komponenten.
- iii) Die Addition bzw. Subtraktion von zwei reellen Matrizen unterschiedlicher Dimensionen ist nicht definiert.
- iv) Beispiel-Codes zur Addition und Subtraktion von Matrizen mit gängiger Software.

MATLAB/Octave	<code>M=A+B; M=A-B;</code>
Mathematica/WolframAlpha	<code>M=A+B; M=A-B;</code>
Python/Numpy	<code>M=A+B; M=A-B;</code>
Python/Sympy	<code>M=A+B; M=A-B;</code>

Beispiele:

- $\begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix} + \begin{bmatrix} 1 & -2 \\ 6 & -5 \end{bmatrix} = \begin{bmatrix} 3 & -3 \\ 13 & 0 \end{bmatrix}$
- $\begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix} - \begin{bmatrix} 1 & -2 \\ 6 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 10 \end{bmatrix}$
- $[0 \ 1] + [-1 \ -1] = [-1 \ 0]$

6.1.2.2 Multiplikation mit einem Skalar

Jede reelle Matrix lässt sich mit einer reellen Zahl multiplizieren.

Definition 6.3 Multiplikation mit einem Skalar

Seien $m, n \in \mathbb{N}^+$, $a \in \mathbb{R}$ und $A \in \mathbb{M}(m, n, \mathbb{R})$, dann ist

$$a \cdot A := \begin{bmatrix} a \cdot A^1_1 & a \cdot A^1_2 & \dots & a \cdot A^1_n \\ a \cdot A^2_1 & a \cdot A^2_2 & \dots & a \cdot A^2_n \\ \vdots & \vdots & \ddots & \vdots \\ a \cdot A^m_1 & a \cdot A^m_2 & \dots & a \cdot A^m_n \end{bmatrix}. \quad (6.6)$$

Bemerkungen:

- i) Für alle $a \in \mathbb{R}$ und $A \in \mathbb{M}(m, n, \mathbb{R})$ gilt $a \cdot A \in \mathbb{M}(m, n, \mathbb{R})$.
- ii) Multipliziert man eine reelle Matrix mit einem Skalar, dann multipliziert man ihre Komponenten mit dem Skalar.
- iii) Es soll keine Rolle spielen, ob der Skalar links oder rechts der reellen Matrix geschrieben wird. Man definiert

$$A \cdot a := a \cdot A. \quad (6.7)$$

- iv) Aus der Multiplikation mit einem Skalar ergibt sich auf natürliche Weise eine Division. Es soll gelten

$$\frac{A}{a} := \frac{1}{a} \cdot A. \quad (6.8)$$

- v) Beispiel-Codes zur *Multiplikation* und *Division* von *Matrizen* mit *Skalaren* mit gängiger Software.

MATLAB/Octave	<code>M=2*A; M=B/3;</code>
Mathematica/WolframAlpha	<code>M=2*A; M=B/3;</code>
Python/Numpy	<code>M=2*A; M=B/3;</code>
Python/Sympy	<code>M=2*A; M=B/3;</code>

Beispiele:

- $2 \cdot \begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ 14 & 10 \end{bmatrix}$
- $\frac{1}{3} \cdot \begin{bmatrix} 9 & -12 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 3 & -4 \\ 1 & 0 \end{bmatrix}$
- $(-1) \cdot \begin{bmatrix} 3 & -3 \end{bmatrix} = \begin{bmatrix} -3 & 3 \end{bmatrix}$

6.1.2.3 Transposition

Jede *reelle Matrix* lässt sich *transponieren*.

Definition 6.4 Transposition

Seien $m, n \in \mathbb{N}^+$ und $A \in \mathbb{M}(m, n, \mathbb{R})$, dann ist

$$A^T := \boxed{\begin{bmatrix} A_{11}^1 & A_{12}^2 & \dots & A_{1n}^m \\ A_{21}^1 & A_{22}^2 & \dots & A_{2n}^m \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1}^1 & A_{n2}^2 & \dots & A_{nn}^m \end{bmatrix}}. \quad (6.9)$$

Bemerkungen:

- i) Beim *Transponieren* einer *reellen Matrix* werden die Spalten mit den Zeilen vertauscht.
- ii) Ist $A \in \mathbb{M}(m, n, \mathbb{R})$, dann gilt $A^T \in \mathbb{M}(n, m, \mathbb{R})$.
- iii) Die *Transposition* ist eine *Involution*, d.h. für jede *Matrix* A gilt

$$(A^T)^T = A. \quad (6.10)$$

- iv) Beispiel-Codes zur Berechnung von *Matrix-Transpositionen* mit gängiger Software.

MATLAB/Octave	<code>M=A'</code>
Mathematica/WolframAlpha	<code>M=Transpose[A]</code>
Python/Numpy	<code>M=A.T</code>
Python/Sympy	<code>M=A.T</code>

Beispiele:

$$\bullet \begin{bmatrix} 2 & -1 & 3 \\ 7 & 5 & -4 \end{bmatrix}^T = \begin{bmatrix} 2 & 7 \\ -1 & 5 \\ 3 & -4 \end{bmatrix}$$

$$\bullet \begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix}^T = \begin{bmatrix} 2 & 7 \\ -1 & 5 \end{bmatrix}$$

$$\bullet [2 \ -1]^T = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

6.1.2.4 Matrix-Multiplikation

Zwei reelle Matrizen mit passenden Dimensionen lassen sich multiplizieren.

Definition 6.5 Matrix-Produkt

Seien $l, m, n \in \mathbb{N}^+$, $A \in \mathbb{M}(l, m, \mathbb{R})$ und $B \in \mathbb{M}(m, n, \mathbb{R})$, dann ist das Matrix-Produkt

$$C = A \cdot B \quad \text{definiert durch} \quad C^i_j := \sum_{s=1}^m A^i_s \cdot B^s_j. \quad (6.11)$$

Bemerkungen:

- i) Das Matrix-Produkt $A \cdot B$ ist also genau dann definiert, wenn A so viele Spalten wie B Zeilen hat.
- ii) Für $A \in \mathbb{M}(l, m, \mathbb{R})$ und $B \in \mathbb{M}(m, n, \mathbb{R})$ gilt $A \cdot B \in \mathbb{M}(l, n, \mathbb{R})$.
- iii) Die Komponenten des Matrix-Produkts $A \cdot B$ sind gerade die Kontraktionen ("Skalarprodukte") der Zeilen von A mit den Spalten von B .
- iv) Die Berechnung eines Matrix-Produkts ist im allgemeinen recht aufwändig.
- v) Beispiel-Codes zur Berechnung von Matrix-Produkten mit gängiger Software.

MATLAB/Octave	<code>M=A*B</code>
Mathematica/WolframAlpha	<code>M=A.B</code>
Python/Numpy	<code>M=A@B</code>
Python/Sympy	<code>M=A*B</code>

Beispiele:

$$\bullet \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$$

$$\bullet \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

- $\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 + 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} 11 \end{bmatrix}$
- $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 \cdot 3 & 1 \cdot 4 \\ 2 \cdot 3 & 2 \cdot 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$

6.1.2.5 Rechenregeln

Für *reelle Matrizen* gelten einige Rechenregeln.

Satz 6.1 Rechenregeln für Matrizen

Es seien A, B, C *reelle Matrizen* und $a, b \in \mathbb{R}$. Sofern alle Operationen gemäss den *Dimensionen* definiert sind, gelten die folgenden Rechenregeln.

- | | |
|---|---|
| (a) $A + B = B + A$ | (g) $A \cdot (B + C) = A \cdot B + A \cdot C$ |
| (b) $(A + B) + C = A + (B + C)$ | (h) $(A + B) \cdot C = A \cdot C + B \cdot C$ |
| (c) $a \cdot (A + B) = a \cdot A + a \cdot B$ | (i) $(A^T)^T = A$ |
| (d) $(a + b) \cdot A = a \cdot A + b \cdot A$ | (j) $(A + B)^T = A^T + B^T$ |
| (e) $(a \cdot A) \cdot B = a \cdot (A \cdot B) = A \cdot (a \cdot B)$ | (k) $(a \cdot A)^T = a \cdot A^T$ |
| (f) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ | (l) $(A \cdot B)^T = B^T \cdot A^T$ |

Bemerkungen:

- i) Die *Matrix-Addition* ist also *assoziativ* und *kommutativ* und es gilt das *Distributivgesetz* sowohl bei der *Multiplikation* mit einem *Skalar* als auch mit einer *reellen Matrix*.
- ii) Die *Matrix-Multiplikation* ist im allgemeinen nicht *kommutativ*. Es gibt *Matrizen* A und B , für welche gilt $A \cdot B = B \cdot A$ aber auch solche für die wir $A \cdot B \neq B \cdot A$ finden.
- iii) Das *Transponierte* eines *Matrix-Produkts* ist das umgekehrte *Matrix-Produkt* der *Transponierten Faktoren*.

6.1.3 Spezielle Matrizen

6.1.3.1 Quadratische Matrix

Eine ganz spezielle Rolle spielen *reelle Matrizen*, die gleich viele Zeilen wie Spalten haben.

Definition 6.6 Quadratische Matrix

Sei $n \in \mathbb{N}^+$. Eine *reelle Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ heisst *quadratische Matrix*.

Bemerkungen:

- i) *Quadratische Matrizen* haben genau n^2 Komponenten.
- ii) Das *Produkt* von zwei *quadratischen Matrizen* ist wieder eine *quadratische Matrix*, d.h. $A, B \in \mathbb{M}(n, n, \mathbb{R}) \Rightarrow A \cdot B \in \mathbb{M}(n, n, \mathbb{R})$.
- iii) *Quadratische Matrizen* können mit sich selbst *multipliziert* werden. So lassen sich *Potenzen* bilden. Für $n, p \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$ ist

$$A^p := \underbrace{A \cdot \dots \cdot A}_{p \text{ Faktoren}}. \quad (6.12)$$

Die *Potenz* ist dann wieder eine *quadratische Matrix*, d.h. $A^p \in \mathbb{M}(n, n, \mathbb{R})$.

Beispiele:

- $\begin{bmatrix} 2 \end{bmatrix}$
- $\begin{bmatrix} 2 & -1 \\ 7 & 5 \end{bmatrix}$
- $\begin{bmatrix} 3 & -6 & 7 \\ 1 & 0 & -2 \\ 1 & 8 & 9 \end{bmatrix}$

Von zwei *quadratischen Matrizen* gleicher *Dimension* können beide *Produkt*, d.h. sowohl $A \cdot B$ als auch $B \cdot A$ gebildet werden. Die *Differenz* ist in vielen Anwendungen von Interesse.

Definition 6.7 Kommutator

Seien $n \in \mathbb{N}^+$ und $A, B \in \mathbb{M}(n, n, \mathbb{R})$. Der *Kommutator* von A und B ist die *Matrix*

$$[A, B] := A \cdot B - B \cdot A. \quad (6.13)$$

Bemerkungen:

- i) Der *Kommutator* von zwei *quadratischen Matrizen* ist wieder eine *quadratische Matrix*, d.h. $A, B \in \mathbb{M}(n, n, \mathbb{R}) \Rightarrow [A, B] \in \mathbb{M}(n, n, \mathbb{R})$.
- ii) Der *Kommutator* verschwindet genau dann, wenn die *Matrizen* kommutieren, d.h. wenn gilt $A \cdot B = B \cdot A$.
- iii) Der *Kommutator* ist als *Operation schiefsymmetrisch*, d.h. es gilt

$$[A, B] = -[B, A]. \quad (6.14)$$

6.1.3.2 Symmetrische & schiefsymmetrische Matrix

Spezielle Bedeutung haben *quadratische Matrizen*, die bei *Transposition* in sich selbst oder in ihr Negatives übergehen.

Definition 6.8 *Symmetrische & schiefsymmetrische Matrix*

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$.

- (a) A ist *symmetrisch*, genau falls $A^T = A$.
- (b) A ist *schiefsymmetrisch*, genau falls $A^T = -A$.

Bemerkungen:

- i) Die *Komponenten* auf der *Hauptdiagonalen* einer *schiefsymmetrischen Matrix* müssen zwingend verschwinden.
- ii) Eine *symmetrische Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ hat die Anzahl unabhängiger *Komponenten* von

$$n_+ = \frac{n \cdot (n + 1)}{2}. \quad (6.15)$$

- iii) Eine *schiefsymmetrische Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ hat die Anzahl unabhängiger *Komponenten* von

$$n_- = \frac{n \cdot (n - 1)}{2}. \quad (6.16)$$

Beispiele:

- $\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$ ist *symmetrisch*
- $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$ ist *schiefsymmetrisch*
- $\begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -3 \\ -2 & 3 & 0 \end{bmatrix}$ ist *schiefsymmetrisch*

6.1.3.3 Nullmatrix

Eine reelle Matrix, deren Komponenten alle verschwinden, heisst *Nullmatrix*.

Definition 6.9 Nullmatrix

Seien $m, n \in \mathbb{N}^+$. Die Matrix $0 \in \mathbb{M}(m, n, \mathbb{R})$ mit

$$0 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (6.17)$$

heisst *Nullmatrix*.

Bemerkungen:

- i) Alle Nullmatrizen werden unabhängig von ihren Dimensionen identifiziert und mit 0 bezeichnet. Es gilt also

$$0 = [0] = [0 \ 0] = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \dots \quad (6.18)$$

- ii) Die Nullmatrix hat die gleichen algebraischen Eigenschaften wie die Zahl Null. Für jede Matrix A gilt

$$A + 0 = A \quad \text{und} \quad 0 \cdot A = 0. \quad (6.19)$$

- iii) Die quadratischen Nullmatrizen sind die einzigen Matrizen die sowohl *symmetrisch* als auch *schiefsymmetrisch* sind.

- iv) Beispiel-Codes zum Erzeugen von Nullmatrizen mit gängiger Software.

MATLAB/Octave	<code>M=zeros(3)</code> <code>M=zeros(2,3)</code>
Mathematica/WolframAlpha	<code>M=ConstantArray[0,{3,3}]</code> <code>M=ConstantArray[0,{2,3}]</code>
Python/Numpy	<code>import numpy as np;</code> <code>M=np.zeros((3,3))</code> <code>M=np.zeros((2,3))</code>
Python/Sympy	<code>import sympy as sp;</code> <code>M=sp.zeros(3)</code> <code>M=sp.zeros(2,3)</code>

6.1.3.4 Einheitsmatrix

Eine ganz spezielle Rolle spielen *quadratische Matrizen*, die auf der *Hauptdiagonalen* nur Einsen und sonst überall nur Nullen haben.

Definition 6.10 Einheitsmatrix

Sei $n \in \mathbb{N}^+$. Die Matrix $\mathbb{1} \in \mathbb{M}(n, n, \mathbb{R})$ mit

$$\mathbb{1} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \quad (6.20)$$

heisst *Einheitsmatrix*.

Bemerkungen:

- i) Alle *Einheitsmatrizen* werden unabhängig von ihrer *Dimension* identifiziert und mit $\mathbb{1}$ bezeichnet. Es gilt also

$$\mathbb{1} = \begin{bmatrix} 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \dots \quad (6.21)$$

- ii) Die *Einheitsmatrix* hat die gleichen algebraischen Eigenschaften wie die Zahl Eins. Für jede *Matrix* A gilt

$$\mathbb{1} \cdot A = A \cdot \mathbb{1} = A. \quad (6.22)$$

- iii) Ist A eine *quadratische Matrix*, dann folgt der *Kommutator*

$$[A, \mathbb{1}] = A \cdot \mathbb{1} - \mathbb{1} \cdot A = A - A = 0. \quad (6.23)$$

Das heisst, die *Einheitsmatrix* *kommutiert* mit jeder *quadratischen Matrix*.

- iv) Die *Einheitsmatrix* ist offensichtlich *symmetrisch*.
- v) Beispiel-Codes zum Erzeugen von *Einheitsmatrizen* mit gängiger Software.

MATLAB/Octave	<code>M=eye(3)</code> <code>M=eye(2,3)</code>
Mathematica/WolframAlpha	<code>M=IdentityMatrix[3]</code>
Python/Numpy	<code>import numpy as np;</code> <code>M=np.eye(3)</code> <code>M=np.eye(2,3)</code>
Python/Sympy	<code>import sympy as sp;</code> <code>M=sp.eye(3)</code> <code>M=sp.eye(2,3)</code>

6.1.3.5 Invertierbare Matrix

Einige *quadratische Matrizen* haben so etwas wie einen *Kehrwert*.

Definition 6.11 Invertierbare Matrix

Sei $n \in \mathbb{N}^+$. Eine *quadratische Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ heisst *invertierbar*, falls es eine *quadratische Matrix* $A^{-1} \in \mathbb{M}(n, n, \mathbb{R})$ gibt, so dass

$$A^{-1} \cdot A = \mathbb{1}. \quad (6.24)$$

Bemerkungen:

- i) *Invertierbare Matrizen* werden auch *regulär* genannt, während *nicht invertierbare Matrizen* als *singulär* bezeichnet werden.
- ii) Die *quadratische Matrix* A^{-1} , falls es die denn gibt, wird *Inverse Matrix* von A genannt.
- iii) Obwohl die *Matrix-Multiplikation* im allgemeinen nicht *kommutativ* ist, lässt sich die Reihenfolge in (6.24) immer vertauschen. Ist A *invertierbar*, dann gilt

$$A^{-1} \cdot A = A \cdot A^{-1} = \mathbb{1}. \quad (6.25)$$

Daraus erhält man den *Kommutator*

$$[A, A^{-1}] = A \cdot A^{-1} - A^{-1} \cdot A = \mathbb{1} - \mathbb{1} = 0. \quad (6.26)$$

- iv) Die *Einheitsmatrix* ist offensichtlich *invertierbar* und ihre *eigene Inverse*, es gilt also

$$\mathbb{1}^{-1} = \mathbb{1}. \quad (6.27)$$

- v) Die *Nullmatrix* ist offensichtlich *singulär*, d.h. *nicht invertierbar*.
- vi) Die *Inversion* einer *reellen Matrix* ist offensichtlich eine *Involution*. Für jede *invertierbare Matrix* gilt

$$(A^{-1})^{-1} = A. \quad (6.28)$$

- vii) Beispiel-Codes zur Berechnung von *inversen Matrizen* mit gängiger Software.

MATLAB/Octave	<code>M=inv(A)</code>
Mathematica/WolframAlpha	<code>M=Inverse[A]</code>
Python/Numpy	<code>import numpy as np;</code> <code>M=np.linalg.inv(A)</code>
Python/Sympy	<code>M=A.inv()</code>

Beispiele:

- $\begin{bmatrix} 2 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} \end{bmatrix}$
- $\begin{bmatrix} 2 & 1 \\ 5 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} 3 & -1 \\ -5 & 2 \end{bmatrix}$

$$\bullet \begin{bmatrix} 2 & 5 \\ -3 & -8 \end{bmatrix}^{-1} = \begin{bmatrix} 8 & 5 \\ -3 & -2 \end{bmatrix}$$

Für *invertierbare Matrizen* gelten einige Rechenregeln.

Satz 6.2 Elementare Rechenregeln der Inversion

Es seien A, B *invertierbare Matrizen* und $a \in \mathbb{R} \setminus \{0\}$. Dann gelten folgende Rechenregeln.

- | | |
|--|--------------------------------------|
| (a) $(a \cdot A)^{-1} = \frac{1}{a} \cdot A^{-1}$ | (c) $(A^T)^{-1} = (A^{-1})^T$ |
| (b) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$ | (d) $(A^{-1})^{-1} = A$ |

6.1.3.6 Diagonale Matrix

Mit *quadratischen Matrizen*, die nur auf der *Hauptdiagonalen* von Null verschiedene *Komponenten* haben, lässt sich besonders einfach rechnen.

Definition 6.12 Diagonal-Matrix

Sei $n \in \mathbb{N}^+$ und $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Eine *Matrix* $D \in \mathbb{M}(n, n, \mathbb{R})$ der Form

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}. \quad (6.29)$$

heisst *diagonal*.

