

Kapitel 8

Integralrechnung

Der historische Zugang zur Integralrechnung fußt auf der praktischen Notwendigkeit, zur Landvermessung Inhalte von Flächen berechnen zu können (Babylonier). Inhalte geradlinig begrenzter Flächen wie Dreiecke, Rechtecke oder Trapeze waren bekannt. Also versuchte man krummlinig begrenzte Flächen durch Unterteilung der Fläche in geradlinig begrenzte Flächen zu approximieren.

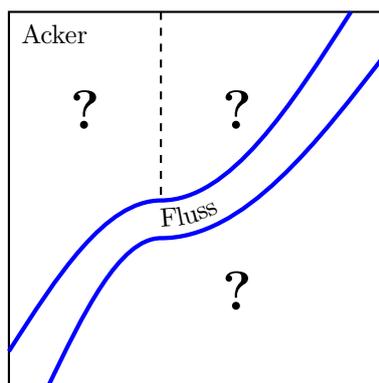


Abbildung 8.1: Fläche eines Ackers

Der griechische Mathematiker Archimedes (gest. 212 vor Chr.) führte den Ansatz ein, dem auch wir hier prinzipiell folgen werden:

Gegeben sei zum Beispiel eine Funktion f mit $f(x) \geq 0$ für $x \in [a, b]$. Gesucht ist die Fläche zwischen Abszisse (x -Achse) und dem Graph der Funktion.

Ansatz: Approximiere mit Rechtecken unterhalb des Graphen und den Graphen einschließenden Rechtecken!

1. *Beobachtung:* Die Summe der unteren Rechteckflächen ist kleiner als der gesamte Flächeninhalt, während die Summe der „oberen“ Rechteckflächen zu groß ist. Wie kann man die Näherung genauer machen?

Idee: Feinere Unterteilung!

2. *Beobachtung:* Durch hinzufügen neuer Punkte wird die Summe unterer Rechtecke größer und die der oberen kleiner!

Spekulation/Idee: Mache die Abstände immer kleiner, dann wird die Approxi-

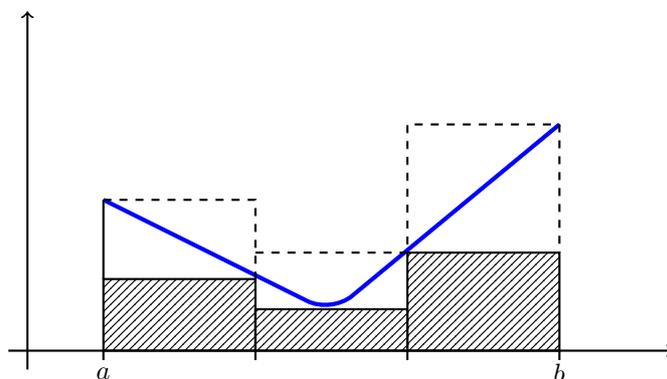


Abbildung 8.2: Approximation der Fläche unterhalb einer Funktion unter der x -Achse durch Rechtecke

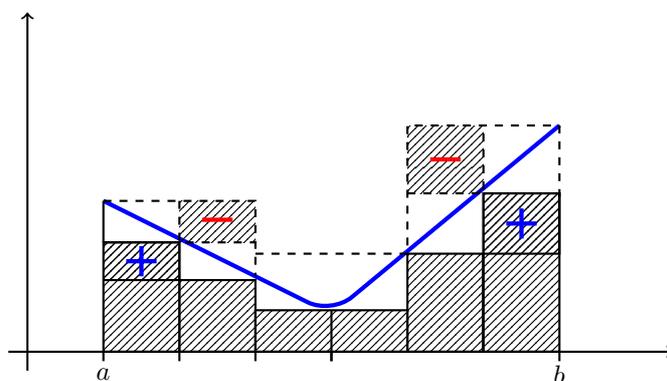


Abbildung 8.3: Eine feinere Unterteilung der Funktion aus Abb. 8.2

mation durch untere und obere Rechtecksummen immer besser – und durch infinitesimal kleine Abstände (Grenzübergang) können die Approximationen exakt werden. Diese Idee wollen wir nun präzisieren.

8.1 Das bestimmte Riemann-Integral

Definition 8.1. Gegeben seien ein Intervall $[a, b] \in \mathbb{R}$ und eine endliche Anzahl von Punkten x_0, x_1, \dots, x_n mit $a = x_0 < x_1 < \dots < x_n = b$. Dann heißt $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$ und $|Z| := \max\{x_i - x_{i-1} : i = 1, \dots, n\}$ das Feinheitsmaß der Zerlegung Z . Eine Zerlegung heißt äquidistant, wenn die Intervalle $[x_{i-1}, x_i]$ für $i = 1, \dots, n$ alle gleich groß sind.