Bemerkungen:

- i)** Die *reellen Zahlen* $\lambda_1, \dots, \lambda_n$ auf der *Hauptdiagonalen* heissen *Eigenwerte* der *Matrix* D .
- ii)** Die *Einheitsmatrix* sowie jede *quadratische Nullmatrix* sind offensichtlich *diagonal*.
- iii)** Jede *diagonale Matrix* ist offensichtlich *symmetrisch*.
- iv)** Zwei beliebige *diagonalen Matrizen kommutieren*, d.h. es gilt

$$[D, \tilde{D}] = 0. \quad (6.30)$$

Aber: *Diagonale Matrizen kommutieren nicht mit allen Matrizen!*

- v)** Sind alle *Eigenwerte* einer *diagonalen Matrix* von Null verschieden, dann ist die *Matrix invertierbar*. Die *Inverse* ist ebenfalls *diagonal* und es gilt

$$D^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda_n} \end{bmatrix}. \quad (6.31)$$

vi) Beispiel-Codes zum Erzeugen von *diagonalen Matrizen* mit gängiger Software.

MATLAB/Octave	<code>M=diag([1,2,3])</code>
Mathematica/WolframAlpha	<code>M=DiagonalMatrix[{1,2,3}]</code>
Python/Numpy	<code>import numpy as np; M=np.diag([1,2,3])</code>
Python/Sympy	<code>import sympy as sp; M=sp.diag(1,2,3)</code>

Beispiele:

- $\begin{bmatrix} 3 \end{bmatrix}$
- $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$
- $\begin{bmatrix} -7 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

6.2 Lineare Abbildungen

6.2.1 Definition

Mit Hilfe von *Matrizen* können *geometrische Abbildungen* beschrieben werden. Wir machen dazu zwei Definitionen.

Definition 6.13 *Lineare Abbildung* - Version 1

Seien $m, n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, m, \mathbb{R})$. Eine *Abbildung* der Form

$$\boxed{\begin{aligned} a : \mathbb{R}^m &\rightarrow \mathbb{R}^n \\ \mathbf{x} &\mapsto a(\mathbf{x}) := A \cdot \mathbf{x} \end{aligned}} \quad (6.32)$$

heisst *lineare Abbildung*.

Definition 6.14 *Lineare Abbildung* - Version 2

Seien $m, n \in \mathbb{N}^+$. Eine *lineare Abbildung* ist eine *Abbildung* des Typs $a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ mit der Eigenschaft, dass für alle $\mathbf{v}, \mathbf{w} \in \mathbb{R}^m$ und $x, y \in \mathbb{R}$ gilt

$$\boxed{a(x \cdot \mathbf{v} + y \cdot \mathbf{w}) = x \cdot a(\mathbf{v}) + y \cdot a(\mathbf{w})}. \quad (6.33)$$

Bemerkungen:

- i) Die beiden Definitionen sind *äquivalent*. Es ist leicht einzusehen, dass die *linearen Abbildungen* gemäss Definition 6.13 die *Haupteigenschaft* (6.33) aus Definition 6.14 erfüllen. Es gilt

$$\begin{aligned} \underline{\underline{a(x \cdot \mathbf{v} + y \cdot \mathbf{w})}} &= A \cdot (x \cdot \mathbf{v} + y \cdot \mathbf{w}) = A \cdot x \cdot \mathbf{v} + A \cdot y \cdot \mathbf{w} = x \cdot A \cdot \mathbf{v} + y \cdot A \cdot \mathbf{w} \\ &= \underline{\underline{x \cdot a(\mathbf{v}) + y \cdot a(\mathbf{w})}}. \end{aligned} \quad (6.34)$$

- ii) Die *Matrix* A , welche gemäss Definition 6.13 eine *lineare Abbildung* a beschreibt, wird *Abbildungsmatrix* genannt.
- iii) Ist die *Abbildungsmatrix* A *quadratisch*, d.h. $m = n$, dann ist a eine *Selbstabbildung* des Typs $a : \mathbb{R}^n \rightarrow \mathbb{R}^n$.
- iv) Bekannte *geometrische Abbildungen* wie *Streckungen*, *Projektionen*, *Spiegelungen* und *Rotationen* sind *lineare Abbildungen*, welche jeweils durch eine *Abbildungsmatrix* ausgedrückt werden können.
- v) Für alle *linearen Abbildungen* gilt offensichtlich

$$\underline{\underline{a(0)}} = A \cdot 0 = \underline{\underline{0}}. \quad (6.35)$$

- vi) Für $n = m = 1$ gibt es eine historisch bedingte Begriffskollision zwischen einer *linearen Funktion* in der *Analysis* und einer *linearen Abbildung* in der *linearen Algebra*.

Analysis: $f(x) = m \cdot x + q$

Lineare Algebra: $a(x) = A \cdot x$. (6.36)

In der *Analysis* heisst eine *Funktion* genau dann *linear*, wenn ihr *Graph* eine *Gerade* ist. Somit ist im *Funktionsterm* nebst dem Term $m \cdot x$ auch noch eine Konstante $q \neq 0$ zulässig. In der *linearen Algebra* würde solch ein q jedoch die Gültigkeit der *Haupt Eigenschaft* (6.33) verhindern.

6.2.2 Eigenschaften

Die *Abbildungsmatrix* der *Verknüpfung* von zwei *linearen Abbildungen* lässt sich auf einfache Weise aus den *Abbildungsmatrizen* der beiden *verknüpften linearen Abbildungen* berechnen.

Satz 6.3 Verknüpfungssatz

Seien $m, n, l \in \mathbb{N}^+$ und $a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ sowie $b : \mathbb{R}^n \rightarrow \mathbb{R}^l$ zwei *lineare Abbildungen* mit *Abbildungsmatrizen* $A \in \mathbb{M}(n, m, \mathbb{R})$ bzw. $B \in \mathbb{M}(l, n, \mathbb{R})$. Dann ist die *Verknüpfung*

$$\begin{aligned} c : \mathbb{R}^m &\rightarrow \mathbb{R}^l \\ \mathbf{x} &\mapsto c(\mathbf{x}) := b(a(\mathbf{x})) \end{aligned} \tag{6.37}$$

ebenfalls eine *lineare Abbildung* mit *Abbildungsmatrix*

$$C = B \cdot A. \tag{6.38}$$

Beweis: Wegen der *Assoziativität* des *Matrix-Produkts* finden wir für alle $\mathbf{x} \in \mathbb{R}^m$

$$\underline{c(\mathbf{x})} = b(a(\mathbf{x})) = b(A \cdot \mathbf{x}) = B \cdot (A \cdot \mathbf{x}) = \underline{(B \cdot A) \cdot \mathbf{x}} =: C \cdot \mathbf{x}, \tag{6.39}$$

wobei gelten muss

$$\underline{\underline{C = B \cdot A.}} \tag{6.40}$$

Damit haben wir den Satz bewiesen. □

Ist eine *lineare Abbildung bijektiv*, dann ist auch die *Umkehrabbildung* wieder eine *lineare Abbildung*, deren *Abbildungsmatrix* gerade die *Inversen* der ursprünglichen *Abbildungsmatrix* ist.

Satz 6.4 Inversionssatz

Seien $m, n \in \mathbb{N}^+$ und $a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine *lineare Abbildung* mit *Abbildungsmatrix* $A \in \mathbb{M}(n, m, \mathbb{R})$. Dann gilt folgendes.

- (a) a bijektiv $\Rightarrow n = m$
- (b) a bijektiv $\Leftrightarrow A$ regulär
- (c) a bijektiv $\Rightarrow a^{-1}(\mathbf{y}) = A^{-1} \cdot \mathbf{y}$.

Bemerkungen:

- i) Die *Verknüpfung* von zwei *linearen Abbildungen* geschieht durch *Matrix-Multiplikation* der *Abbildungsmatrizen*. Dabei muss die Reihenfolge beachtet werden, d.h.

$$a(\mathbf{x}) := a_N \left(\dots a_2(a_1(\mathbf{x})) \right) \Rightarrow A = A_N \cdot \dots \cdot A_2 \cdot A_1. \tag{6.41}$$

- ii) Die *Umkehrung* einer *bijektiven linearen Abbildung* geschieht durch *Inversion* der *Abbildungsmatrix*.

6.2.3 Spalten-Vektor-Konstruktionsverfahren

Wir betrachten eine *lineare Abbildung* $a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ deren geometrische Wirkung auf *Vektoren* in \mathbb{R}^m wir kennen, z.B. eine *Streckung*, *Projektion*, *Spiegelung* oder *Drehung*. Wie lässt sich nun die zugehörige *Abbildungsmatrix* A berechnen? Um die überraschend einfache Antwort auf diese Frage zu finden, betrachten wir zunächst eine allgemeine 2×2 -Matrix sowie die *Standard-Einheitsvektoren* in \mathbb{R}^2 , d.h.

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{bzw.} \quad \hat{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \hat{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (6.42)$$

Es gilt

$$a(\hat{\mathbf{e}}_1) = A \cdot \hat{\mathbf{e}}_1 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \cdot 1 + b \cdot 0 \\ c \cdot 1 + d \cdot 0 \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix} \quad (6.43)$$

$$a(\hat{\mathbf{e}}_2) = A \cdot \hat{\mathbf{e}}_2 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \cdot 0 + b \cdot 1 \\ c \cdot 0 + d \cdot 1 \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}. \quad (6.44)$$

Die *Bilder* der *Standard-Einheitsvektoren* unter der *linearen Abbildung* a sind demnach gerade die *Spalten* der *Abbildungsmatrix* A . Dies ist eine allgemeingültige Tatsache, die für beliebige *lineare Abbildungen* gilt.

Satz 6.5 Spalten-Vektor-Satz

Seien $m, n \in \mathbb{N}^+$, $a : \mathbb{R}^m \rightarrow \mathbb{R}^n$ eine *lineare Abbildung* mit *Abbildungsmatrix* $A \in \mathbb{M}(n, m, \mathbb{R})$ und

$$\hat{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \hat{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad \hat{\mathbf{e}}_m = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}. \quad (6.45)$$

Dann gilt

$$A = [a(\hat{\mathbf{e}}_1) \quad a(\hat{\mathbf{e}}_2) \quad \dots \quad a(\hat{\mathbf{e}}_m)]. \quad (6.46)$$

Beweis: Die *Bilder* der *Einheitsvektoren* aus (6.45) unter der *linearen Abbildung* a sind

$$a(\hat{\mathbf{e}}_1) = A \cdot \hat{\mathbf{e}}_1 = \begin{bmatrix} A^1_1 & A^1_2 & \dots & A^1_m \\ A^2_1 & A^2_2 & \dots & A^2_m \\ A^3_1 & A^3_2 & \dots & A^3_m \\ \vdots & \vdots & \vdots & \vdots \\ A^n_1 & A^n_2 & \dots & A^n_m \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} A^1_1 \\ A^2_1 \\ A^3_1 \\ \vdots \\ A^n_1 \end{bmatrix} \quad (6.47)$$

$$a(\hat{\mathbf{e}}_2) = A \cdot \hat{\mathbf{e}}_2 = \begin{bmatrix} A^1_1 & A^1_2 & \dots & A^1_m \\ A^2_1 & A^2_2 & \dots & A^2_m \\ A^3_1 & A^3_2 & \dots & A^3_m \\ \vdots & \vdots & \vdots & \vdots \\ A^n_1 & A^n_2 & \dots & A^n_m \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} A^1_2 \\ A^2_2 \\ A^3_2 \\ \vdots \\ A^n_2 \end{bmatrix} \quad (6.48)$$

$$a(\hat{\mathbf{e}}_m) = A \cdot \hat{\mathbf{e}}_m = \begin{bmatrix} A^1_1 & A^1_2 & \dots & A^1_m \\ A^2_1 & A^2_2 & \dots & A^2_m \\ A^3_1 & A^3_2 & \dots & A^3_m \\ \vdots & \vdots & \vdots & \vdots \\ A^n_1 & A^n_2 & \dots & A^n_m \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} A^1_m \\ A^2_m \\ A^3_m \\ \vdots \\ A^n_m \end{bmatrix}. \quad (6.49)$$

Daraus erhalten wir

$$\underline{\underline{A}} = \begin{bmatrix} A^1_1 & A^1_2 & \dots & A^1_m \\ A^2_1 & A^2_2 & \dots & A^2_m \\ \vdots & \vdots & \vdots & \vdots \\ A^n_1 & A^n_2 & \dots & A^n_m \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} A^1_1 \\ A^2_1 \\ \vdots \\ A^n_1 \end{bmatrix} & \begin{bmatrix} A^1_2 \\ A^2_2 \\ \vdots \\ A^n_2 \end{bmatrix} & \dots & \begin{bmatrix} A^1_m \\ A^2_m \\ \vdots \\ A^n_m \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a(\hat{\mathbf{e}}_1) & a(\hat{\mathbf{e}}_2) & \dots & a(\hat{\mathbf{e}}_m) \end{bmatrix}. \quad (6.50)$$

Damit haben wir den Satz bewiesen. \square

Bemerkungen:

- i) Um die *Abbildungsma*trix einer *linearen Abbildung* zu bestimmen, braucht man also nur ihre *geometrische Wirkung* auf die *Standard-Einheitsvektoren* zu kennen.
- ii) Um den *Spalten-Vektor-Satz* auf eine *geometrische Abbildung* anwenden zu können, muss man jedoch schon wissen bzw. mit Hilfe der *Haupteigenschaft* (6.33) zuerst prüfen, dass die *Abbildung* auch wirklich *linear* ist.

6.2.4 Beispiele

Bei folgenden Beispielen von *linearen Abbildungen* des Typs $a : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, lässt sich die *Abbildungsma*trix sehr leicht mit Hilfe des *Spalten-Vektor-Konstruktionsverfahrens* finden.

- *Identität*: $\mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
- *Streckung* am *Ursprung* um den *Faktor* $\lambda \in \mathbb{R}$: $Z_\lambda = \lambda \cdot \mathbb{1} = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$
- *Punktspiegelung* am *Ursprung*: $P = -\mathbb{1} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$
- *Projektion* auf die *x-Achse*: $P_x := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
- *Projektion* auf die *y-Achse*: $P_y := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
- *Spiegelung* an der *x-Achse*: $S_x := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

- Spiegelung an der y -Achse: $S_y := \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
- Spiegelung an der Geraden $y = x$: $S_{xy} := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
- Drehung um den Ursprung um $\pi/2$: $R(\pi/2) = \mathbb{i} := \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$
- Drehung um den Ursprung um $-\pi/2$: $R(-\pi/2) = -\mathbb{i} := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$
- Drehung um den Ursprung um φ : $R(\varphi) = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix}$

Die Matrix \mathbb{i} , welche die Drehung um den Ursprung um $\pi/2$ beschreibt, heisst *symplektische Matrix*. Sie hat analoge algebraische Eigenschaften wie die *imaginäre Einheit* in den *komplexen Zahlen*. Insbesondere gilt

$$\mathbb{i}^2 = R^2(\pi/2) = R(\pi/2) \cdot R(\pi/2) = R(\pi) = P = -\mathbb{1}. \quad (6.51)$$

6.3 Orthogonale Matrizen

6.3.1 Definition

Es gibt *reguläre Matrizen*, deren *Inverser* gerade auch ihre *Transponierte* ist. Dies ist eine sehr interessante Eigenschaft, die sowohl *algebraische* als auch *geometrische* Konsequenzen hat.

Definition 6.15 Orthogonale Matrix

Sei $n \in \mathbb{N}^+$. Eine *reguläre Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ heisst *orthogonal*, falls

$$A^{-1} = A^T. \quad (6.52)$$

Ferner definieren wir die *Menge* aller *orthogonalen* $n \times n$ -*Matrizen*.

Definition 6.16 Orthogonale Gruppe

Sei $n \in \mathbb{N}^+$. Die *orthogonale Gruppe* in \mathbb{N}^D ist die *Menge*

$$\mathrm{O}(n) := \{A \in \mathbb{M}(n, n, \mathbb{R}) \mid A^{-1} = A^T\}. \quad (6.53)$$

Wir betrachten den folgenden Satz.

Satz 6.6 Orthogonale Gruppe

Sei $n \in \mathbb{N}^+$, dann bildet $\mathrm{O}(n)$ eine *algebraische Gruppe*, d.h. für alle $A, B, C \in \mathrm{O}(n)$ gilt folgendes.

(a) Endogenität:

$$A \cdot B \in \mathrm{O}(n)$$

(c) Neutrales Element:

$$\mathbb{1} \in \mathrm{O}(n)$$

(b) Assoziativität:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C)$$

(d) Inverse Elemente:

$$A^{-1} \in \mathrm{O}(n)$$

Bemerkungen:

i) Offensichtlich gilt $\mathbb{1} \in \mathrm{O}(n)$ für alle $n \in \mathbb{N}^+$, denn es gilt

$$\mathbb{1}^{-1} = \mathbb{1} = \mathbb{1}^T. \quad (6.54)$$

ii) Ist eine *Matrix orthogonal und symmetrisch*, dann folgt

$$\underline{\underline{A}}^{-1} = \underline{\underline{A}}^T = \underline{\underline{A}} \Rightarrow \underline{\underline{A}}^2 = \underline{\underline{A}} \cdot \underline{\underline{A}} = \underline{\underline{A}}^{-1} \cdot \underline{\underline{A}} = \underline{\underline{1}}. \quad (6.55)$$

Die *orthogonalen symmetrischen Matrizen* verhalten sich wie *Wurzeln der Einheitsmatrix*.

iii) Ist eine *Matrix orthogonal und schiefsymmetrisch*, dann folgt

$$\underline{\underline{A}}^{-1} = \underline{\underline{A}}^T = \underline{\underline{-A}} \Rightarrow \underline{\underline{A}}^2 = -\underline{\underline{A}}^{-1} \cdot \underline{\underline{A}} = \underline{\underline{-1}}. \quad (6.56)$$

Die *orthogonalen schiefsymmetrischen Matrizen* verhalten sich ähnlich wie die *imaginäre Einheit* $i \in \mathbb{C}$.

- iv)** Ganz besonders bequem lässt sich ein LGLS lösen, wenn die linke Seite eine *orthogonale Matrix* ist. Gilt $A \in O(n)$, dann folgt

$$A \cdot \mathbf{u} = \mathbf{b} \quad \Big| A^T. \quad (6.57)$$

$$\Leftrightarrow A^T \cdot A \cdot \mathbf{u} = A^{-1} \cdot A \cdot \mathbf{u} = A^T \cdot \mathbf{b} \quad (6.58)$$

$$\Leftrightarrow \mathbb{1} \cdot \mathbf{u} = A^T \cdot \mathbf{b}. \quad (6.59)$$

Daraus erhalten wir die *Lösung*

$$\underline{\mathbf{u} = A^T \cdot \mathbf{b}}. \quad (6.60)$$

Das LGLS ist demnach *eindeutig lösbar* und die *Lösung* lässt sich ganz einfach durch *Multiplikation* der rechten Seite von links mit der *Transponierten* von A berechnen.

Beispiele:

- $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
- $\frac{1}{\sqrt{13}} \cdot \begin{bmatrix} -2 & 3 \\ 3 & 2 \end{bmatrix}$
- $\frac{1}{5} \cdot \begin{bmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

6.3.2 Geometrische Eigenschaften

Die *Spalten-Vektoren* einer *orthogonalen Matrix* haben bemerkenswerte geometrische Eigenschaften.

Satz 6.7 *Orthonormalität der Spalten-Vektoren*

Seien $n \in \mathbb{N}^+$ und $A \in O(n)$ mit *Spaltenvektoren* $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{R}^n$, d.h.

$$A = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n]. \quad (6.61)$$

Dann gilt

$$\langle \mathbf{a}_i, \mathbf{a}_j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases} \quad (6.62)$$

Beweis: Weil A *orthogonal* ist, gilt

$$\underline{\underline{1}} = A^{-1} \cdot A = A^T \cdot A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix} \cdot [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_n] = \begin{bmatrix} \mathbf{a}_1^T \cdot \mathbf{a}_1 & \mathbf{a}_1^T \cdot \mathbf{a}_2 & \dots & \mathbf{a}_1^T \cdot \mathbf{a}_n \\ \mathbf{a}_2^T \cdot \mathbf{a}_1 & \mathbf{a}_2^T \cdot \mathbf{a}_2 & \dots & \mathbf{a}_2^T \cdot \mathbf{a}_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n^T \cdot \mathbf{a}_1 & \mathbf{a}_n^T \cdot \mathbf{a}_2 & \dots & \mathbf{a}_n^T \cdot \mathbf{a}_n \end{bmatrix}$$

$$= \begin{bmatrix} \langle \mathbf{a}_1, \mathbf{a}_1 \rangle & \langle \mathbf{a}_1, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{a}_1, \mathbf{a}_n \rangle \\ \langle \mathbf{a}_2, \mathbf{a}_1 \rangle & \langle \mathbf{a}_2, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{a}_2, \mathbf{a}_n \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \mathbf{a}_n, \mathbf{a}_1 \rangle & \langle \mathbf{a}_n, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{a}_n, \mathbf{a}_n \rangle \end{bmatrix}. \quad (6.63)$$

Damit haben (6.62) und den Satz bewiesen. \square

Bemerkungen:

- i) Die *Spalten-Vektoren* einer *orthogonalen Matrix* sind *Einheitsvektoren*, die paarweise *senkrecht* aufeinander stehen.
- ii) Gemäss *Spalten-Vektor-Konstruktionsverfahren* werden durch eine *orthogonale Abbildung* die *Standard-Einheitsvektoren* $\hat{\mathbf{e}}_1, \dots, \hat{\mathbf{e}}_n \in \mathbb{R}^n$ auf paarweise aufeinander *senkrecht* stehende *Einheitsvektoren* abgebildet.
- iii) Die Schreibweise der *Komponenten* der *Einheitsmatrix* als *Koeffizienten* δ_{ij} wird auch *KRONECKER-Symbol* genannt.

Bevor wir die geometrischen Eigenschaften von *orthogonalen Abbildungen* weiter untersuchen können, benötigen wir eine Rechenregel.

Satz 6.8 Metrische Adjunktion

Seien $n \in \mathbb{N}^+$, $A \in \mathbb{M}(n, n, \mathbb{R})$ und $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, dann gilt

$$\boxed{\langle \mathbf{v}, A \cdot \mathbf{w} \rangle = \langle A^T \cdot \mathbf{v}, \mathbf{w} \rangle}. \quad (6.64)$$

Beweis: Weil das *Skalar-Produkt* von zwei *Vektoren* eine *reelle Zahl* ist, die als 1×1 -Matrix aufgefasst werden kann und alle 1×1 -Matrizen *symmetrisch* sind, erhalten wir

$$\begin{aligned} \underline{\langle \mathbf{v}, A \cdot \mathbf{w} \rangle} &= \langle \mathbf{v}, A \cdot \mathbf{w} \rangle^T = (\mathbf{v}^T \cdot A \cdot \mathbf{w})^T = \mathbf{w}^T \cdot A^T \cdot \mathbf{v} = \langle \mathbf{w}, A^T \cdot \mathbf{v} \rangle \\ &= \underline{\langle A^T \cdot \mathbf{v}, \mathbf{w} \rangle}. \end{aligned} \quad (6.65)$$

Damit haben (6.64) und den Satz bewiesen. \square

Bemerkungen:

- i) Die Rechenregel (6.64) hat nichts mit *Orthogonalität* zu tun, sie gilt für alle *quadratischen Matrizen*.
- ii) Weil die *Transposition* eine *Involution* ist, gilt auch

$$\langle \mathbf{v}, A^T \cdot \mathbf{w} \rangle = \left\langle (A^T)^T \cdot \mathbf{v}, \mathbf{w} \right\rangle = \langle A \cdot \mathbf{v}, \mathbf{w} \rangle. \quad (6.66)$$

- iii) Gemäss (6.64) darf man eine *Matrix* vor einem *Vektor* in einem *Skalar-Produkt* auf die andere Seite “schieben”, wenn man sie *transponiert*.

Die folgenden geometrischen Eigenschaften von *orthogonalen Abbildungen* bzw. *Matrizen* zeichnen diese aus.

Satz 6.9 *Invarianzen von orthogonalen Abbildungen*

Seien $n \in \mathbb{N}^+$ und $A \in O(n)$. Dann gilt für alle Vektoren $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ folgendes.

(a) *Skalar-Produkt-Invarianz*:

$$\langle A \cdot \mathbf{v}, A \cdot \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad (6.67)$$

(b) *Längen-Invarianz*:

$$|A \cdot \mathbf{v}| = |\mathbf{v}| \quad (6.68)$$

(c) *Winkel-Invarianz*:

$$\angle(A \cdot \mathbf{v}, A \cdot \mathbf{w}) = \angle(\mathbf{v}, \mathbf{w}) \quad (6.69)$$

Beweis: Mit Hilfe der Rechenregel (6.64) und weil A *orthogonal* ist, finden wir

$$\underline{\underline{\langle A \cdot \mathbf{v}, A \cdot \mathbf{w} \rangle}} = \langle A^T \cdot A \cdot \mathbf{v}, \mathbf{w} \rangle = \langle A^{-1} \cdot A \cdot \mathbf{v}, \mathbf{w} \rangle = \underline{\underline{\langle \mathbf{v}, \mathbf{w} \rangle}}. \quad (6.70)$$

Daraus folgt sofort auch

$$\underline{\underline{|A \cdot \mathbf{v}|}} = \sqrt{\langle A \cdot \mathbf{v}, A \cdot \mathbf{v} \rangle} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \underline{\underline{|\mathbf{v}|}}. \quad (6.71)$$

Wir betrachten die Fälle $0 \in \{\mathbf{v}, \mathbf{w}\}$ und $0 \notin \{\mathbf{v}, \mathbf{w}\}$ getrennt.

Fall 1: Es sei $0 \in \{\mathbf{v}, \mathbf{w}\}$. In diesem Fall gilt auch $0 \in \{A \cdot \mathbf{v}, A \cdot \mathbf{w}\}$ und es folgt

$$\underline{\underline{\angle(A \cdot \mathbf{v}, A \cdot \mathbf{w})}} = \frac{\pi}{2} = \underline{\underline{\angle(\mathbf{v}, \mathbf{w})}}. \quad (6.72)$$

Fall 2: Es sei $0 \notin \{\mathbf{v}, \mathbf{w}\}$. In diesem Fall gilt auch $0 \notin \{A \cdot \mathbf{v}, A \cdot \mathbf{w}\}$ und es folgt

$$\underline{\underline{\angle(A \cdot \mathbf{v}, A \cdot \mathbf{w})}} = \arccos\left(\frac{\langle A \cdot \mathbf{v}, A \cdot \mathbf{w} \rangle}{|A \cdot \mathbf{v}| \cdot |A \cdot \mathbf{w}|}\right) = \arccos\left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{v}| \cdot |\mathbf{w}|}\right) = \underline{\underline{\angle(\mathbf{v}, \mathbf{w})}}. \quad (6.73)$$

Damit haben wir alle Aussagen und den Satz bewiesen. □

Bemerkungen:

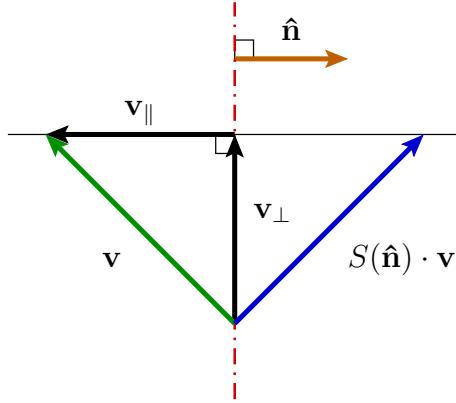
i) Die letzten beiden *Invarianz-Eigenschaften* aus Satz 6.9 werden auch *Längentreue* und *Winkeltreue* genannt. *Orthogonale Abbildungen* sind demnach *Kongruenz-Abbildungen* im Sinne der klassischen Geometrie. Das Umgekehrte gilt jedoch nicht, denn es gibt auch *Kongruenz-Abbildungen* die nicht *linear* sind und folglich auch nicht *orthogonal* sein können, z.B. *Translationen*.

ii) Tatsächlich sind die *Abbildungen* in $O(n)$ gerade die *Spiegelungen* und *Drehungen* in \mathbb{R}^n sowie deren *Verknüpfungen*.

6.3.3 Spezielle orthogonale Matrizen

6.3.3.1 Spiegelungen in nD

Wir betrachten $n \in \mathbb{N}^+$ und die *Spiegelung* in \mathbb{R}^n an einem $(n - 1)$ -dimensionalen Unterraum, welcher *senkrecht* steht zur *Richtung* $\hat{\mathbf{n}} \in \mathbb{R}^n$. Die Situation ist in der folgenden Skizze dargestellt.



Die *Spiegelungsmatrix* lässt sich durch eine einfache Formel aus $\hat{\mathbf{n}}$ berechnen.

Satz 6.10 HOUSEHOLDER-Formel

Die *Spiegelung* wird beschrieben durch die *Matrix*

$$S(\hat{\mathbf{n}}) = \mathbb{1} - 2 \cdot \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^T. \quad (6.74)$$

Beweis: Um das *Bild* eines Vektors $\mathbf{v} \in \mathbb{R}^n$ unter der *Spiegelung* zu berechnen, zerlegen wir \mathbf{v} in seine Anteile *parallel* und *senkrecht* zu $\hat{\mathbf{n}}$. Es sei also

$$\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp} \quad (6.75)$$

mit

$$\mathbf{v}_{\parallel} = \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \cdot \hat{\mathbf{n}} \quad \text{und} \quad \mathbf{v}_{\perp} = \mathbf{v} - \mathbf{v}_{\parallel}. \quad (6.76)$$

Für die *Spiegelungsmatrix* $S(\hat{\mathbf{n}})$ gilt gemäss Skizze

$$\begin{aligned} S(\hat{\mathbf{n}}) \cdot \mathbf{v} &= \mathbf{v} - 2 \cdot \mathbf{v}_{\parallel} = \mathbf{v} - 2 \cdot \langle \mathbf{v}, \hat{\mathbf{n}} \rangle \cdot \hat{\mathbf{n}} = \mathbf{v} - 2 \cdot \hat{\mathbf{n}} \cdot \langle \mathbf{v}, \hat{\mathbf{n}} \rangle = \mathbf{v} - 2 \cdot \hat{\mathbf{n}} \cdot \langle \hat{\mathbf{n}}, \mathbf{v} \rangle \\ &= \mathbf{v} - 2 \cdot \hat{\mathbf{n}} \cdot (\hat{\mathbf{n}}^T \cdot \mathbf{v}) = \mathbb{1} \cdot \mathbf{v} - 2 \cdot (\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^T) \cdot \mathbf{v} = \underline{\underline{(\mathbb{1} - 2 \cdot \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^T) \cdot \mathbf{v}}}. \end{aligned} \quad (6.77)$$

Dies impliziert (6.74) und wir haben den Satz bewiesen. \square

Bemerkungen:

- i) Es gilt $S(\hat{\mathbf{n}}) \in O(n)$.
- ii) In jedem Fall gilt

$$S(-\hat{\mathbf{n}}) = S(\hat{\mathbf{n}}) \quad (6.78)$$

d.h. es spielt keine Rolle auf welche Seite man $\hat{\mathbf{n}}$ wählt.

iii) Die Spiegelungsmatrix $S(\hat{\mathbf{n}})$ ist symmetrisch und es folgt

$$\underline{\underline{S^{-1}(\hat{\mathbf{n}}) = S^T(\hat{\mathbf{n}}) = S(\hat{\mathbf{n}})}} \quad (6.79)$$

Wie geometrisch offensichtlich, ist die Spiegelungsmatrix ihre eigene Inverse.

iv) Betrachtet man einen Normalen-Vektor $\mathbf{n} \in \mathbb{R}^n \setminus \{0\}$ der noch nicht normiert ist, dann kann man diese Normierung auch direkt in die HOUSEHOLDER-Formel (6.74) einbauen. Es gilt

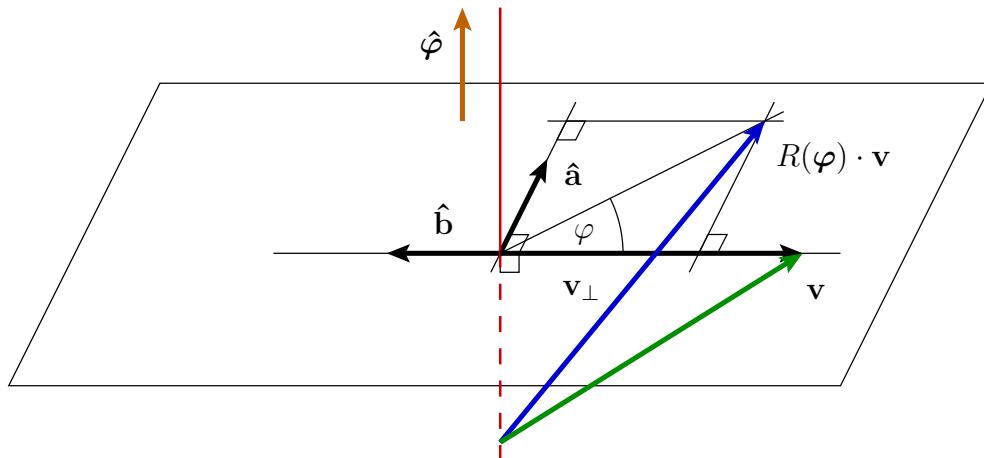
$$\underline{\underline{S(\mathbf{n}) = \mathbb{1} - 2 \cdot \hat{\mathbf{n}} \cdot \hat{\mathbf{n}}^T = \mathbb{1} - 2 \cdot \frac{\mathbf{n}}{|\mathbf{n}|} \cdot \frac{\mathbf{n}^T}{|\mathbf{n}|} = \mathbb{1} - 2 \cdot \frac{\mathbf{n} \cdot \mathbf{n}^T}{|\mathbf{n}|^2} = \mathbb{1} - 2 \cdot \frac{\mathbf{n} \cdot \mathbf{n}^T}{\langle \mathbf{n}, \mathbf{n} \rangle}}} \quad (6.80)$$

v) Beispiel-Codes zum Erzeugen von HOUSEHOLDER-Spiegelungsmatrizen gemäss (6.80) mit gängiger Software.

MATLAB/Octave	<code>S=@(n) eye(length(n))-2*n*n'/dot(n,n);</code>
Python/Numpy	<code>import numpy as np;</code> <code>def S(n):</code> <code>nn=n/np.linalg.norm(n);</code> <code>M=np.eye(n.shape[0])-2*nn@nn.T;</code> <code>return M;</code>

6.3.3.2 Rotationen in 3D

Wir betrachten die *Rotation* in \mathbb{R}^3 um den Winkel $\varphi \in \mathbb{R}$ rechtshändig um die Drehachse in Richtung $\hat{\varphi} \in \mathbb{R}^3$. Die Situation ist in der folgenden Skizze dargestellt.



Zunächst bemerken wir, dass sich aus den Bestimmungsstücken φ und $\hat{\varphi}$ der Rotation auf kanonische Weise ein Vektor bilden lässt gemäss

$$\varphi = \varphi \cdot \hat{\varphi}. \quad (6.81)$$

Dieser Vektor zeigt demnach in Richtung der Drehachse und hat gerade die Länge des Drehwinkels. Als nächstes betrachten wir die sogenannte Generator-Matrix.

Definition 6.17 Rotationsgenerator

Für jedes $\mathbf{w} \in \mathbb{R}^3$ ist der *Rotationsgenerator* die *Matrix*

$$J(\mathbf{w}) := \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}. \quad (6.82)$$

Im folgenden Satz stellen wir die wichtigsten *algebraischen* Eigenschaften des *Rotationsgenerators* zusammen.

Satz 6.11 Eigenschaften des Rotationsgenerators

Seien $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ mit $w = |\mathbf{w}|$, dann gilt folgendes.

- | | |
|---|---|
| (a) $J(\mathbf{w})$ ist <i>schiefsymmetrisch</i> | (d) $J^T(\mathbf{w}) = -J(\mathbf{w}) = J(-\mathbf{w})$ |
| (b) $J^2(\mathbf{w})$ ist <i>symmetrisch</i> | (e) $J^3(\mathbf{w}) = -w^2 \cdot J(\mathbf{w})$ |
| (c) $J(\mathbf{w}) \cdot \mathbf{v} = \mathbf{w} \times \mathbf{v}$ | (f) $J^4(\mathbf{w}) = -w^2 \cdot J^2(\mathbf{w})$ |

Bemerkungen:

- i) Gemäss Eigenschaft (c) beschreibt $J(\mathbf{w})$ gerade das GRASSMANN- *Vektor-Produkt* von links mit dem Vektor \mathbf{w} .
- ii) Für alle $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ gilt die berühmte *Drehimpuls-Kommutatonsrelation*

$$[J(\mathbf{v}), J(\mathbf{w})] = J(\mathbf{v} \times \mathbf{w}). \quad (6.83)$$

Mit Hilfe des *Rotationsgenerators* lässt sich die *Abbildungsmatrix* einer *Rotation* in 3D einfach ausdrücken.

Satz 6.12 RODRIGUES-Formel

Die *Rotation* in \mathbb{R}^3 um den Winkel $\varphi \in \mathbb{R}$ *rechtshändig* um die *Drehachse* in *Richtung* $\hat{\boldsymbol{\varphi}} \in \mathbb{R}^3$ wird beschrieben durch die *Abbildungsmatrix*

$$R(\varphi) = \mathbb{1} + (1 - \cos(\varphi)) \cdot J^2(\hat{\boldsymbol{\varphi}}) + \sin(\varphi) \cdot J(\hat{\boldsymbol{\varphi}}). \quad (6.84)$$

Beweis: Um das *Bild* eines Vektors $\mathbf{v} \in \mathbb{R}^3$ unter der *Rotation* zu berechnen, zerlegen wir \mathbf{v} in seine Anteile *parallel* und *senkrecht* zu $\hat{\boldsymbol{\varphi}}$. Gemäss Skizze und den Eigenschaften des GRASSMANN- *Vektor-Produkts* gilt

$$\hat{\boldsymbol{\varphi}} \times \mathbf{v} = \hat{\boldsymbol{\varphi}} \times \mathbf{v}_\perp = |\hat{\boldsymbol{\varphi}}| \cdot |\mathbf{v}_\perp| \cdot \hat{\mathbf{a}} = 1 \cdot v_\perp \cdot \hat{\mathbf{a}} = v_\perp \cdot \hat{\mathbf{a}} \quad (6.85)$$

$$\hat{\boldsymbol{\varphi}} \times (\hat{\boldsymbol{\varphi}} \times \mathbf{v}) = \hat{\boldsymbol{\varphi}} \times (v_\perp \cdot \hat{\mathbf{a}}) = v_\perp \cdot \hat{\boldsymbol{\varphi}} \times \hat{\mathbf{a}} = v_\perp \cdot \hat{\mathbf{b}}. \quad (6.86)$$

Mit Hilfe der Skizze und durch Einsetzen des *Rotationsgenerators* erhalten wir daraus

$$\begin{aligned} \underline{R(\varphi)} \cdot \mathbf{v} &= \mathbf{v} + v_\perp \cdot \hat{\mathbf{b}} - v_\perp \cdot \cos(\varphi) \cdot \hat{\mathbf{b}} + v_\perp \cdot \sin(\varphi) \cdot \hat{\mathbf{a}} \\ &= \mathbf{v} + (1 - \cos(\varphi)) \cdot v_\perp \cdot \hat{\mathbf{b}} + \sin(\varphi) \cdot v_\perp \cdot \hat{\mathbf{a}} \\ &= \mathbf{v} + (1 - \cos(\varphi)) \cdot \hat{\boldsymbol{\varphi}} \times (\hat{\boldsymbol{\varphi}} \times \mathbf{v}) + \sin(\varphi) \cdot \hat{\boldsymbol{\varphi}} \times \mathbf{v} \end{aligned}$$

$$\begin{aligned}
&= \mathbf{v} + (1 - \cos(\varphi)) \cdot J(\hat{\boldsymbol{\varphi}}) \cdot (J(\hat{\boldsymbol{\varphi}}) \cdot \mathbf{v}) + \sin(\varphi) \cdot J(\hat{\boldsymbol{\varphi}}) \cdot \mathbf{v} \\
&= \mathbf{v} + (1 - \cos(\varphi)) \cdot (J(\hat{\boldsymbol{\varphi}}) \cdot J(\hat{\boldsymbol{\varphi}})) \cdot \mathbf{v} + \sin(\varphi) \cdot J(\hat{\boldsymbol{\varphi}}) \cdot \mathbf{v} \\
&= \mathbb{1} \cdot \mathbf{v} + (1 - \cos(\varphi)) \cdot J^2(\hat{\boldsymbol{\varphi}}) \cdot \mathbf{v} + \sin(\varphi) \cdot J(\hat{\boldsymbol{\varphi}}) \cdot \mathbf{v} \\
&= \left(\mathbb{1} + (1 - \cos(\varphi)) \cdot J^2(\hat{\boldsymbol{\varphi}}) + \sin(\varphi) \cdot J(\hat{\boldsymbol{\varphi}}) \right) \cdot \mathbf{v}.
\end{aligned} \tag{6.87}$$

Dies impliziert (6.84) und wir haben den Satz bewiesen. \square

Bemerkungen:

i) Es gilt $R(\boldsymbol{\varphi}) \in O(n)$.

ii) In jedem Fall gilt

$$\underline{R^{-1}(\boldsymbol{\varphi}) = R^T(\boldsymbol{\varphi}) = R(-\boldsymbol{\varphi})}. \tag{6.88}$$

Wie geometrisch offensichtlich, ist die *Rotation* um den *Winkel* $-\varphi$ die *Inverse* der *Rotation* um den *Winkel* φ .