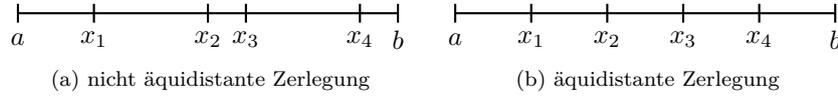


Abbildung 8.4: Zerlegungen

Definition 8.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt ($\forall x \in [a, b] : |f(x)| \leq K < \infty$) und Z eine Zerlegung auf $[a, b]$. Wir nennen

$$s(Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \inf\{f(x) : x \in [x_{i-1}, x_i]\} \text{ die Untersumme und}$$

$$S(Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \sup\{f(x) : x \in [x_{i-1}, x_i]\} \text{ die Obersumme}$$

von f bezüglich der Zerlegung Z .

Definition 8.3. Eine Zerlegung \tilde{Z} wird eine Verfeinerung von Zerlegung Z genannt (in Zeichen: $Z \leq \tilde{Z}$) wenn \tilde{Z} alle Punkte von Z enthält. Eine Zerlegung \hat{Z} , die genau die Punkte von Z und \tilde{Z} enthält, soll Überlagerung von Z und \tilde{Z} heißen und mit $\hat{Z} = Z + \tilde{Z}$ bezeichnet werden.

Lemma 8.4. Sei $f(x)$ auf $[a, b]$ beschränkt mit $|f(x)| \leq K$ und Z eine Zerlegung von $[a, b]$. Die Zerlegung \tilde{Z} entstehe aus Z durch Hinzunahme eines zusätzlichen Punktes. Dann gilt:

1. $s(Z) \leq s(\tilde{Z}) \leq s(Z) + 2K|Z|$
2. $S(Z) \geq S(\tilde{Z}) \geq S(Z) - 2K|Z|$

Beweis: Sei $Z = (x_0, \dots, x_n)$ und $\tilde{x} \in (x_{k-1}, x_k)$ für ein beliebiges $k = 1, \dots, n$.

Also: $\tilde{Z} = (x_0, \dots, x_{k-1}, \tilde{x}, x_k, \dots, x_n)$.

ad a) $m_k = \inf\{f(x) : x \in [x_{k-1}, x_k]\}$

$\tilde{m}_k = \inf\{f(x) : x \in [x_{k-1}, \tilde{x}]\} \geq m_k$

$\tilde{m}_2 = \inf\{f(x) : x \in [\tilde{x}, x_k]\} \geq m_k$

$$\begin{aligned} s(\tilde{Z}) &= s(Z) - (x_k - x_{k-1})m_k + (\tilde{x} - x_{k-1})\tilde{m}_1 + (x_k - \tilde{x})\tilde{m}_2 \\ \Leftrightarrow s(\tilde{Z}) - s(Z) &= (\tilde{x} - x_{k-1})\tilde{m}_1 + (x_k - \tilde{x})\tilde{m}_2 - (x_k - x_{k-1})m_k \\ &= \tilde{x}\tilde{m}_1 - x_{k-1}\tilde{m}_1 + x_k\tilde{m}_1 + x_k\tilde{m}_2 - \tilde{x}\tilde{m}_2 - x_k m_k \\ &\quad + x_{k-1}m_k + \underbrace{\tilde{x}m_k - \tilde{x}m_k}_{\text{Trick!}} \\ &= \underbrace{(\tilde{x} - x_{k-1})}_{\geq 0} \underbrace{(\tilde{m}_1 - m_k)}_{\geq 0} + \underbrace{(x_k - \tilde{x})}_{\geq 0} \underbrace{(\tilde{m}_2 - m_k)}_{\geq 0} \geq 0 \\ &\leq (\tilde{x} - x_{k-1}) \cdot 2K + (x_k - \tilde{x}) \cdot 2K \\ &= \underbrace{(x_k - x_{k-1})}_{\leq |Z|} \cdot 2K \leq 2K \cdot |Z| \end{aligned}$$

ad b) analog. □

Satz 8.5. Jede Untersumme ist kleiner oder gleich jeder Obersumme.

Beweis: Seien Z und \tilde{Z} zwei beliebige Zerlegungen. Dann gilt

$$s(Z) \stackrel{\text{Lem.}}{\leq} s(Z + \tilde{Z}) \stackrel{\text{Def.}}{\leq} S(Z + \tilde{Z}) \stackrel{\text{Lem.}}{\leq} S(\tilde{Z}) \quad \square$$

Definition 8.6. Die Funktion $f(x)$ sei auf $[a, b]$ beschränkt. Man nennt $\int_{-a}^b f(x) dx := \sup\{s(Z) : Z \text{ ist eine Zerlegung von } [a, b]\}$ das untere (Riemann-)Integral und $\int_a^b f(x) dx := \inf\{S(Z) : Z \text{ ist eine Zerlegung von } [a, b]\}$ das obere (Riemann-)Integral. Sind unteres und oberes Riemann-Integral gleich, dann heißt $f(x)$ über $[a, b]$ (Riemann-)integrierbar und $\int_a^b f(x) dx := \sup_Z s(Z) = \inf_Z S(Z)$ das (Riemann-)Integral von f über $[a, b]$. Man nennt a (bzw. b) die untere (bzw. obere) Integrationsgrenze und x die Integrationsvariable.

Anmerkung: Um auszudrücken, dass f Riemann-integrierbar über $[a, b]$ ist, schreiben wir auch $f \in R[a, b]$.

Satz 8.7. Es gilt $\int_{-a}^b f(x) dx \leq \int_a^b f(x) dx$ bzw. $\sup_Z s(Z) \leq \inf_Z S(Z)$.

Beweis: Für zwei beliebige Zerlegungen Z und \tilde{Z} gilt nach Lemma 8.5 stets $s(Z) \leq S(\tilde{Z})$. Also gilt für ein beliebig fixiertes \tilde{Z} auch $\sup_Z s(Z) \leq S(\tilde{Z})$. Da \tilde{Z} beliebig, kann es derart gewählt werden, dass $S(\tilde{Z}) = \inf_Z S(Z)$. Folglich $\sup_Z s(Z) \leq \inf_Z S(Z)$. \square

Das folgende Resultat wird uns die Bestimmung von unterem und oberem Riemann-Integral spürbar erleichtern, da nicht mehr zwingend Supremum und Infimum der Summen über *alle möglichen* Zerlegungen gefunden werden müssen. Vielmehr reicht es dann, den Grenzwert *irgendeiner Zerlegungsnullfolge* für Unter- und Obersumme zu berechnen.

Satz 8.8. Sei f auf $[a, b]$ beschränkt und Z_n eine Folge von Zerlegungen mit $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$. Dann gilt

a) $\lim_{n \rightarrow \infty} s(Z_n) = \sup_Z s(Z)$ und

b) $\lim_{n \rightarrow \infty} S(Z_n) = \inf_Z S(Z)$

Beweis:

ad a) Sei \tilde{Z} eine beliebige Zerlegung mit p inneren Teilpunkten. Wegen Lemma 8.4 gilt dann $s(Z_n) \leq s(Z_n + \tilde{Z}) \leq s(Z_n) + 2pK|Z_n|$.

Wenn $\lim_{n \rightarrow \infty} s(Z_n) = \alpha$ existiert, dann $\underbrace{s(Z_n)}_{\rightarrow \alpha} + \underbrace{2pK|Z_n|}_{\rightarrow 0} \rightarrow \alpha$ und schließlich

$s(Z_n + \tilde{Z}) \rightarrow \alpha$ für $n \rightarrow \infty$. Da aber auch $s(\tilde{Z}) \leq \underbrace{s(Z_n + \tilde{Z})}_{\rightarrow \alpha} \leq \sup_Z s(Z)$ gilt,

folgt $s(\tilde{Z}) \leq \alpha \leq \sup_Z s(Z)$.

Weil \tilde{Z} beliebig, kann \tilde{Z} auch so gewählt werden, dass $s(\tilde{Z}) = \sup_Z s(Z)$. Also muss $\alpha = \sup_Z s(Z)$ gelten.

ad b) Analog. □

Beispiel 8.1. Notation: $M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\}$ und $m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$ mit $i = (1, 2, \dots, n)$.

1. $f(x) = c = \text{const.}$ für alle $x \in [a, b]$. Offensichtlich ist $m_i = M_i = c$ für alle $i = 1, 2, \dots, n$ und deshalb $s(Z) = S(Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot c \stackrel{\text{Teleskopsumme}}{=} c \cdot (x_n - x_0) = c(b - a)$ für jede beliebige Zerlegung Z von $[a, b]$.

2. $f(x) = x$ auf $[a, b]$. Wähle äquidistante Zerlegung $Z_n : x_i = i \cdot \frac{a}{n}$ für alle $i = 1, 2, \dots, n$.

$$m_i = \inf_{\substack{f(x) \\ =x}}\{f(x) : x \in [x_{i-1}, x_i]\} = x_{i-1} = (i-1) \frac{a}{n}$$

$$M_i = \sup\{x : x \in [x_{i-1}, x_i]\} = x_i = i \cdot \frac{a}{n}$$

$$\begin{aligned} \Rightarrow s(Z_n) &= \sum_{i=1}^n \underbrace{(x_i - x_{i-1})}_{a/n} \cdot m_i = \sum_{i=1}^n \frac{a}{n} \cdot (i-1) \cdot \frac{a}{n} = \frac{a^2}{n^2} \sum_{i=1}^n (i-1) \\ &= \frac{a^2}{n^2} \sum_{i=1}^{n-1} i = \frac{a^2}{n^2} \cdot \frac{(n-1)n}{2} = \frac{a^2}{2} \left(1 - \frac{1}{n}\right) \end{aligned}$$