- iii) In der RODRIGUES-Formel (6.84) steht in beiden Termen jeweils $J(\hat{\boldsymbol{\varphi}})$ und nicht $J(\boldsymbol{\varphi})$. Die Verwechslung der beiden ist eine berüchtigte Fehlerquelle.
- iv) Beispiel-Codes zum Erzeugen von RODRIGUES-Rotationsmatrizen gemäss (6.84) mit gängiger Software.

MATLAB/Octave	<pre>J=@(w)[0,-w(3),w(2);w(3),0,-w(1);-w(2),w(1),0]; R=@(n,phi)eye(3)+(1-cos(phi))*J(n/norm(n))^2 +sin(phi)*J(n/norm(n));</pre>
Python/Numpy	<pre>import numpy as np; def J(w): M=np.array([[0,-w[2],w[1]],[w[2],0,-w[0]], [-w[1],w[0],0]]); return M; def R(phi,n): nn=n/np.linalg.norm(n); M=np.eye(3)+(1-np.cos(phi))*J(nn)@J(nn) +np.sin(phi)*J(nn); return M;</pre>

6.4 Spur & Determinante

6.4.1 Einleitung

Jede *lineare Abbildung* $a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ wird durch eine *Abbildungsmatrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ beschrieben. Aus diesen insgesamt n^2 *reelle Zahlen* lassen sich auf geschickte Weise *Kennzahlen* der *Matrix* berechnen, die bei der Untersuchung der algebraischen und geometrischen Eigenenschaften der zugehörigen *linearen Abbildung* eine grosse Rolle spielen. Zwei besonders wichtige *Kennzahlen* einer *Matrix* sind ihre *Spur* und ihre *Determinante*.

6.4.2 Spur

6.4.2.1 Definition

Für beliebige *quadratische Matrizen* machen wir folgende Definition.

Definition 6.18 Spur

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$. Die *Spur* der *Matrix* A ist die *reelle Zahl*

$$\text{tr}(A) = A_{11}^1 + A_{22}^2 + \dots + A_{nn}^n. \quad (6.89)$$

Bemerkungen:

- i) Die *Spur* einer *Matrix* ist also ganz einfach die *Summe* ihrer *Diagonalelemente*. Die Abkürzung tr kommt von der englischen Bezeichnung *trace*.
- ii) Offensichtlich gilt

$$\text{tr}(0) = 0 + \dots + 0 = n \cdot 0 = 0 \quad (6.90)$$

$$\text{tr}(\mathbb{1}) = 1 + \dots + 1 = n \cdot 1 = n. \quad (6.91)$$

- iii) Die *Spur* einer *diagonalen Matrix* ist gerade die *Summe* ihrer *Eigenwerte*. Es gilt also

$$\text{tr}(D) = \lambda_1 + \lambda_2 + \dots + \lambda_n. \quad (6.92)$$

- iv) Weil die *Diagonalelemente* einer *schiefsymmetrischen Matrix* alle verschwinden, muss gelten

$$A^T = -A \Rightarrow \text{tr}(A) = 0 + \dots + 0 = n \cdot 0 = 0. \quad (6.93)$$

- v) Je nach Art der *linearen Abbildung* (*Streckung*, *Spiegelung*, *Rotation*, *Projektion*, etc..) die eine *Abbildungsmatrix* beschreibt, kann ihre *Spur* ganz unterschiedliche geometrische Bedeutungen haben.
- vi) Für eine *Rotation* in \mathbb{R}^3 um den *Winkel* $\varphi \in \mathbb{R}$ *rechtshändig* um die *Drehachse* in *Richtung* $\hat{\varphi} \in \mathbb{R}^3$ findet man aus der *RODRIGUES-Formel* die *Spur*

$$\text{tr}(R(\varphi)) = 1 + 2 \cdot \cos(\varphi). \quad (6.94)$$

Man kann also aus der *Spur* den *Drehwinkel* φ ablesen.

vii) Ganz besondere Bedeutung kommt der *Spur* in der Quantenphysik und in der Chemie beim Aufstellen von *Charaktertafeln* zu, mit deren Hilfe man gewisse Eigenschaften von Molekülen beurteilen kann.

viii) Beispiel-Codes zur Berechnung der *Spur* mit gängiger Software.

MATLAB/Octave	<code>s=trace(A)</code>
Mathematica/WolframAlpha	<code>s=Tr[A]</code>
Python/Numpy	<code>import numpy as np; s=np.trace(A)</code>
Python/Sympy	<code>import sympy as sp; s=sp.trace(A)</code>

Beispiele:

- $\text{tr}\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\right) = 1 + 4 = 5$
- $\text{tr}\left(\begin{bmatrix} 1 & -5 & 9 \\ 8 & 2 & 7 \\ -3 & 1 & -3 \end{bmatrix}\right) = 1 + 2 - 3 = 0$
- $\text{tr}\left(\begin{bmatrix} 5 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 8 \end{bmatrix}\right) = 5 - 9 + 8 = 4$

6.4.2.2 Eigenschaften

Für die *Spur* einer *Matrix* gelten einige einfache *Rechenregeln*.

Satz 6.13 Rechenregeln der *Spur*

Seien $n \in \mathbb{N}^+$, $A, B \in \mathbb{M}(n, n, \mathbb{R})$ und $a \in \mathbb{R}$. Dann gelten folgende Rechenregeln.

- | | |
|---|--|
| (a) $\text{tr}(A^T) = \text{tr}(A)$ | (c) $\text{tr}(a \cdot A) = a \cdot \text{tr}(A)$ |
| (b) $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$ | (d) $\text{tr}(B \cdot A) = \text{tr}(A \cdot B)$ |

Beweis: Die Matrizen A und A^T haben die gleichen *Diagonalelemente* und folglich auch die gleiche *Spur*. Es gilt

$$\underline{\underline{\text{tr}(A + B)}} = \sum_{s=1}^n (A_s^s + B_s^s) = \sum_{s=1}^n A_s^s + \sum_{s=1}^n B_s^s = \underline{\underline{\text{tr}(A) + \text{tr}(B)}}. \quad (6.95)$$

Ebenso erhalten wir

$$\underline{\underline{\text{tr}(a \cdot A)}} = \sum_{s=1}^n a \cdot A_s^s = a \cdot \sum_{s=1}^n A_s^s = \underline{\underline{a \cdot \text{tr}(A)}}. \quad (6.96)$$

Ferner berechnen wir

$$\underline{\underline{\text{tr}(B \cdot A)}} = \sum_{s=1}^n \sum_{r=1}^n B_s^r \cdot A_s^r = \sum_{s=1}^n \sum_{r=1}^n A_s^r \cdot B_r^s = \sum_{r=1}^n \sum_{s=1}^n A_s^r \cdot B_r^s = \underline{\underline{\text{tr}(A \cdot B)}}. \quad (6.97)$$

Damit haben wir alle Aussagen und den Satz bewiesen. \square

Bemerkungen:

- i) Die *Spur* kann selbst als *lineare Abbildung* des Typs $\text{tr} : \mathbb{M}(n, n, \mathbb{R}) \rightarrow \mathbb{R}$ aufgefasst werden.
- ii) Bei der Anwendung der Regel (d) aus Satz 6.13 auf *Produkte* von mehr als zwei *Matrizen* ist Vorsicht geboten. Für drei *Matrizen* gilt

$$\text{tr}(A \cdot B \cdot C) = \text{tr}(C \cdot A \cdot B) = \text{tr}(B \cdot C \cdot A) \quad (6.98)$$

$$\text{tr}(A \cdot C \cdot B) = \text{tr}(B \cdot A \cdot C) = \text{tr}(C \cdot B \cdot A). \quad (6.99)$$

Die Gleichheit der Werte der beiden Zeilen muss jedoch nicht gelten. Die *Spur* bleibt im allgemeinen nur erhalten, wenn man die *Faktoren* eines *Matrix-Produkts zyklisch vertauscht*. Bei beliebigen Änderungen der Reihenfolge kann der Wert der *Spur* sich auch ändern.

6.4.3 Determinante

6.4.3.1 Definition

Die Definition der *Determinante* nach der bekannten *LEIBNIZ-Formel* basiert auf dem Begriff der *Permutationen*. Dies benötigt ein paar Vorbereitungen.

Definition 6.19 Permutation & symmetrische Gruppe

Sei $n \in \mathbb{N}^+$. Die *symmetrische Gruppe* vom *Grad* n ist die *Menge* der n -stelligen *Permutationen*, d.h.

$$S_n := \left\{ p : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid p \text{ ist bijektiv} \right\} \quad (6.100)$$

Bemerkungen:

- i) Eine *Permutation* $p \in S_n$ kann interpretiert werden als Umordnung bzw. Umsortierung der *natürlichen Zahlen* in $\{1, \dots, n\}$.
- ii) *Permutationen* haben ihre Hauptanwendung in der *Kombinatorik*, wenn es darum geht n unterscheidbare Objekte auf n unterscheidbare Plätze zu verteilen. S_6 entspricht dabei z.B. der *Menge* aller Möglichkeiten um 6 Autos auf 6 Parkplätze zu verteilen.
- iii) Üblicherweise wird eine *Permutation* $p \in S_n$ als $2 \times n$ -*Matrix* dargestellt gemäss

$$\begin{bmatrix} 1 & 2 & \dots & n \\ p(1) & p(2) & \dots & p(n) \end{bmatrix}. \quad (6.101)$$

Beispiele:

- $S_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

- $S_2 = \left\{ \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \right\}$
- $S_3 = \left\{ \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \right\}$

Es stellt sich natürlich die Frage, wie man die *Kardinalitäten* der *symmetrischen Gruppen* S_n möglichst einfach aus $n \in \mathbb{N}^+$ berechnen kann.

Satz 6.14 Kardinalität der *symmetrischen Gruppen*

Sei $n \in \mathbb{N}^+$. Es gilt

$$\# S_n = n! = 1 \cdot 2 \cdot \dots \cdot n. \quad (6.102)$$

Beweis: Um ein $p \in S_n$ zu definieren, wählen wir für jedes $k \in \{1, \dots, n\}$ einen *Funktionswert* $p(k) \in \{1, \dots, n\}$ aus. Weil p *bijektiv* ist, müssen wir dabei jedes *Element* von $\{1, \dots, n\}$ genau einmal verwenden. Für $p(1)$ haben wir also n Werte zur Auswahl, für $p(2)$ dann noch $(n-1)$, für $p(3)$ nur noch $(n-2)$ und immer so weiter, bis am Schluss für $p(n)$ nur noch ein einziger Wert übrig ist. Durch *Kombination* erhalten wir insgesamt die Anzahl Möglichkeiten von

$$N = n \cdot (n-1) \cdot (n-2) \cdot \dots \cdot 2 \cdot 1 = n!. \quad (6.103)$$

Weil jede dieser $N = n!$ Möglichkeiten genau eine der *Permutationen* aus S_n definiert, haben wir damit den Satz bewiesen. \square

Manche Eigenschaften einer *Permutation* lassen sich am Wert von zwei Zahlen ablesen. Dazu machen wir folgende Definitionen.

Definition 6.20 Inversionszahl & Vorzeichen

Seien $n \in \mathbb{N}^+$ und $p \in S_n$.

(a) Die *Inversionszahl* von p ist

$$\rho(p) := \# \left\{ (i; j) \in \{1, \dots, n\}^2 \mid i < j \wedge p(i) > p(j) \right\}. \quad (6.104)$$

(b) Das *Vorzeichen* von p ist

$$\sigma(p) := (-1)^{\rho(p)}. \quad (6.105)$$

Bemerkungen:

- i) Die *Inversionszahl* einer *Permutation* ist also die Anzahl der Pärchen aus Zahlen in $\{1, \dots, n\}$ für welche die *Permutation* die Reihenfolge umdreht.
- ii) Die *Inversionszahl* wird auch *Fehlstandszahl* genannt und für das *Vorzeichen* wird auch das lateinische Wort *Signum* verwendet.
- iii) In jedem Fall gilt

$$\rho(p) \in \left\{ 0, 1, 2, \dots, \frac{n \cdot (n-1)}{2} \right\}. \quad (6.106)$$

iv) In jedem Fall gilt $\sigma(p) \in \{-1, +1\}$. Es werden folgende Bezeichnungen verwendet:

$$\begin{aligned}\sigma(p) = -1 &\Leftrightarrow p \text{ ist eine } \textit{ungerade Permutation} \\ \sigma(p) = +1 &\Leftrightarrow p \text{ ist eine } \textit{gerade Permutation}.\end{aligned}\quad (6.107)$$

Mit diesen Vorbereitungen sind wir nun bereit, die *Determinante* einer beliebigen *quadratischen Matrix* mit Hilfe der *LEIBNIZ-Formel* zu definieren.

Definition 6.21 Determinante

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$. Die *Determinante* der *Matrix A* ist die *reelle Zahl*

$$\det(A) := \sum_{p \in S_n} \sigma(p) \cdot A^{p(1)}_1 \cdot \dots \cdot A^{p(n)}_n. \quad (6.108)$$

Bemerkungen:

i) Gemäss *LEIBNIZ-Formel* (6.108) wird die *Determinante* einer *quadratischen Matrix A* durch folgende Schritte gebildet.

- S1** Man sucht sich alle Möglichkeiten, um n *Komponenten* aus A auszuwählen, so dass aus jeder Zeile und jeder Spalte genau eine *Komponente* vertreten ist.
- S2** Man berechnet für jede Möglichkeit aus Schritt S1 das *Produkt* der gefundenen *Komponenten*.
- S3** Man multipliziert jedes dieser *Produkte* aus Schritt S2 mit dem *Vorzeichen* der *Permutation* zwischen Zeilen- und Spaltenindizes.
- S4** Man *addiert* alle *Produkte* mit den entsprechenden *Vorzeichen*.

ii) Für $n = 2$ gibt es nur die folgenden $n! = 2! = 1 \cdot 2 = 2$ *Permutationen*.

k	$p_k(1)$	$p_k(2)$	$\rho(p_k)$	$\sigma(p_k)$
1	1	2	0	+1
2	2	1	1	-1

(6.109)

Die *Determinante* einer 2×2 -*Matrix* ist folglich

$$\det(A) = A^1_1 \cdot A^2_2 - A^2_1 \cdot A^1_2. \quad (6.110)$$

iii) Für $n = 3$ gibt es die folgenden $n! = 3! = 1 \cdot 2 \cdot 3 = 6$ *Permutationen*.

k	$p_k(1)$	$p_k(2)$	$p_k(3)$	$\rho(p_k)$	$\sigma(p_k)$
1	1	2	3	0	+1
2	2	3	1	2	+1
3	3	1	2	2	+1
4	3	2	1	3	-1
5	1	3	2	1	-1
6	2	1	3	1	-1

(6.111)

Die *Determinante* einer 3×3 -Matrix ist folglich

$$\det(A) = A_{11}^1 \cdot A_{22}^2 \cdot A_{33}^3 + A_{12}^2 \cdot A_{23}^3 \cdot A_{31}^1 + A_{13}^3 \cdot A_{21}^1 \cdot A_{32}^2 - A_{11}^3 \cdot A_{22}^2 \cdot A_{33}^1 - A_{11}^1 \cdot A_{23}^3 \cdot A_{32}^2 - A_{12}^2 \cdot A_{21}^1 \cdot A_{33}^3. \quad (6.112)$$

iv) Offensichtlich gilt in allen *Dimensionen*

$$\det(0) = 0 \quad \text{und} \quad \det(\mathbb{1}) = 1. \quad (6.113)$$

v) Die *Determinante* einer *diagonalen Matrix* ist gerade das *Produkt* ihrer *Eigenwerte*. Es gilt also

$$\det(D) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n. \quad (6.114)$$

vi) Die *Determinante* einer *linken unteren oder rechten oberen Dreiecksmatrix* reduziert sich auf das *Produkt* ihrer *Diagonalelemente*. Es gilt also

$$\det(L) = \det \begin{pmatrix} L_1^1 & 0 & 0 & \dots & 0 \\ L_1^2 & L_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ L_1^n & L_2^n & L_3^n & \dots & L_n^n \end{pmatrix} = L_1^1 \cdot L_2^2 \cdot \dots \cdot L_n^n \quad (6.115)$$

$$\det(R) = \det \begin{pmatrix} R_1^1 & R_1^2 & R_1^3 & \dots & R_1^n \\ 0 & R_2^2 & R_2^3 & \dots & R_2^n \\ \vdots & \dots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & R_n^n \end{pmatrix} = R_1^1 \cdot R_2^2 \cdot \dots \cdot R_n^n. \quad (6.116)$$

vii) Beispiel-Codes zur Berechnung der *Determinante* mit gängiger Software.

MATLAB/Octave	<code>d=det(A)</code>
Mathematica/WolframAlpha	<code>d=Det[A]</code>
Python/Numpy	<code>import numpy as np;</code> <code>d=np.linalg.det(A)</code>
Python/Sympy	<code>import sympy as sp;</code> <code>d=sp.det(A)</code>

Beispiele:

- $\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 1 \cdot 4 - 3 \cdot 2 = 4 - 6 = 2$
- $\det \begin{pmatrix} 2 & 4 \\ 3 & 6 \end{pmatrix} = 2 \cdot 6 - 3 \cdot 4 = 12 - 12 = 0$
- $\det \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = 2 \cdot 3 - 0 \cdot 0 = 6 - 0 = 6$

$$\bullet \det \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 4 \end{pmatrix} = 3 \cdot (-2) \cdot 4 = -24$$

$$\bullet \det \begin{pmatrix} 3 & 0 & 0 \\ 7 & -2 & 0 \\ 5 & -8 & 4 \end{pmatrix} = 3 \cdot (-2) \cdot 4 = -24$$

6.4.3.2 Eigenschaften

Für die *Determinante* einer *Matrix* gelten einige einfache *Rechenregeln*.

Satz 6.15 Rechenregeln der Determinante

Seien $n \in \mathbb{N}^+$, $A, B \in \mathbb{M}(n, n, \mathbb{R})$ und $a \in \mathbb{R}$. Dann gelten folgende Rechenregeln.

$$(a) \det(A^T) = \det(A)$$

$$(c) \boxed{\det(A \cdot B) = \det(A) \cdot \det(B)}$$

$$(b) \det(a \cdot A) = a^n \cdot \det(A)$$

$$(d) \det(A^{-1}) = \frac{1}{\det(A)} \quad \text{falls } A \text{ regulär}$$

Beweis: Die Aussagen (a) und (b) sind klar, während (c) nur mit grossem Aufwand gezeigt werden kann. Falls A regulär ist, dann hat sie eine *Inversen* $A^{-1} \in \mathbb{M}(n, n, \mathbb{R})$ und es gilt

$$A^{-1} \cdot A = \mathbb{1} \quad \boxed{\det(\dots)} \quad (6.117)$$

$$\Rightarrow \det(A^{-1} \cdot A) = \det(\mathbb{1}) \quad (6.118)$$

$$\Rightarrow \det(A^{-1}) \cdot \det(A) = 1 \quad \boxed{: \det(A)}. \quad (6.119)$$

Daraus folgt $\det(A) \neq 0$ und nach der *Division* erhalten wir

$$\underline{\underline{\det(A^{-1}) = \frac{1}{\det(A)}}}. \quad (6.120)$$

Damit haben wir die Aussage (d) bewiesen. □

Bemerkungen:

i) Für die *Determinante* gibt es keine allgemeingültige *Summen-Regel*. Es gibt *quadratische Matrizen* A und B , für welche gilt $\det(A + B) = \det(A) + \det(B)$ aber auch solche für die wir $\det(A + B) \neq \det(A) + \det(B)$ finden.

ii) Die Rechenregel (c) aus Satz 6.15 ist eine äusserst wichtige Eigenschaft der *Determinante*. Sie hat unter anderem zur Folge, dass gilt

$$\det(B \cdot A) = \det(B) \cdot \det(A) = \det(A) \cdot \det(B) = \det(A \cdot B). \quad (6.121)$$

Das heisst, man darf innerhalb einer *Determinante* die *Faktoren* eines *Matrix-Produkts* vertauschen. Im Gegensatz zur Situation bei der *Spur* gilt dies auch bei mehr als zwei *Faktoren* für beliebige Änderungen der Reihenfolge der *Faktoren*. Für drei *Faktoren* erhalten wir

$$\begin{aligned} \det(A \cdot B \cdot C) &= \det(C \cdot A \cdot B) = \det(B \cdot C \cdot A) = \det(A \cdot C \cdot B) = \det(B \cdot A \cdot C) \\ &= \det(C \cdot B \cdot A) = \det(A) \cdot \det(B) \cdot \det(C). \end{aligned} \quad (6.122)$$

- iii)** Eine *quadratische Matrix* U heisst *unimodular*, falls gilt $\det(U) = 1$. Für eine beliebige *quadratische Matrix* A folgt dann

$$\det(U \cdot A) = \det(U) \cdot \det(A) = 1 \cdot \det(A) = \det(A). \quad (6.123)$$

Ferner ist das *Produkt* von zwei *unimodularen Matrizen* wieder *unimodular*.

Gewisse einfache Modifikationen an einer *Matrix* führen auch zu einfachen Veränderungen ihrer *Determinante*.

Satz 6.16 Modifikationsregeln der *Determinante*

Sei $n \in \mathbb{N}^+$, $A \in \mathbb{M}(n, n, \mathbb{R})$ und $a \in \mathbb{R}$. Dann gelten folgende Modifikationsregeln.

- (a) Zeilentausch $\Rightarrow \det(A) \mapsto -\det(A)$.
- (b) Spaltentausch $\Rightarrow \det(A) \mapsto -\det(A)$.
- (c) Multiplikation einer Zeile mit $a \Rightarrow \det(A) \mapsto a \cdot \det(A)$.
- (d) Multiplikation einer Spalte mit $a \Rightarrow \det(A) \mapsto a \cdot \det(A)$.

Besonders nützlich ist auch die folgende Modifikationsregel.

Satz 6.17 *Invarianz* der *Determinante* unter einem *GAUSS-Schritt*

Subtrahiert man von einer Zeile einer $n \times n$ -*Matrix* ein beliebiges *Vielfaches* einer andern Zeile, dann ändert sich die *Determinante* der *Matrix* nicht.

Die Modifikationsregeln ermöglichen es, *Determinanten* mit Hilfe des *GAUSS-Verfahrens* zu berechnen. Als Beispiel berechnen wir die *Determinante* der *Matrix*

$$A = \begin{bmatrix} 1 & 2 & -3 \\ 2 & -4 & 1 \\ 2 & 2 & 8 \end{bmatrix}. \quad (6.124)$$

Wir erhalten

$$\begin{aligned} \underline{\underline{\det(A)}} &= \begin{vmatrix} 1 & 2 & -3 \\ 2 & -4 & 1 \\ 2 & 2 & 8 \end{vmatrix} = 2 \begin{vmatrix} [1] & 2 & -3 \\ 2 & -4 & 1 \\ 1 & 1 & 4 \end{vmatrix} \cdot 2 = \begin{vmatrix} [1] & 2 & -3 \\ 0 & -8 & 7 \\ 0 & -1 & 7 \end{vmatrix} \cdot 2 \\ &= 8 \begin{vmatrix} [1] & 2 & -3 \\ 0 & [1] & -7 \\ 0 & 8 & -7 \end{vmatrix} \cdot 2 \cdot (-1)^3 = \begin{vmatrix} [1] & 2 & -3 \\ 0 & [1] & -7 \\ 0 & 0 & [49] \end{vmatrix} \cdot 2 \cdot (-1) = 1 \cdot 1 \cdot 49 \cdot 2 \cdot (-1) = \underline{\underline{-98}}. \end{aligned} \quad (6.125)$$

Eine der Hauptanwendungen von *Determinanten* ist die Prüfung einer *quadratischen Matrix* auf *Singularität* bzw. *Regularität*. Es gilt nämlich der folgende, bemerkenswerte Satz.

Satz 6.18 *Regularitätssatz*

Eine *quadratische Matrix* A ist genau dann *regulär*, wenn gilt $\det(A) \neq 0$.

Bemerkungen:

- i) Gilt $\det(A) \neq 0$, dann ist die *quadratische Matrix* A *regulär*, d.h. sie hat eine *Inverse* A^{-1} und die zugehörige *lineare Abbildung* a ist *bijektiv* und hat eine *Umkehrabbildung* a^{-1} .

- ii)** Gilt $\det(A) = 0$, dann ist die *quadratische Matrix A singulär*, d.h. sie hat keine *Inverse* und die zugehörige *lineare Abbildung a* ist nicht *bijektiv*.

Wir werden später sehen, dass der *Determinante* auch eine geometrische Bedeutung zukommt. In diesem Abschnitt erwähnen wir dazu nur den folgenden Satz über *Determinanten von orthogonalen Matrizen*.

Satz 6.19 *Determinante einer orthogonalen Matrix*

Seien $n \in \mathbb{N}^+$ und $A \in O(n)$, dann gilt $\det(A) \in \{-1, 1\}$.

Beweis: Für eine *orthogonale Matrix A* muss gelten

$$\det(A) = \det(A^T) = \det(A^{-1}) = \frac{1}{\det(A)} \quad | \cdot \det(A) \quad (6.126)$$

$$\Leftrightarrow (\det(A))^2 = 1. \quad (6.127)$$

Daraus folgt

$$\underline{\underline{\det(A) \in \{-1, 1\}}}. \quad (6.128)$$

Damit haben den Satz bewiesen. \square

Bemerkungen:

- i)** Die Umkehrung des Satzes 6.19 gilt nicht. Dazu betrachten wir z.B. die *Matrix*

$$A = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix}. \quad (6.129)$$

Wir berechnen leicht, dass zwar

$$\det(A) = \frac{1}{2} \cdot 2 - 0 \cdot 0 = 1 - 0 = 1, \quad (6.130)$$

aber deswegen ist A noch lange nicht *orthogonal*, denn es gilt

$$A^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \neq \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 2 \end{bmatrix} = A^T. \quad (6.131)$$

- ii)** Die *Determinante* teilt die *Menge* $O(n)$ in zwei Teile auf. Man definiert

$$O^\pm(n) := \{A \in O(n) \mid \det(A) = \pm 1\}. \quad (6.132)$$

Ferner bezeichnet man

$$SO(n) := O^+(n) \quad (6.133)$$

als *spezielle orthogonale Gruppe* in n Dimensionen.

- iii)** Man kann zeigen, dass $O^-(n)$ alle *Spiegelungsmatrizen* und $O^+(n)$ alle *Rotationsmatrizen* enthält.

6.4.4 Mass-Formeln

Mit Hilfe der *Determinante* lassen sich Begriffe wie *Länge*, *Fläche*, *Volumen* und entsprechende Verallgemeinerungen in beliebig *dimensionalen EUKLID-Räumen* einführen und berechnen. Dazu machen wir folgende Definition.

Definition 6.22 GRAM-Matrix

Seien $m, n \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$. Die *GRAM-Matrix* dieser Vektoren ist

$$G(\mathbf{v}_1; \dots; \mathbf{v}_m) := \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_1, \mathbf{v}_m \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_2, \mathbf{v}_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \mathbf{v}_m, \mathbf{v}_1 \rangle & \langle \mathbf{v}_m, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_m, \mathbf{v}_m \rangle \end{bmatrix}. \quad (6.134)$$

Bemerkungen:

- i) Die *Komponenten* der *GRAM-Matrix* sind gerade alle möglichen *Skalar-Produkte*, die sich aus den Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_m$ bilden lassen. Davon gibt es insgesamt m^2 und konsequenterweise gilt $G \in \mathbb{M}(m, m, \mathbb{R})$.
- ii) Wegen der *Symmetrie* des *Skalar-Produkts* muss gelten

$$G^T = G, \quad (6.135)$$

d.h. G ist *symmetrisch*.

- iii) Wegen der *positiven Definitheit* des *Skalar-Produkts* muss gelten

$$\det(G) \geq 0. \quad (6.136)$$

- iv) Für die *Standard-Einheitsvektoren* gilt

$$G(\hat{\mathbf{e}}_1; \dots; \hat{\mathbf{e}}_n) = \mathbb{1}. \quad (6.137)$$

- v) Es sei $A \in \mathbb{M}(n, m, \mathbb{R})$ die *Matrix*, deren *Spalten* gerade die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_m$ sind, d.h.

$$A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_m]. \quad (6.138)$$

Dann gilt

$$\begin{aligned} \underline{\underline{A^T \cdot A}} &= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_m^T \end{bmatrix} \cdot \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_m \end{bmatrix} = \begin{bmatrix} \mathbf{v}_1^T \cdot \mathbf{v}_1 & \mathbf{v}_1^T \cdot \mathbf{v}_2 & \dots & \mathbf{v}_1^T \cdot \mathbf{v}_m \\ \mathbf{v}_2^T \cdot \mathbf{v}_1 & \mathbf{v}_2^T \cdot \mathbf{v}_2 & \dots & \mathbf{v}_2^T \cdot \mathbf{v}_m \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{v}_m^T \cdot \mathbf{v}_1 & \mathbf{v}_m^T \cdot \mathbf{v}_2 & \dots & \mathbf{v}_m^T \cdot \mathbf{v}_m \end{bmatrix} \\ &= \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_1, \mathbf{v}_m \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_2, \mathbf{v}_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \mathbf{v}_m, \mathbf{v}_1 \rangle & \langle \mathbf{v}_m, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_m, \mathbf{v}_m \rangle \end{bmatrix} = \underline{\underline{G}}. \end{aligned} \quad (6.139)$$

Wir betrachten die folgende Definition.

Definition 6.23 *Mass*

Seien $m, n \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{R}^n$ mit GRAM-Matrix G . Das *Mass* der Vektoren ist

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) := \sqrt{\det(G)}. \quad (6.140)$$

Bemerkungen:

- i) Wegen $\det(G) \geq 0$ kann die *Wurzel* immer in \mathbb{R} gezogen werden.
- ii) Das *Mass* ist die Verallgemeinerung der Begriffe *Länge*, *Fläche* und *Volumen* auf beliebige *Dimensionen*. Insbesondere gilt folgendes.

$$\mu(\mathbf{v}) = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = |\mathbf{v}| \equiv \text{Länge von } \mathbf{v} \quad (6.141)$$

$$\mu(\mathbf{v}_1; \mathbf{v}_2) \equiv \text{Fläche des von } \mathbf{v}_1 \text{ und } \mathbf{v}_2 \text{ aufgespannten Parallelogramms} \quad (6.142)$$

$$\mu(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3) \equiv \text{Volumen des von } \mathbf{v}_1, \mathbf{v}_2 \text{ und } \mathbf{v}_3 \text{ aufgespannten Spats} \quad (6.143)$$

- iii) Für die *Standard-Einheitsvektoren* gilt

$$\mu(\hat{\mathbf{e}}_1; \dots; \hat{\mathbf{e}}_n) = \sqrt{\det(\mathbb{1})} = \sqrt{1} = 1. \quad (6.144)$$

- iv) Für $m > n$ ist die GRAM-Matrix G in jedem Fall *singulär* und es folgt

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) = 0. \quad (6.145)$$

Falls die Anzahl *Vektoren* gerade der *Dimension* gleicht, dann lässt sich die Berechnung des *Masses* vereinfachen.

Satz 6.20 *Mass-Formel* in voller Dimension

Seien $n \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbb{R}^n$ gerade die *Spalten* der Matrix

$$A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n] \in \mathbb{M}(n, n, \mathbb{R}). \quad (6.146)$$

Dann gilt

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_n) := |\det(A)|. \quad (6.147)$$

Beweis: Weil A eine *quadratische Matrix* ist, folgt

$$\begin{aligned} \underline{\mu(\mathbf{v}_1; \dots; \mathbf{v}_n)} &= \sqrt{\det(G)} = \sqrt{\det(A^T \cdot A)} = \sqrt{\det(A^T) \cdot \det(A)} = \sqrt{\det(A) \cdot \det(A)} \\ &= \sqrt{\det^2(A)} = \underline{|\det(A)|}. \end{aligned} \quad (6.148)$$

Damit haben wir den Satz bewiesen. □

6.5 Eigenwerte & Eigenvektoren

6.5.1 Einleitung

Wir betrachten

$$A = \begin{bmatrix} 7 & 2 \\ 3 & 8 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \mathbf{E}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \mathbf{E}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \quad (6.149)$$

Dann gilt

$$A \cdot \mathbf{u} = \begin{bmatrix} 7 & 2 \\ 3 & 8 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \cdot 2 + 2 \cdot 1 \\ 3 \cdot 2 + 8 \cdot 1 \end{bmatrix} = \begin{bmatrix} 16 \\ 14 \end{bmatrix} \quad (6.150)$$

$$\underline{\underline{A}} \cdot \underline{\underline{\mathbf{E}_1}} = \begin{bmatrix} 7 & 2 \\ 3 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \cdot 1 + 2 \cdot (-1) \\ 3 \cdot 1 + 8 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix} = 5 \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \underline{\underline{5}} \cdot \underline{\underline{\mathbf{E}_1}} \quad (6.151)$$

$$\underline{\underline{A}} \cdot \underline{\underline{\mathbf{E}_2}} = \begin{bmatrix} 7 & 2 \\ 3 & 8 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 7 \cdot 2 + 2 \cdot 3 \\ 3 \cdot 2 + 8 \cdot 3 \end{bmatrix} = \begin{bmatrix} 20 \\ 30 \end{bmatrix} = 10 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \underline{\underline{10}} \cdot \underline{\underline{\mathbf{E}_2}}. \quad (6.152)$$

Beobachtungen: Bei der Wirkung von A auf \mathbf{u} fällt nichts besonderes auf. Die Wirkung von A auf die Vektoren \mathbf{E}_1 und \mathbf{E}_2 ist jedoch eine einfache *Streckung* um die Faktoren 5 bzw. 10.

6.5.2 Definition & Eigenschaften

Wir machen folgende Definition.

Definition 6.24 *Eigenwert & Eigenvektor*

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$. Ein Vektor $\mathbf{E} \in \mathbb{R}^n \setminus \{0\}$ heisst *Eigenvektor* von A zum *Eigenwert* $\lambda \in \mathbb{R}$, falls gilt

$$\boxed{A \cdot \mathbf{E} = \lambda \cdot \mathbf{E}}. \quad (6.153)$$

Bemerkungen:

- i) Weil $A \cdot 0 = 0$ für jede Matrix A gilt, zählt $0 \in \mathbb{R}^n$ nicht als *Eigenvektor* und wird bei der Definition explizit ausgeschlossen. Die Zahl $0 \in \mathbb{R}$ kann jedoch als *Eigenwert* auftreten.
- ii) Die Menge aller *Eigenwerte* einer Matrix A wird *Spektrum* von A genannt. Man schreibt

$$\text{Spec}(A) := \{\lambda_1, \dots, \lambda_m\}. \quad (6.154)$$

- iii) Sind \mathbf{E}_1 und \mathbf{E}_2 *Eigenvektoren* von A zum gleichen *Eigenwert* λ und $a, b \in \mathbb{R}$, dann gilt

$$\begin{aligned} A \cdot (a \cdot \mathbf{E}_1 + b \cdot \mathbf{E}_2) &= a \cdot A \cdot \mathbf{E}_1 + b \cdot A \cdot \mathbf{E}_2 = a \cdot \lambda \cdot \mathbf{E}_1 + b \cdot \lambda \cdot \mathbf{E}_2 \\ &= \lambda \cdot (a \cdot \mathbf{E}_1 + b \cdot \mathbf{E}_2). \end{aligned} \quad (6.155)$$

Das heisst auch jede *Linearkombination* der Form $(a \cdot \mathbf{E}_1 + b \cdot \mathbf{E}_2)$ ist wieder *Eigenvektor* von A zum gleichen *Eigenwert* λ . Die *Eigenvektoren* zu einem *Eigenwert* bilden daher wieder einen *Vektorraum*, den sogenannten *Eigenraum* E_λ zum *Eigenwert* λ .

- iv) Weil alle Vielfache eines *Eigenvektors* wieder *Eigenvektoren* zum gleichen *Eigenwert* sind, gibt es zu jedem *Eigenwert* λ_k einer Matrix auch einen *Einheitseigenvektor* $\hat{\mathbf{E}}_k$.

6.5.3 Charakteristisches Polynom

Wir machen folgende Definition.

Definition 6.25 *Charakteristisches Polynom*

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$. Die *Funktion*

$$\boxed{\begin{aligned} p_A : \mathbb{R} &\rightarrow \mathbb{R} \\ \lambda &\mapsto p_A(\lambda) := \det(\lambda \cdot \mathbb{1} - A) \end{aligned}} \quad (6.156)$$

heisst *charakteristisches Polynom* von A .

Bemerkungen:

- i) Der *Funktionsterm* von p_A wirkt zunächst irritierend. Setzt man jedoch die *Komponenten* von $\mathbb{1}$ und A ein und rechnet die *Determinante* aus, dann erhält man tatsächlich ein *Polynom* in λ .
- ii) Offensichtlich gilt

$$p_0(\lambda) = \det(\lambda \cdot \mathbb{1} - 0) = \det(\lambda \cdot \mathbb{1}) = \lambda^n \cdot \det(\mathbb{1}) = \lambda^n \cdot 1 = \lambda^n \quad (6.157)$$

$$\begin{aligned} p_{\mathbb{1}}(\lambda) &= \det(\lambda \cdot \mathbb{1} - \mathbb{1}) = \det((\lambda - 1) \cdot \mathbb{1}) = (\lambda - 1)^n \cdot \det(\mathbb{1}) = (\lambda - 1)^n \cdot 1 \\ &= (\lambda - 1)^n. \end{aligned} \quad (6.158)$$

Das *charakteristische Polynom* hat ein paar wichtige Eigenschaften.

Satz 6.21 Eigenschaften des *charakteristischen Polynoms*

Seien $n \in \mathbb{N}^+$ und $A \in \mathbb{M}(n, n, \mathbb{R})$. Dann ist p_A ein *Polynom* vom *Grad* n der Form

$$p_A(\lambda) = a_n \cdot \lambda^n + a_{n-1} \cdot \lambda^{n-1} + \dots + a_1 \cdot \lambda + a_0, \quad (6.159)$$

wobei in jedem Fall gilt

$$\mathbf{(a)} \quad a_n = 1 \qquad \mathbf{(b)} \quad a_{n-1} = -\text{tr}(A) \qquad \mathbf{(c)} \quad a_0 = (-1)^n \cdot \det(A)$$

Beweis: Es gilt

$$\underline{\underline{a_0}} = p_A(0) = \det(0 \cdot \mathbb{1} - A) = \det(-A) = \underline{\underline{(-1)^n \cdot \det(A)}} \quad (6.160)$$

Damit haben wir die Aussage (c) bewiesen. □

Bemerkungen:

- i) Für eine *singuläre Matrix* gilt $\det(A) = 0$ und somit $a_0 = 0$. Das *charakteristische Polynom* hat dann die Form

$$p_A(\lambda) = \lambda^n - \text{tr}(A) \cdot \lambda^{n-1} + \dots + a_1 \cdot \lambda = \lambda \cdot (\lambda^{n-1} - \text{tr}(A) \cdot \lambda^{n-2} + \dots + a_1). \quad (6.161)$$

- ii) Für $n = 2$ ist p_A eine *quadratische Funktion* und es folgt

$$p_A(\lambda) = \lambda^2 - \text{tr}(A) \cdot \lambda + \det(A). \quad (6.162)$$

iii) Beispiel-Codes zur Berechnung des *charakteristischen Polynoms* mit gängiger Software.

MATLAB/Octave	<code>p=charpoly(A)</code>
Mathematica/WolframAlpha	<code>p=CharacteristicPolynomial[A,x]</code>
Python/Numpy	<code>import numpy as np; p=np.poly(A)</code>
Python/Sympy	<code>import sympy as sp; p=A.charpoly()</code>

Beispiele:

- Wir betrachten die *Matrix*

$$A = \begin{bmatrix} 7 & 2 \\ 3 & 8 \end{bmatrix}. \quad (6.163)$$

Es folgt

$$\text{tr}(A) = 7 + 8 = 15 \quad (6.164)$$

$$\det(A) = 7 \cdot 8 - 3 \cdot 2 = 56 - 6 = 50 \quad (6.165)$$

$$\underline{\underline{p_A(\lambda)}} = \lambda^2 - \text{tr}(A) \cdot \lambda + \det(A) = \underline{\underline{\lambda^2 - 15\lambda + 50}}. \quad (6.166)$$

- $A = \begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix} \Rightarrow \dots \Rightarrow p_A(\lambda) = \lambda^2 - 10\lambda + 24$
- $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \Rightarrow \dots \Rightarrow p_A(\lambda) = \lambda^2 - 1$

6.5.4 Eigenwerte & Eigenvektoren berechnen

Um die *Eigenwerte* und *Eigenvektoren* einer *Matrix* zu berechnen, hilft uns der folgende Satz weiter.

Satz 6.22 *Eigenwerte* sind Nullstellen des *charakteristischen Polynoms*

Seien $n \in \mathbb{N}^+$, $A \in \mathbb{M}(n, n, \mathbb{R})$ und $\lambda \in \mathbb{R}$. Dann gilt

$$\boxed{\lambda \in \text{Spec}(A) \Leftrightarrow p_A(\lambda) = 0.} \quad (6.167)$$

Beweis: Für einen *Eigenwert* $\lambda \in \mathbb{R}$ und einen zugehörigen *Eigenvektor* $\mathbf{E} \in \mathbb{R}^n \setminus \{0\}$ gilt

$$A \cdot \mathbf{E} = \lambda \cdot \mathbf{E} \quad | - A \cdot \mathbf{E} \quad (6.168)$$

$$\Leftrightarrow \lambda \cdot \mathbf{E} - A \cdot \mathbf{E} = 0 \quad (6.169)$$

$$\Leftrightarrow (\lambda \cdot \mathbb{1} - A) \cdot \mathbf{E} = 0. \quad (6.170)$$

Dies ist ein *homogenes, lineares Gleichungssystem* für \mathbf{E} , das genau dann von 0 verschiedene Lösungen hat, wenn die *Matrix* $\lambda \cdot \mathbb{1} - A$ singulär ist. Demnach muss gelten

$$\underline{\underline{0}} = \det(\lambda \cdot \mathbb{1} - A) = \underline{\underline{p_A(\lambda)}}. \quad (6.171)$$

Damit haben wir den Satz bewiesen. □

Bemerkungen:

- i) Die *Eigenwerte* sind gerade die *Nullstellen* des *charakteristischen Polynoms*.
- ii) Ein *Polynom* vom *Grad* n kann maximal n *Nullstellen* haben. Dementsprechend kann eine *Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ auch höchstens n *Eigenwerte* haben.
- iii) Die *Nullstellen* eines *Polynoms* lassen sich bekanntlich nur für $n \in \{1, 2, 3, 4\}$ direkt durch Formeln berechnen. Man kann beweisen(!), dass es für *Polynome* vom *Grad* fünf oder höher keine direkten Formeln für die *Nullstellen* geben kann. Somit ist die Berechnung der *Eigenwerte* für $n \geq 5$ mit Hilfe des *charakteristischen Polynoms* sehr schwierig.
- iv) Für $n = 2$ kann die *Mitternachtsformel* für *quadratische Gleichungen* eingesetzt werden, um die *Eigenwerte* zu berechnen. Aus (6.162) folgt, dass eine *Matrix* $A \in \mathbb{M}(2, 2, \mathbb{R})$ genau dann *Eigenwerte* hat, wenn gilt

$$D = \text{tr}^2(A) - 4 \cdot \det(A) \geq 0. \quad (6.172)$$

In diesem Fall erhält man

$$\lambda_{1,2} = \frac{\text{tr}(A) \pm \sqrt{D}}{2}. \quad (6.173)$$

- v) Sind die *Eigenwerte* einer *Matrix* erst einmal bekannt, dann können die zugehörigen *Eigenvektoren* durch lösen des *linearen Gleichungssystems* (6.170) gefunden werden.
- vi) Beispiel-Codes zur Berechnung der *Eigenwerte* und *Eigenvektoren* mit gängiger Software.

MATLAB/Octave	<code>[E,D]=eig(A)</code>
Mathematica/WolframAlpha	<code>Eigenvalues[A]</code> <code>Eigenvectors[A]</code> <code>Eigensystem[A]</code>
Python/Numpy	<code>import numpy as np;</code> <code>[S,E]=np.linalg.eig(A)</code>
Python/Sympy	<code>import sympy as sp;</code> <code>S=A.eigenvals()</code> <code>E=A.eigenvects()</code> <code>[E,D]=A.diagonalize()</code>

Um die *Eigenwerte* einer *Matrix* $A \in \mathbb{M}(n, n, \mathbb{R})$ zu berechnen, können wir also nach den folgenden Schritten vorgehen.

S1 Berechnen des *charakteristischen Polynoms*

$$p_A(\lambda) = \det(\lambda \cdot \mathbb{1} - A). \quad (6.174)$$

S2 Bestimmen der *Eigenwerte*, d.h. der *Nullstellen* von $p_A(\lambda)$.

S3 Für jeden *Eigenwert* $\lambda \in \text{Spec}(A)$ die *Lösungen* des *linearen Gleichungssystems*

$$(\lambda \cdot \mathbb{1} - A) \cdot \mathbf{E} = 0. \quad (6.175)$$

berechnen.

6.5.5 Spezielle Eigenwerte & Eigenvektoren

Für einige *Abbildungsmatrizen* kann man die *Eigenwerte* und *Eigenvektoren* direkt aus ihren *geometrischen* Eigenschaften ablesen. Wir machen dazu folgende Bemerkungen.

- i)** Es gilt $\text{Spec}(0) = \{0\}$ und jedes $\mathbf{E} \in \mathbb{R}^n$ ist *Eigenvektor* der *Nullmatrix*.
- ii)** Es gilt $\text{Spec}(\mathbb{1}) = \{1\}$ und jedes $\mathbf{E} \in \mathbb{R}^n$ ist *Eigenvektor* der *Einheitsmatrix*.
- iii)** A singulär $\Leftrightarrow 0 \in \text{Spec}(A)$
- iv)** $A \in O(n) \Rightarrow \text{Spec}(A) \subseteq \{-1, 1\}$
- v)** A eine *Spiegelung* $\Rightarrow -1 \in \text{Spec}(A)$
- vi)** Für eine *Diagonalmatrix* gilt:

$$D = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix} \Rightarrow \text{Spec}(D) \subseteq \{\lambda_1, \dots, \lambda_n\}. \quad (6.176)$$

Die *Einheitseigenvektoren* zu jedem *Eigenwert* sind gerade die *Einheitsvektoren* entlang der Achsen, d.h.

$$\hat{\mathbf{E}}_k = \hat{\mathbf{e}}_k. \quad (6.177)$$

Kapitel 7

Vektorräume

7.1 Vektorraumstruktur

7.1.1 Definition

Die fundamentale Struktur der *linearen Algebra* ist der *Vektorraum*. Dieser wird, wie in der modernen Mathematik allgemein üblich, mit Hilfe einer überschaubaren Aufzählung von grundlegenden Eigenschaften, den sogenannten *Axiomen*, definiert.

Definition 7.1 Vektorraum

Ein *Vektorraum* ist ein *Quadrupel* $(V, \mathbb{K}, +, \cdot)$, bestehend aus einer *Menge* V , einem *Zahlenkörper* \mathbb{K} und zwei Operationen

$$\begin{array}{ccc} + : V \times V & \rightarrow V & \cdot : \mathbb{K} \times V & \rightarrow V \\ (\mathbf{v}; \mathbf{w}) & \mapsto \mathbf{v} + \mathbf{w} & \text{und} & (a; \mathbf{v}) & \mapsto a \cdot \mathbf{v}, \end{array} \quad (7.1)$$

so dass für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $a, b \in \mathbb{K}$ die folgenden *Axiome* gelten.

VR-1 $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$

VR-2 Es gibt ein $\mathbf{0} \in V$ mit $\mathbf{0} + \mathbf{v} = \mathbf{v}$ für alle $\mathbf{v} \in V$.

VR-3 Für jedes $\mathbf{v} \in V$ gibt es ein $-\mathbf{v} \in V$ mit $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

VR-4 $\mathbf{w} + \mathbf{v} = \mathbf{v} + \mathbf{w}$

VR-5 $a \cdot (\mathbf{v} + \mathbf{w}) = a \cdot \mathbf{v} + a \cdot \mathbf{w}$

VR-6 $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$

VR-7 $(a \cdot b) \cdot \mathbf{v} = a \cdot (b \cdot \mathbf{v})$

VR-8 $1 \cdot \mathbf{v} = \mathbf{v}$

Entwickelt man eine mathematische Theorie ausgehend von *Axiomen*, dann müssen auch Aussagen, deren Gültigkeit in praktischen Anwendungen “selbstverständlich” ist, sorgfältig aus diesen *Axiomen* bewiesen werden. Ein schönes Beispiel ist die sogenannte *Null-Koinzidenz*, welche in allen *Vektorräumen* gilt.

Satz 7.1 Null-Koinzidenz

In jedem Vektorraum $(V, \mathbb{K}, +, \cdot)$ gilt für alle Vektoren $\mathbf{v} \in V$

$$0 \cdot \mathbf{v} = \mathbf{0}. \quad (7.2)$$

Beweis: Gemäss VR-8, VR-6 und wieder VR-8 und dann VR-1 sowie VR-3 gilt

$$0 \cdot \mathbf{v} + \mathbf{v} = 0 \cdot \mathbf{v} + 1 \cdot \mathbf{v} = (0 + 1) \cdot \mathbf{v} = 1 \cdot \mathbf{v} = \mathbf{v} \quad | + (-\mathbf{v}) \quad (7.3)$$

$$\Leftrightarrow (0 \cdot \mathbf{v} + \mathbf{v}) + (-\mathbf{v}) = \mathbf{v} + (-\mathbf{v}) \quad (7.4)$$

$$\Leftrightarrow 0 \cdot \mathbf{v} + (\mathbf{v} + (-\mathbf{v})) = \mathbf{0} \quad (7.5)$$

$$\Leftrightarrow 0 \cdot \mathbf{v} + \mathbf{0} = \mathbf{0}. \quad (7.6)$$

Aus VR-2 folgt schliesslich

$$\underline{\underline{0 \cdot \mathbf{v} = \mathbf{0}}}. \quad (7.7)$$

Damit haben wir (7.2) und den Satz bewiesen. \square

Bemerkungen:

- i) Die *Elemente* des *Zahlenkörpers* \mathbb{K} werden *Skalare* und die *Elemente* der *Menge* V werden *Vektoren* genannt.
- ii) Die *Vektoren* in einem *Vektorraum* erfüllen oft weitere *Strukturen*, sie können z.B. *Punkte* in der Geometrie oder *Funktionen* in der Analysis sein. Insbesondere bilden alle möglichen physikalischen Größen einen *Vektorraum*.
- iii) Auch für die Struktur des *Zahlenkörpers* \mathbb{K} gibt es ein System von *Axiomen*, die wir hier einfach stillschweigend anwenden. In diesem Kurs gilt fast immer $\mathbb{K} = \mathbb{R}$. In diesem Fall spricht man auch von einem *reellen Vektorraum*. Denkbar wären aber auch $\mathbb{K} \in \{\mathbb{Q}, \mathbb{C}\}$.
- iv) Wegen der *Null-Koinzidenz* ist es sinnvoll, $0 \in \mathbb{K}$ und $\mathbf{0} \in V$ zu identifizieren. Wir schreiben von nun an einfach nur noch 0 .
- v) Wenn in Anwendungen “klar” ist, welcher *Zahlenkörper* \mathbb{K} und welche Operationen $+$ bzw. \cdot gemeint ist, dann wird der *Vektorraum* meist nur noch durch die *Menge* V der *Vektoren* bezeichnet und nicht mehr das ganze *Quadrupel* $(V, \mathbb{K}, +, \cdot)$ geschrieben.
- vi) Die 8 *Axiome* des *Vektorraums* zusammen mit ersten Folgerungen wie der *Null-Koinzidenz* sind das, was man gemeinhin unter den “üblichen Rechenregeln” für *Vektoren* versteht.
- vii) Aufgrund der *Axiome* lassen sich in jedem *Vektorraum* beliebige *Linearkombinationen* von *Vektoren* bilden. Es gelten dabei die “üblichen Rechenregeln” und insbesondere ist jede *Linearkombination* von *Vektoren* wieder ein *Vektor*, d.h. *Vektorräume* sind unter der Bildung von *Linearkombinationen* abgeschlossen.

Beispiele:

- Der einfachste *Vektorraum* ist der *triviale Vektorraum*, der aus einem beliebigen *Zahlenkörper* aber nur einem einzigen *Vektor* besteht, nämlich 0 , d.h. $(\{0\}, \mathbb{K}, +, \cdot)$.
- Für $n \in \mathbb{N}^+$ die bekannten *EUKLID-Räume* $(\mathbb{R}^n, \mathbb{R}, +, \cdot)$.

- Für $n \in \mathbb{N}^+$ die *rationalen* und *komplexen EUKLID-Räume* $(\mathbb{Q}^n, \mathbb{Q}, +, \cdot)$ bzw. $(\mathbb{C}^n, \mathbb{C}, +, \cdot)$.
- Für $n, m \in \mathbb{N}^+$ die *Matrix-Räume* $(\mathbb{M}(m, n, \mathbb{K}), \mathbb{K}, +, \cdot)$ aus *Matrizen* mit *Komponenten* aus dem *Zahlenkörper* \mathbb{K} .
- In der Geometrie lässt sich jede *Gerade*, *Ebene*, *Hyperebene*, etc.. als *Vektorraum* beschreiben.
- Für $n \in \mathbb{N}^+$ der *Funktionenraum* der *Polynome* mit *reellen Koeffizienten* vom *Grad* kleiner oder gleich n , d.h. $(\mathcal{P}_n(\mathbb{R}), \mathbb{R}, +, \cdot)$ mit

$$\mathcal{P}_n(\mathbb{R}) := \{ f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ ist ein Polynom vom Grad } p \leq n \}. \quad (7.8)$$

- Der *Funktionenraum* der *stetigen, reellwertigen Funktionen*, d.h. $(C(\mathbb{R}), \mathbb{R}, +, \cdot)$ mit

$$C(\mathbb{R}) := \{ f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ ist stetig} \}. \quad (7.9)$$

- Für $p \in \mathbb{N}^+$ die *LEBESGUE-Funktionenräume*, d.h. $(\mathcal{L}^p(\mathbb{R}), \mathbb{R}, +, \cdot)$ mit

$$\mathcal{L}^p(\mathbb{R}) := \left\{ f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ ist integrierbar} \wedge \int_{-\infty}^{\infty} |f(x)|^p dx < \infty \right\}. \quad (7.10)$$

In der Physik spielt der *LEBESGUE-Funktionenraum* $\mathcal{L}^2(\mathbb{R})$ eine sehr wichtige Rolle.

7.1.2 Linearkombinationen, Basis & Dimension

Die Operationen $+$ und \cdot und die gemäss den *Vektorraum-Axiomen* dafür geforderten Rechenregeln garantieren, dass sich in jedem *Vektorraum* auf sinnvolle Weise *Linearkombinationen* bilden lassen.

Definition 7.2 Linearkombination

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum*, $m \in \mathbb{N}^+$, $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \subseteq V$ und $\{x_1, \dots, x_m\} \subseteq \mathbb{K}$. Eine *Linearkombination* der *Vektoren* $\mathbf{v}_1, \dots, \mathbf{v}_m$ ist eine *Formel* der Form

$$\mathbf{w} = \sum_{k=1}^m x_k \cdot \mathbf{v}_k. \quad (7.11)$$

Sind *Linearkombinationen* erst einmal definiert, dann stellt sich die Frage, welche *Teilmenge* des *Vektorraums* durch *Linearkombinieren* von einigen *Vektoren* mit beliebigen *Koeffizienten* aus dem *Zahlenkörper* erzeugt wird.

Definition 7.3 Lineare Hülle

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* und $m \in \mathbb{N}^+$. Die *lineare Hülle* von $\{\mathbf{v}_1, \dots, \mathbf{v}_m\} \subseteq V$ ist

$$\text{span}(\{\mathbf{v}_1, \dots, \mathbf{v}_m\}) := \left\{ \sum_{k=1}^m x_k \cdot \mathbf{v}_k \mid x_1, \dots, x_m \in \mathbb{K} \right\}. \quad (7.12)$$

Bemerkungen:

- i) In der Literatur wird die *lineare Hülle* auch *Spann* oder *Spannweite* genannt.

- ii)** Das Konzept der *linearen Hülle* lässt sich auf eine beliebige *Teilmenge* $A \subseteq V$ eines *Vektorraums* ausweiten. In jedem Fall gilt

$$\text{span}(A) \equiv \text{Alle mögliche Linearkombinationen von Vektoren in } A. \quad (7.13)$$

Dabei ist zu beachten, dass auch dann wenn A unendlich viele Vektoren enthält in jeder Linearkombination nur endlich viele davon auftreten dürfen.

Definition 7.4 linear unabhängig, erzeugend und Basis

Seien $(V, \mathbb{K}, +, \cdot)$ ein Vektorraum, $n \in \mathbb{N}^+$.

- (a)** Die Vektoren in $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ heissen *linear unabhängig*, falls

$$0 = \sum_{k=1}^n x_k \cdot \mathbf{v}_k \Leftrightarrow 0 = x_1 = \dots = x_n. \quad (7.14)$$

- (b)** Die Vektoren in $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ heissen *erzeugend*, falls

$$\text{span}(\{\mathbf{v}_1, \dots, \mathbf{v}_n\}) = V. \quad (7.15)$$

- (c)** Die Vektoren in $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq V$ bilden eine *Basis* von V , falls sie *linear unabhängig* und *erzeugend* sind.

Bemerkungen:

- i)** Die Vektoren in $\{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subseteq V$ heissen *linear abhängig*, genau dann wenn sie nicht *linear unabhängig* sind.
- ii)** In einer Menge von *linear unabhängigen* Vektoren lässt sich keiner dieser Vektoren als Linearkombination der andern darstellen. Jeder Vektor trägt so etwas wie eine “neue Richtung” bei, welche durch eine Linearkombination der andern nicht “beschriften” werden kann.
- iii)** Ist eine Menge von Vektoren *erzeugend*, dann lässt sich jeder Vektor im Vektorraum als Linearkombination von Vektoren aus dieser Menge darstellen.

Bei einer *Basis* kommen die Eigenschaften *linear unabhängig* und *erzeugend* zusammen. Dies führt auf folgendes Ergebnis, das für die Praxis äusserst wertvoll ist.

Satz 7.2 Eindeutigkeit der *Basis-Darstellung*

Sei $(V, \mathbb{K}, +, \cdot)$ ein Vektorraum, $n \in \mathbb{N}^+$, $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq V$ eine *Basis* und $\mathbf{v} \in V$. Dann gibt es eindeutige Koeffizienten $v^1, \dots, v^n \in \mathbb{K}$, so dass

$$\mathbf{v} = \sum_{k=1}^n v^k \cdot \mathbf{e}_k. \quad (7.16)$$

Beweis: Weil die Vektoren in B eine *Basis* des Vektorraums bilden, sind sie nach Definition 7.4 *erzeugend*. Somit gibt es Koeffizienten $v^1, \dots, v^n \in \mathbb{K}$, so dass

$$\mathbf{v} = \sum_{k=1}^n v^k \cdot \mathbf{e}_k. \quad (7.17)$$

Betrachten wir einen zweiten Satz *Koeffizienten* $\tilde{v}^1, \dots, \tilde{v}^n \in \mathbb{K}$, so dass ebenfalls gilt

$$\mathbf{v} = \sum_{k=1}^n \tilde{v}^k \cdot \mathbf{e}_k, \quad (7.18)$$

dann erhalten wir

$$0 = \mathbf{v} - \mathbf{v} = \sum_{k=1}^n \tilde{v}^k \cdot \mathbf{e}_k - \sum_{k=1}^n v^k \cdot \mathbf{e}_k = \sum_{k=1}^n (\tilde{v}^k \cdot \mathbf{e}_k - v^k \cdot \mathbf{e}_k) = \sum_{k=1}^n (\tilde{v}^k - v^k) \cdot \mathbf{e}_k. \quad (7.19)$$

Weil die *Vektoren* in B gemäss Definition 7.4 auch *linear unabhängig* sind, folgt für jedes $k \in \{1, \dots, n\}$, dass

$$\tilde{v}^k - v^k = 0 \quad \left| + v^k \right. \quad (7.20)$$

$$\Leftrightarrow \quad \tilde{v}^k = v^k. \quad (7.21)$$

Wir haben also sowohl die *Existenz* als auch die *Eindeutigkeit* der *Koeffizienten* $v^1, \dots, v^n \in \mathbb{K}$ gezeigt und damit den Satz bewiesen. \square

Nicht jeder *Vektorraum* kann durch *endlich viele Vektoren erzeugt* werden. Betrachtet man hingegen auch *Mengen* aus *unendlich vielen Vektoren*, dann kann man in jedem *nichttrivialen Vektorraum* eine *Basis* finden.

Satz 7.3 Existenz einer Basis

In jedem *nichttrivialen Vektorraum* existiert eine *Basis* entweder aus *endlich* oder *unendlich vielen Vektoren*.

Von besonderem Interesse ist der folgende Satz über die Anzahl *Vektoren* in einer *Basis*.

Satz 7.4 Eindeutigkeit der Anzahl *Basis-Vektoren*

Jede *Basis* eines *Vektorraums* besteht aus der gleichen Anzahl *Basis-Vektoren*.

Die Tatsache, dass jede *Basis* eines *Vektorraums* aus gleich vielen *Vektoren* besteht, macht diese Anzahl zu einer wichtigen Kenngrösse, die einen eigenen Namen verdient.

Definition 7.5 Dimension

Die Anzahl *Basis-Vektoren* in jeder *Basis* eines *Vektorraums* heisst *Dimension* des *Vektorraums*.

Bemerkungen:

- i) Insbesondere *Funktionenräume* können oft nicht *endlich erzeugt* werden und haben somit *unendliche Dimension*.
- ii) Bemerkenswert ist die Tatsache, dass der Begriff *Dimension* hier rein *algebraisch* definiert wird. Es wird dafür keinerlei *Geometrie* oder gar *geometrische Anschauung* benötigt.
- iii) Ist $n \in \mathbb{N}^+$ die *Dimension* eines *Vektorraums* $(V, \mathbb{K}, +, \cdot)$, dann schreibt man formell

$$\dim(V) = n. \quad (7.22)$$

Beispiele:

- Für jedes $n \in \mathbb{N}^+$ bilden die *Standard-Einheitsvektoren*

$$\left\{ \hat{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \hat{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \hat{\mathbf{e}}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\} \quad (7.23)$$

eine *Basis* von \mathbb{K}^n , die sogenannte *Standard-Basis*. Somit folgt

$$\dim(\mathbb{K}^n) = n. \quad (7.24)$$

- Für jedes $n \in \mathbb{N}^+$ bilden die *Monome*

$$\{1, x, x^2, x^3, \dots, x^n\} \quad (7.25)$$

eine *Basis* des *Polynomraums* $\mathcal{P}_n(\mathbb{R})$. Somit folgt

$$\dim(\mathcal{P}_n(\mathbb{R})) = n + 1. \quad (7.26)$$

- Die *Funktionenräume* $C(\mathbb{R})$ und $\mathcal{L}^2(\mathbb{R})$ können nicht *endlich erzeugt* werden und haben daher *unendliche Dimension*.

Ist in einem *Vektorraum* $(V, \mathbb{K}, +, \cdot)$ eine *Basis* $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ gewählt, dann lässt sich jeder *Vektor* eindeutig *Linearkombination* des *Basis-Vektoren* schreiben gemäss

$$\mathbf{v} = \sum_{k=1}^n v^k \cdot \mathbf{e}_k. \quad (7.27)$$

Die *Koeffizienten* $v^1, \dots, v^n \in \mathbb{K}$ können wie im Falle der *EUKLID-Räume* \mathbb{R}^n in eine *Spalten-Matrix* geschrieben werden, d.h. man identifiziert

$$\mathbf{v} = \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}. \quad (7.28)$$

Die Operationen $+$ und \cdot im *Vektorraum* lassen sich dann wieder durch die gewohnten Operationen in den *Komponenten* realisieren.

Beispiel:

- Wir betrachten den *Vektorraum* $\mathcal{P}_2(\mathbb{R})$ der *quadratischen Funktionen* mit *reellen Koeffizienten* und wählen die *Basis* $\{1, x, x^2\}$. Dann identifizieren wir

$$f(x) = 3x^2 + 5x + 7 \mapsto \mathbf{f} = \begin{bmatrix} 7 \\ 5 \\ 3 \end{bmatrix}. \quad (7.29)$$

7.1.3 Unterräume

Eine *Teilmenge* eines *Vektorraums* kann selbst wieder ein *Vektorraum* sein. Dazu machen wir die folgende Definition.

Definition 7.6 Unterraum

Sei $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum*. Eine *Teilmenge* $W \subseteq V$ heisst *Unterraum* von V , falls $(W, \mathbb{K}, +, \cdot)$ ebenfalls ein *Vektorraum* ist.

Bemerkungen:

- i) In der Literatur werden *Unterräume* auch *Teilräume* genannt.
- ii) Um kompakt auszudrücken, dass $W \subseteq V$ nicht nur eine *Teilmenge* sondern ein *Unterraum* von V ist, verwendet man die Schreibweise

$$W \leq V. \quad (7.30)$$

- iii) Jeder *Unterraum* eines *Vektorraums* V muss mindestens 0 enthalten.
- iv) Jeder *Vektorraum* hat zumindest sich selbst und den *trivialen Vektorraum* als *Unterraum*. Für jeden *Vektorraum* V gilt also

$$\{0\} \leq V \quad \text{und} \quad V \leq V. \quad (7.31)$$

Mit Hilfe des folgenden Satzes kann sehr einfach getestet werden, ob eine *Teilmenge* eines *Vektorraums* ein *Unterraum* ist.

Satz 7.5 Test auf *Unterraum*

Sei $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* und $W \subseteq V$. Dann gilt

$$W \leq V \Leftrightarrow \text{span}(W) = W. \quad (7.32)$$

Bemerkungen:

- i) Eine *Teilmenge* eines *Vektorraums* ist also genau dann ein *Unterraum*, wenn sie abgeschlossen ist unter der Bildung von *Linearkombinationen*.
- ii) Um nachzuweisen, das $W \leq V$ muss man also zeigen, dass

$$\mathbf{w}_1, \mathbf{w}_2 \in W, x_1, x_2 \in \mathbb{K} \Rightarrow x_1 \cdot \mathbf{w}_1 + x_2 \cdot \mathbf{w}_2 \in W. \quad (7.33)$$

Beispiele:

- Für jedes $m, n \in \mathbb{N}^+$ mit $m \leq n$ gilt $\mathbb{K}^m \leq \mathbb{K}^n$.
- Für jedes $m, n \in \mathbb{N}^+$ mit $m \leq n$ gilt $\mathcal{P}_m(\mathbb{R}) \leq \mathcal{P}_n(\mathbb{R})$.
- Jede *Gerade* in der *Ebene* kann als *Unterraum* der *Ebene* aufgefasst werden.
- Jede *Gerade* oder *Ebene* im *Raum* kann als *Unterraum* des *Raumes* aufgefasst werden.

7.2 Lineare Abbildungen

7.2.1 Definition

Der Begriff *lineare Abbildung* lässt sich zwischen allgemeinen *Vektorräumen* definieren.

Definition 7.7 Lineare Abbildung

Seien $(V, \mathbb{K}, +, \cdot)$ und $(W, \mathbb{K}, +, \cdot)$ zwei *Vektorräume* über dem gleichen *Zahlenkörper* \mathbb{K} . Eine *Abbildung* der Form

$$a : V \rightarrow W \quad (7.34)$$

heisst *lineare Abbildung*, falls für alle $\mathbf{u}, \mathbf{v} \in V$ und $x, y \in \mathbb{K}$ gilt

$$a(x \cdot \mathbf{u} + y \cdot \mathbf{v}) = x \cdot a(\mathbf{u}) + y \cdot a(\mathbf{v}). \quad (7.35)$$

Bemerkungen:

- i)** Für diese Definition ist es sehr wichtig, dass die *Vektorräume* V und W über dem gleichen *Zahlenkörper* \mathbb{K} definiert sind.
- ii)** Diese Definition entspricht der Definition 6.14 aus Abschnitt 6.2.1.
- iii)** Eine *lineare Abbildung* erkennt man daran, dass sie die Struktur einer *Linearkombination* respektiert.
- iv)** Für alle *linearen Abbildungen* gilt offensichtlich

$$\underline{\underline{a(0)}} = a(0 \cdot 0) = 0 \cdot a(0) = \underline{\underline{0}}. \quad (7.36)$$

Beispiele:

- Die bereits bekannten *geometrisch* definierten *linearen Abbildungen* in \mathbb{R}^n wie *Spiegelungen*, *Drehungen*, *Projektionen* etc...
- Die *Ableitung* $d : \mathcal{P}_n(\mathbb{R}) \rightarrow \mathcal{P}_n(\mathbb{R})$ oder $d : \mathcal{P}_n(\mathbb{R}) \rightarrow \mathcal{P}_{n-1}(\mathbb{R})$.
- Die *Orthogonal-Projektion* in $\mathcal{L}^2(\mathbb{R})$.

7.2.2 Matrix-Darstellung

Betrachtet man eine *lineare Abbildung* zwischen zwei *endlich dimensionalen Vektorräumen* und wählt in beiden jeweils eine *Basis*, dann kann man die *lineare Abbildung* durch eine *Abbildungs-matrix* darstellen.

Definition 7.8 Abbildungsmatrix

Seien $(V, \mathbb{K}, +, \cdot)$ und $(W, \mathbb{K}, +, \cdot)$ zwei *Vektorräume* über dem gleichen *Zahlenkörper* \mathbb{K} mit den *endlichen Dimensionen* $\dim(V) = n \in \mathbb{N}^+$ bzw. $\dim(W) = m \in \mathbb{N}^+$ und *Basen* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq V$ bzw. $\{\mathbf{E}_1, \dots, \mathbf{E}_m\} \subseteq W$ sowie $a : V \rightarrow W$ eine *lineare Abbildung*. Die *Abbildungsmatrix* von a bezüglich der gewählten *Basen* ist die *Matrix* $A \in \mathbb{M}(m, n, \mathbb{K})$ mit den *Komponenten* $A^i_j \in \mathbb{K}$, so dass für alle $j \in \{1, \dots, n\}$ gilt

$$a(\mathbf{e}_j) = \sum_{i=1}^m A^i_j \cdot \mathbf{E}_i. \quad (7.37)$$

Bemerkenswerterweise gilt nun der folgende Satz.

Satz 7.6 Berechnung mit der *Abbildungsmatrix*

Seien $(V, \mathbb{K}, +, \cdot)$ und $(W, \mathbb{K}, +, \cdot)$ zwei *Vektorräume* über dem gleichen *Zahlenkörper* \mathbb{K} mit den *endlichen Dimensionen* $\dim(V) = n \in \mathbb{N}^+$ bzw. $\dim(W) = m \in \mathbb{N}^+$ und *Basen* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subseteq V$ bzw. $\{\mathbf{E}_1, \dots, \mathbf{E}_m\} \subseteq W$ sowie $a : V \rightarrow W$ eine *lineare Abbildung* mit *Abbildungsmatrix* $A \in \mathbb{M}(m, n, \mathbb{K})$. Ferner seien

$$\mathbf{v} = \sum_{j=1}^n v^j \cdot \mathbf{e}_j \in V \quad \text{und} \quad \mathbf{w} = \sum_{i=1}^m w^i \cdot \mathbf{E}_i \in W, \quad (7.38)$$

für welche gilt

$$\mathbf{w} = a(\mathbf{v}). \quad (7.39)$$

Für die *Komponenten* von \mathbf{v} und \mathbf{w} gilt dann die Beziehung

$$\boxed{\begin{bmatrix} w^1 \\ \vdots \\ w^m \end{bmatrix} = \begin{bmatrix} A^1_1 & \dots & A^1_n \\ \vdots & \vdots & \vdots \\ A^m_1 & \dots & A^m_n \end{bmatrix} \cdot \begin{bmatrix} v^1 \\ \vdots \\ v^n \end{bmatrix}}. \quad (7.40)$$

Beweis: Durch Einsetzen der *Basis-Darstellung* von \mathbf{v} und weil a eine *lineare Abbildung* ist, erhalten wir

$$\begin{aligned} \sum_{i=1}^m w^i \cdot \mathbf{E}_i &= \mathbf{w} = a(\mathbf{v}) = a\left(\sum_{j=1}^n v^j \cdot \mathbf{e}_j\right) = \sum_{j=1}^n v^j \cdot a(\mathbf{e}_j) = \sum_{j=1}^n v^j \cdot \sum_{i=1}^m A^i_j \cdot \mathbf{E}_i \\ &= \sum_{i=1}^m \underbrace{\sum_{j=1}^n A^i_j \cdot v^j}_{= w^i} \cdot \mathbf{E}_i. \end{aligned} \quad (7.41)$$

Wegen der Eindeutigkeit der *Basis-Darstellung* folgt aus einem *Koeffizientenvergleich* für alle $i \in \{1, \dots, m\}$, dass

$$w^i = \sum_{j=1}^n A^i_j \cdot v^j. \quad (7.42)$$

In *Matrix-Schreibweise* entspricht dies genau (7.40) und wir haben den Satz bewiesen. \square

Bemerkungen:

- i)** Sind in den *Vektorräumen* V und W jeweils *Basen* gewählt, dann reduziert sich die Anwendung einer *linearen Abbildung* auf eine *Matrix-Multiplikation* analog zur Situation für eine *lineare Abbildung* des Typs $a : \mathbb{R}^n \rightarrow \mathbb{R}^m$.
- ii)** Die Eigenschaften einer *linearen Abbildung* können aus den Eigenschaften ihrer *Abbildungsmatrix* abgelesen werden.
- iii)** Analog zur Situation für eine *lineare Abbildung* des Typs $a : \mathbb{R}^n \rightarrow \mathbb{R}^m$ gelten auch hier der *Verknüpfungssatz* und der *Inversionssatz* für *bijektive, lineare Abbildungen*
- iv)** Analog zur Situation für eine *lineare Abbildung* des Typs $a : \mathbb{R}^n \rightarrow \mathbb{R}^m$ kann die *Abbildungsmatrix* mit Hilfe des *Spalten-Vektor-Konstruktionsverfahrens* gefunden werden.
- v)** Wählt man in V bzw. W eine andere *Basis*, dann wird die gleiche *lineare Abbildung* durch eine andere *Abbildungsmatrix* beschrieben.

7.2.3 Bild & Kern

Für eine *lineare Abbildung* lassen sich zwei charakteristische *Mengen* definieren.

Definition 7.9 Bild & Kern

Seien $(V, \mathbb{K}, +, \cdot)$ und $(W, \mathbb{K}, +, \cdot)$ zwei *Vektorräume* über dem gleichen *Zahlenkörper* \mathbb{K} und $a : V \rightarrow W$ eine *lineare Abbildung*.

(a) Das *Bild* von a ist die *Menge*

$$\text{img}(a) := a(V) := \{\mathbf{w} \in W \mid \text{Es gibt ein } \mathbf{v} \in V \text{ mit } a(\mathbf{v}) = \mathbf{w}\}. \quad (7.43)$$

(b) Der *Kern* von a ist die *Menge*

$$\ker(a) := \{\mathbf{v} \in V \mid a(\mathbf{v}) = 0\}. \quad (7.44)$$

Bemerkungen:

- i)** Der Begriff des *Bildes* in der *linearen Algebra* stimmt überein mit dem entsprechenden Begriff aus der allgemeinen Theorie der *Funktionen*.
- ii)** Der *Kern* einer *linearen Abbildung* besteht gerade aus jenen *Vektoren*, die auf 0 abgebildet werden.
- iii)** Um Verwechslungen vorzubeugen, sei nochmals betont, dass gilt

$$\text{img}(a) \subseteq W \quad \text{aber} \quad \ker(a) \subseteq V. \quad (7.45)$$

- iv)** In jedem Fall gilt $0 \in \ker(a)$.
- v)** Gilt $\ker(a) = \{0\}$, dann sagt man, a hat einen *trivialen Kern*.

Charakteristisch für *lineare Abbildungen* ist der folgende Satz.

Satz 7.7 Dimensionssatz

Seien $(V, \mathbb{K}, +, \cdot)$ und $(W, \mathbb{K}, +, \cdot)$ zwei *Vektorräume* über dem gleichen *Zahlenkörper* \mathbb{K} und $a : V \rightarrow W$ eine *lineare Abbildung*. Dann gilt $\text{img}(a) \leq W$ und $\ker(a) \leq V$ sowie

$$\dim(\text{img}(a)) + \dim(\ker(a)) = \dim(V). \quad (7.46)$$

Bemerkungen:

- i)** *Bild* und *Kern* sind demnach nicht nur *Teilmengen* sondern *Unterräume* von W bzw. V .
- ii)** Der *Dimensionssatz* ist eine Art ‘Erhaltungssatz’ für *Dimensionen* unter der Wirkung einer *linearen Abbildung*. Die *Dimension* von V wird aufgeteilt auf *Bild* und *Kern*.

Satz 7.8 Umkehrbarkeitssatz

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* und $a : V \rightarrow V$ eine *lineare Abbildung*. Dann gilt

$$a \text{ bijektiv} \Leftrightarrow \ker(a) = \{0\}. \quad (7.47)$$

7.3 Skalar-Produkt & Metrik

7.3.1 Skalar-Produkt

7.3.1.1 Definition

Wir betrachten die folgende Definition.

Definition 7.10 Skalar-Produkt

Sei $(V, \mathbb{K}, +, \cdot)$ ein Vektorraum über dem Zahlenkörper \mathbb{K} . Ein Skalar-Produkt auf V ist eine Operation der Form

$$\begin{aligned} V \times V &\rightarrow \mathbb{K} \\ (\mathbf{v}; \mathbf{w}) &\mapsto \langle \mathbf{v}, \mathbf{w} \rangle, \end{aligned} \tag{7.48}$$

so dass für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $a, b \in \mathbb{K}$ die folgenden *Axiome* gelten.

SP-1 Linearität im 2. Argument:

$$\langle \mathbf{u}, a \cdot \mathbf{v} + b \cdot \mathbf{w} \rangle = a \cdot \langle \mathbf{u}, \mathbf{v} \rangle + b \cdot \langle \mathbf{u}, \mathbf{w} \rangle \tag{7.49}$$

SP-2 Symmetrie:

$$\langle \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle^* \tag{7.50}$$

SP-3 Nicht-Degeneriertheit:

$$\langle \mathbf{v}, \mathbf{p} \rangle = 0 \text{ für alle } \mathbf{p} \in V \Leftrightarrow \mathbf{v} = 0 \tag{7.51}$$

Wir betrachten die folgende Definition.

Definition 7.11 Positive Definitheit

Sei $(V, \mathbb{K}, +, \cdot)$ ein Vektorraum über dem Zahlenkörper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, dann heisst ein Skalar-Produkt $\langle \cdot, \cdot \rangle$ auf V positiv definit, falls für alle $\mathbf{v} \in V$ gilt

$$\langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \quad \text{und} \quad \langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = 0. \tag{7.52}$$

Bemerkungen:

- i) In der Literatur sind die Begriffe Skalar-Produkt und inneres Produkt synonym.
- ii) Aus der positiven Definitheit folgt sofort SP-3. Deshalb wird bei positiv definiten Skalar-Produkten das Axiom SP-3 durch die Eigenschaft der positiven Definitheit ersetzt.
- iii) Je nach Wahl von $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ vereinfacht sich SP-2. Es gilt

$$\mathbb{K} = \mathbb{R} \Rightarrow \langle \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \tag{7.53}$$

$$\mathbb{K} = \mathbb{C} \Rightarrow \langle \mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle^*. \tag{7.54}$$

- iv)** Für *reelle Vektorräume*, d.h. für $\mathbb{K} = \mathbb{R}$ folgt aus der Kombination von SP-1 und SP-2 die *Linearität im 1. Argument* und damit die *Bilinearität des Skalar-Produkts*. Für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $a, b \in \mathbb{R}$ gilt

$$\langle a \cdot \mathbf{u} + b \cdot \mathbf{v}, \mathbf{w} \rangle = a \cdot \langle \mathbf{u}, \mathbf{w} \rangle + b \cdot \langle \mathbf{v}, \mathbf{w} \rangle. \quad (7.55)$$

- v)** Für *komplexe Vektorräume*, d.h. für $\mathbb{K} = \mathbb{C}$ folgt aus der Kombination von SP-1 und SP-2 die *Semilinearität im 1. Argument* und damit die *Sesquilinearität des Skalar-Produkts*. Für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $a, b \in \mathbb{C}$ gilt

$$\langle a \cdot \mathbf{u} + b \cdot \mathbf{v}, \mathbf{w} \rangle = a^* \cdot \langle \mathbf{u}, \mathbf{w} \rangle + b^* \cdot \langle \mathbf{v}, \mathbf{w} \rangle. \quad (7.56)$$

Beispiele:

- GRAM-RIEMANN-Skalar-Produkt auf \mathbb{K}^n (*positiv definit*):

$$\text{auf } V = \mathbb{R}^n: \quad \langle \mathbf{v}, \mathbf{w} \rangle := v_1 \cdot w_1 + v_2 \cdot w_2 + \dots + v_n \cdot w_n \quad (7.57)$$

$$\text{auf } V = \mathbb{C}^n: \quad \langle \mathbf{v}, \mathbf{w} \rangle := v_1^* \cdot w_1 + v_2^* \cdot w_2 + \dots + v_n^* \cdot w_n. \quad (7.58)$$

Anwendungen: *Geometrie, Datenanalyse*

- LORENTZ-MINKOWSKI-Skalar-Produkt auf \mathbb{R}^{1+3} (*nicht positiv definit*):

$$\langle \mathbf{v}, \mathbf{w} \rangle := v^0 \cdot w^0 - v^1 \cdot w^1 - v^2 \cdot w^2 - v^3 \cdot w^3 \quad (7.59)$$

Anwendungen: *Relativitätstheorie*

- SCHUR-Skalar-Produkt auf $\mathbb{M}(n, n, \mathbb{R})$ (*positiv definit*):

$$\langle A, B \rangle := \text{tr}(A^T \cdot B). \quad (7.60)$$

Anwendungen: *Gruppen-Theorie, Datenanalyse*

- Wir betrachten den *Funktionenraum der komplexen, integrierbaren, periodischen Funktionen* auf \mathbb{R} mit *Periode* $T > 0$ gemäss

$$V = \{ f : \mathbb{R} \rightarrow \mathbb{C} \mid f \text{ ist integrierbar} \wedge f(t+T) = f(t) \text{ für alle } t \in \mathbb{R} \}. \quad (7.61)$$

L²-Skalar-Produkt auf V (*positiv definit*):

$$(f, g) := \frac{1}{T} \int_T f^*(t) \cdot g(t) dt. \quad (7.62)$$

Anwendungen: *FOURIER-Entwicklungen, Signalverarbeitung*

- *L²-Skalar-Produkt* auf den LEBESGUE-Funktionenräumen $\mathcal{L}^2(\mathbb{R}, \mathbb{K})$ (*positiv definit*):

$$\text{auf } V = \mathcal{L}^2(\mathbb{R}, \mathbb{R}): \quad (f, g) := \int_{-\infty}^{\infty} f(t) \cdot g(t) dt \quad (7.63)$$

$$\text{auf } V = \mathcal{L}^2(\mathbb{R}, \mathbb{C}): \quad (f, g) := \int_{-\infty}^{\infty} f^*(t) \cdot g(t) dt. \quad (7.64)$$

Anwendungen: *FOURIER-Transformation, LAPLACE-Transformation, Signalverarbeitung, Variationsrechnung, FEM-Simulationen, Quantenphysik*

7.3.1.2 Mass-Formeln

Mit Hilfe eines *Skalar-Produkts* lassen sich *Masse* (*Längen, Flächen, Volumen, etc..*) definieren. Wir betrachten dazu die folgende Definition.

Definition 7.12 GRAM-Matrix

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* über dem *Zahlenkörper* \mathbb{K} , $m \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$. Die *GRAM-Matrix* dieser *Vektoren* ist

$$G(\mathbf{v}_1; \dots; \mathbf{v}_m) := \begin{bmatrix} \langle \mathbf{v}_1, \mathbf{v}_1 \rangle & \langle \mathbf{v}_1, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_1, \mathbf{v}_m \rangle \\ \langle \mathbf{v}_2, \mathbf{v}_1 \rangle & \langle \mathbf{v}_2, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_2, \mathbf{v}_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle \mathbf{v}_m, \mathbf{v}_1 \rangle & \langle \mathbf{v}_m, \mathbf{v}_2 \rangle & \dots & \langle \mathbf{v}_m, \mathbf{v}_m \rangle \end{bmatrix}. \quad (7.65)$$

Bemerkungen:

- i) Die *Komponenten* der *GRAM-Matrix* sind gerade alle möglichen *Skalar-Produkte*, die sich aus den *Vektoren* $\mathbf{v}_1, \dots, \mathbf{v}_m$ bilden lassen. Davon gibt es insgesamt m^2 und konsequenterweise gilt $G \in \mathbb{M}(m, m, \mathbb{K})$.
- ii) Wegen SP-2 muss gelten

$$G^T = G^*. \quad (7.66)$$

- iii) Je nach Wahl von $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ ergeben sich daraus unterschiedliche Eigenschaften der *GRAM-Matrix*. Es gilt

$$\mathbb{K} = \mathbb{R} \Rightarrow G^T = G \quad (\text{symmetrisch}) \quad (7.67)$$

$$\mathbb{K} = \mathbb{C} \Rightarrow G^T = G^* \quad (\text{hermitesch}). \quad (7.68)$$

- iv) Für $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und ein *positiv definites Skalar-Produkt* muss gelten

$$\det(G) \geq 0. \quad (7.69)$$

Wir betrachten die folgende Definition.

Definition 7.13 Mass

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* über dem *Zahlenkörper* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $m \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$. Das *Mass* der *Vektoren* ist

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) := \sqrt{|\det(G)|}. \quad (7.70)$$

Wir betrachten den folgenden Satz.

Satz 7.9 Regularität des Masses

Seien $(V, \mathbb{K}, +, \cdot)$ ein *Vektorraum* über dem *Zahlenkörper* $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $m \in \mathbb{N}^+$ und $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$. Dann gilt

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) = 0 \Leftrightarrow G(\mathbf{v}_1; \dots; \mathbf{v}_m) \text{ singulär} \Leftrightarrow \{\mathbf{v}_1, \dots, \mathbf{v}_m\} \text{ linear abhängig.} \quad (7.71)$$

Bemerkungen:

i) In jedem Fall gilt

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) \geq 0. \quad (7.72)$$

ii) Weil für ein *positiv definites Skalar-Produkt* gilt $\det(G) \geq 0$, kann der *Betrag* in diesem Fall weggelassen werden.

iii) Das *Mass* ist die Verallgemeinerung der Begriffe *Länge*, *Fläche* und *Volumen* auf beliebige *reelle* und *komplexe Vektorräume* mit *Skalar-Produkt*. Insbesondere gilt folgendes.

$$\mu(\mathbf{v}) = \sqrt{|\langle \mathbf{v}, \mathbf{v} \rangle|} = |\mathbf{v}| \equiv \text{Länge von } \mathbf{v} \quad (7.73)$$

$$\mu(\mathbf{v}_1; \mathbf{v}_2) \equiv \text{Fläche des von } \mathbf{v}_1 \text{ und } \mathbf{v}_2 \text{ aufgespannten Parallelogramms} \quad (7.74)$$

$$\mu(\mathbf{v}_1; \mathbf{v}_2; \mathbf{v}_3) \equiv \text{Volumen des von } \mathbf{v}_1, \mathbf{v}_2 \text{ und } \mathbf{v}_3 \text{ aufgespannten Spats} \quad (7.75)$$

iv) Für $m > \dim(V)$ ist die GRAM-Matrix G in jedem Fall *singulär* und es folgt

$$\mu(\mathbf{v}_1; \dots; \mathbf{v}_m) = 0. \quad (7.76)$$

v) Anhand ihres *Masses* lässt sich beurteilen, ob eine *Menge* von *Vektoren linear abhängig* oder *linear unabhängig* ist.

7.3.1.3 Eigenschaften reeller Skalar-Produkte

In diesem Abschnitt betrachten wir ausschliesslich *reelle Vektorräume* mit *Skalar-Produkt*. In jedem Fall gelten dann *binomische Formeln*. Wir betrachten dazu den folgenden Satz.

Satz 7.10 Binomische Formeln

Seien $(V, \mathbb{R}, +, \cdot)$ ein *reeller Vektorraum* mit *Skalar-Produkt* $\langle \cdot, \cdot \rangle$ und $\mathbf{v}, \mathbf{w} \in V$, dann gelten die folgenden *binomischen Formeln*.

(a) $\langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle + 2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle$

(b) $\langle \mathbf{v} - \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle - 2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle$

(c) $\langle \mathbf{v} + \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{w}, \mathbf{w} \rangle$

Beweis: Übung. □

Wir betrachten den folgenden Satz.

Satz 7.11 CAUCHY-SCHWARZ-Ungleichung

Seien $(V, \mathbb{R}, +, \cdot)$ ein *reeller Vektorraum* mit *positiv definitem Skalar-Produkt* $\langle \cdot, \cdot \rangle$ und $\mathbf{v}, \mathbf{w} \in V$, dann gilt die *CAUCHY-SCHWARZ-Ungleichung*

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq |\mathbf{v}| \cdot |\mathbf{w}|. \quad (7.77)$$

Beweis: Für alle $x \in \mathbb{R}$ betrachten wir den Vektor

$$\mathbf{u}(x) := x \cdot \mathbf{v} + \mathbf{w}. \quad (7.78)$$

Wegen der *positiven Definitheit* des Skalar-Produkts gilt

$$0 \leq \langle \mathbf{u}(x), \mathbf{u}(x) \rangle = \langle x \cdot \mathbf{v} + \mathbf{w}, x \cdot \mathbf{v} + \mathbf{w} \rangle \quad (7.79)$$

$$0 \leq \langle x \cdot \mathbf{v}, x \cdot \mathbf{v} \rangle + 2 \cdot \langle x \cdot \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \quad (7.80)$$

$$0 \leq x^2 \cdot \langle \mathbf{v}, \mathbf{v} \rangle + 2 \cdot x \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle \quad (7.81)$$

$$0 \leq \langle \mathbf{v}, \mathbf{v} \rangle \cdot x^2 + 2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle \cdot x + \langle \mathbf{w}, \mathbf{w} \rangle =: f(x). \quad (7.82)$$

Offensichtlich ist $f(x)$ eine *quadratische Funktion* mit *Grund-Form-Parameter*

$$a = \langle \mathbf{v}, \mathbf{v} \rangle, \quad b = 2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle \quad \text{und} \quad c = \langle \mathbf{w}, \mathbf{w} \rangle. \quad (7.83)$$

Für die *Diskriminante* von f muss gelten

$$0 \geq D = b^2 - 4 \cdot a \cdot c \quad (7.84)$$

$$0 \geq (2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle)^2 - 4 \cdot \langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \quad (7.85)$$

$$0 \geq 4 \cdot \langle \mathbf{v}, \mathbf{w} \rangle^2 - 4 \cdot \langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \quad | : 4 \quad (7.86)$$

$$0 \geq \langle \mathbf{v}, \mathbf{w} \rangle^2 - \langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \quad | + \langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \quad (7.87)$$

$$\langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle \geq \langle \mathbf{v}, \mathbf{w} \rangle^2 \quad | \sqrt{\dots} \quad (7.88)$$

Daraus folgt

$$|\underline{\langle \mathbf{v}, \mathbf{w} \rangle}| \leq \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle \cdot \langle \mathbf{w}, \mathbf{w} \rangle} = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \sqrt{\langle \mathbf{w}, \mathbf{w} \rangle} = |\underline{\mathbf{v}}| \cdot |\underline{\mathbf{w}}|. \quad (7.89)$$

Damit haben wir den Satz bewiesen. \square

Bemerkungen:

- i) Die CAUCHY-SCHWARZ-*Ungleichung* gilt nur für *positiv definite Skalar-Produkte*.
- ii) Die CAUCHY-SCHWARZ-*Ungleichung* gilt auch für *positiv definite Skalar-Produkte in komplexen Vektorräumen*, d.h. für $\mathbb{K} = \mathbb{C}$.
- iii) Im deutschsprachigen Raum ausserhalb Bayerns wird die CAUCHY-SCHWARZ-*Ungleichung* üblicherweise abgekürzt durch CSU.

Wir betrachten die folgende Definition.

Definition 7.14 Winkel zwischen Vektoren

Seien $(V, \mathbb{R}, +, \cdot)$ ein *reeller Vektorraum* mit *positiv definitem Skalar-Produkt* $\langle \cdot, \cdot \rangle$ und $\mathbf{v}, \mathbf{w} \in V$. Der *Winkel* zwischen \mathbf{v} und \mathbf{w} ist

$$\angle(\mathbf{v}; \mathbf{w}) := \begin{cases} \arccos\left(\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{v}| \cdot |\mathbf{w}|}\right) & 0 \notin \{\mathbf{v}, \mathbf{w}\} \\ \frac{\pi}{2} & 0 \in \{\mathbf{v}, \mathbf{w}\}. \end{cases} \quad (7.90)$$

Bemerkungen:

- i) Für alle *Skalar-Produkte* und alle $\mathbf{v}, \mathbf{w} \in V$ gilt die wichtige Konvention

$$\mathbf{v} \perp \mathbf{w} \Leftrightarrow \langle \mathbf{v}, \mathbf{w} \rangle = 0. \quad (7.91)$$

- ii) Für *positiv definite Skalar-Produkte* kann die Formel (7.90) in jeden Fall angewendet werden, denn aus der *CAUCHY-SCHWARZ-Ungleichung* folgt für alle $\mathbf{v}, \mathbf{w} \in V \setminus \{0\}$

$$\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{|\mathbf{v}| \cdot |\mathbf{w}|} \in [-1, 1]. \quad (7.92)$$

- iii) Für nicht *positiv definite Skalar-Produkte* lässt sich im allgemeinen keine universell gültige Formel für den *Winkel* zwischen zwei *Vektoren* finden.

7.3.2 Metrik

Wir betrachten die folgende Definition.

Definition 7.15 Metrik

Seien $(V, \mathbb{R}, +, \cdot)$ ein *reeller Vektorraum* mit *endlicher Dimension* $n \in \mathbb{N}^+$ und *Skalar-Produkt* $\langle \cdot, \cdot \rangle$ und $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subset V$ eine *Basis* von V . Die *Metrik* $g \in \mathbb{M}(n, n, \mathbb{R})$ ist die *GRAM-Matrix* der *Basis-Vektoren* in B , d.h.

$$g = G(\mathbf{e}_1; \dots; \mathbf{e}_n) := \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle & \dots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle \\ \vdots & \ddots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle & \dots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle \end{bmatrix}. \quad (7.93)$$

Bemerkungen:

- i) Die *Komponenten* der *Metrik* werden auch *metrische Koeffizienten* genannt. Für alle $i, j \in \{1, \dots, n\}$ gilt

$$g_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle. \quad (7.94)$$

- ii) Das Symbol g wird sowohl für die *Metrik* als auch für deren *Determinante* verwendet. Welche Bedeutung gerade gemeint ist, muss aus dem Kontext abgelesen werden.

- iii) Die *Metrik* ist in jedem Fall *regulär*, weil die *Basis-Vektoren* nach Definition *linear unabhängig* sein müssen.

- iv) Die *Metrik* ist genau dann *diagonal*, wenn die *Basis-Vektoren* paarweise *senkrecht* aufeinander stehen.

- v) In \mathbb{R}^n mit *GRAM-RIEMANN-Skalar-Produkt* haben die *Standard-Einheitsvektoren* $\{\hat{\mathbf{e}}_1, \dots, \hat{\mathbf{e}}_n\}$ die *Metrik*

$$\hat{g} = \mathbb{1} \quad \text{bzw.} \quad g_{ij} = \delta_{ij}. \quad (7.95)$$

Wir betrachten den folgenden Satz.

Satz 7.12 Metrische Skalar-Produkt-Formel

Seien $(V, \mathbb{R}, +, \cdot)$ ein reeller Vektorraum mit endlicher Dimension $n \in \mathbb{N}^+$ und Skalar-Produkt $\langle \cdot, \cdot \rangle$ und $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\} \subset V$ eine Basis von V mit Metrik $g \in \mathbb{M}(n, n, \mathbb{R})$ und $\mathbf{v}, \mathbf{w} \in V$ mit Basis-Darstellungen

$$\mathbf{v} = \sum_{r=1}^n v^r \cdot \mathbf{e}_r \quad \text{bzw.} \quad \mathbf{w} = \sum_{s=1}^n w^s \cdot \mathbf{e}_s \quad (7.96)$$

Dann gilt

$$\boxed{\langle \mathbf{v}, \mathbf{w} \rangle = \begin{bmatrix} v^1 & \dots & v^n \end{bmatrix} \cdot \begin{bmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \vdots & \vdots \\ g_{n1} & \dots & g_{nn} \end{bmatrix} \cdot \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix} = \mathbf{v}^T \cdot g \cdot \mathbf{w}.} \quad (7.97)$$

Beweis: Durch Einsetzen der Basis-Darstellungen und mit Hilfe der Bilinearität des Skalar-Produkts erhalten wir

$$\begin{aligned} \underline{\underline{\langle \mathbf{v}, \mathbf{w} \rangle}} &= \left\langle \sum_{r=1}^n v^r \cdot \mathbf{e}_r, \sum_{s=1}^n w^s \cdot \mathbf{e}_s \right\rangle = \sum_{r=1}^n \sum_{s=1}^n v^r \cdot w^s \cdot \langle \mathbf{e}_r, \mathbf{e}_s \rangle = \sum_{r=1}^n \sum_{s=1}^n v^r \cdot w^s \cdot g_{rs} \\ &= \sum_{r=1}^n v^r \cdot \sum_{s=1}^n g_{rs} \cdot w^s = \begin{bmatrix} v^1 & \dots & v^n \end{bmatrix} \cdot \begin{bmatrix} g_{11} \cdot w^1 + g_{12} \cdot w^2 + \dots + g_{1n} \cdot w^n \\ \vdots \\ g_{n1} \cdot w^1 + g_{n2} \cdot w^2 + \dots + g_{nn} \cdot w^n \end{bmatrix} \\ &= \begin{bmatrix} v^1 & \dots & v^n \end{bmatrix} \cdot \begin{bmatrix} g_{11} & \dots & g_{1n} \\ \vdots & \vdots & \vdots \\ g_{n1} & \dots & g_{nn} \end{bmatrix} \cdot \begin{bmatrix} w^1 \\ \vdots \\ w^n \end{bmatrix} = \mathbf{v}^T \cdot g \cdot \mathbf{w}. \end{aligned} \quad (7.98)$$

Damit haben wir den Satz bewiesen. □

Bemerkungen:

- i) Durch die *metrische Skalar-Produkt-Formel* kann das *Skalar-Produkt* in einem beliebigen Vektorraum mit fix gewählter *Basis* aus den Komponenten der Vektoren als *Matrix-Produkt* mit drei *Faktoren* berechnet werden.
- ii) In einem Vektorraum kann ein *Skalar-Produkt* definiert werden durch Angabe einer *Basis* und deren *Metrik*.
- iii) In \mathbb{R}^n mit GRAM-RIEMANN-Skalar-Produkt folgt

$$\underline{\underline{\langle \mathbf{v}, \mathbf{w} \rangle}} = \mathbf{v}^T \cdot g \cdot \mathbf{w} = \mathbf{v}^T \cdot \mathbb{1} \cdot \mathbf{w} = \underline{\underline{\mathbf{v}^T \cdot \mathbf{w}}} \quad (7.99)$$