$$\begin{aligned} S(Z_n) &= \sum_{i=1}^n \underbrace{(x_i - x_{i-1})}_{a/n} \cdot M_i = \sum_{i=1}^n \frac{a}{n} \cdot i \cdot \frac{a}{n} = \frac{a^2}{n^2} \sum_{i=1}^n i = \frac{a^2}{n^2} \cdot \frac{n(n+1)}{2} \\ &= \frac{a^2}{2} \left(1 + \frac{1}{n}\right) \end{aligned}$$

$$\left. \begin{aligned} \sup_n s(Z_n) &= \sup_n \frac{a^2}{2} \left(1 - \frac{1}{n}\right) = \frac{a^2}{2} \\ \inf_n S(Z_n) &= \inf_n \frac{a^2}{2} \left(1 + \frac{1}{n}\right) = \frac{a^2}{2} \end{aligned} \right\} \Rightarrow f(x) \in R[0, a]$$

$$\longrightarrow \int_0^a x \, dx = \frac{a^2}{2}$$

3. $f(x) = e^x$ in $[0, a]$. Wähle wieder äquidistante Zerlegung $Z_n : x_i = i \frac{a}{n}$ für $i = 1, 2, \dots, n$. Wegen Monotonie:

$$\begin{aligned} m_i &= e^{x_{i-1}} = e^{(i-1) \frac{a}{n}} = \left(e^{\frac{a}{n}}\right)^{i-1} =: q^{i-1} \text{ mit } q := e^{\frac{a}{n}} \\ M_i &= e^{x_i} = e^{i \cdot \frac{a}{n}} = \left(e^{\frac{a}{n}}\right)^i = q^i \end{aligned}$$

$$\begin{aligned} s(Z_n) &= \sum_{i=1}^n (x_i - x_{i-1}) m_i = \sum_{i=1}^n \frac{a}{n} \cdot q^{i-1} = \frac{a}{n} \sum_{i=1}^n q^{i-1} = \frac{a}{n} \sum_{i=1}^{n-1} q^i \\ &= \frac{a}{n} \cdot \frac{1-q^n}{1-q} = \frac{a}{n} \cdot \frac{1-(e^{\frac{a}{n}})^n}{1-e^{\frac{a}{n}}} = (e^a - 1) \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1} \end{aligned}$$

$$\begin{aligned} S(Z_n) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot M_i = \sum_{i=1}^n \frac{a}{n} \cdot q^i = \frac{a}{n} \cdot q \cdot \sum_{i=1}^n q^{i-1} \\ &= \frac{a}{n} q \sum_{i=1}^{n-1} q^i = \frac{a}{n} \cdot q \cdot \frac{1-q^n}{1-q} = \frac{a}{n} \cdot e^{\frac{a}{n}} \cdot \frac{1-(e^{\frac{a}{n}})^n}{1-e^{\frac{a}{n}}} \\ &= (e^a - 1) \cdot e^{\frac{a}{n}} \cdot \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1} \end{aligned}$$

$$\begin{aligned} \sup_n s(Z_n) &= (e^a - 1) \cdot \sup_n \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1} \\ &\stackrel{\text{S. 8.8}}{=} (e^a - 1) \lim_{n \rightarrow \infty} \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1} \quad (h := \frac{a}{n}) \\ &= (e^a - 1) \lim_{h \rightarrow 0} \frac{h}{e^h - 1} \quad (\text{l'Hospital}) \\ &= (e^a - 1) \underbrace{\lim_{h \rightarrow 0} \frac{1}{e^h}}_{=1} = e^a - 1 \end{aligned}$$

$$\begin{aligned} \inf_n S(Z_n) &= (e^a - 1) \lim_{n \rightarrow \infty} e^{\frac{a}{n}} \cdot \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1} \quad (\text{s. o.}) \\ &= (e^a - 1) \underbrace{\left[\lim_{n \rightarrow \infty} e^{\frac{a}{n}} \right]}_{=1} \cdot \underbrace{\lim_{n \rightarrow \infty} \frac{\frac{a}{n}}{e^{\frac{a}{n}} - 1}}_{=1 \text{ (s.o.)}} = e^a - 1 \end{aligned}$$

Also ist e^x auf $[0, a]$ Riemann-integrierbar und es ist $\int_0^a e^x dx = e^a - 1$.

$$4. f(x) = \begin{cases} 0 & \text{für rationale } x \in [0, 1] \\ 1 & \text{für irrationale } x \in [0, 1] \end{cases}$$

Offensichtlich ist $m_i = 0$ und $M_i = 1$ für $i = 1, \dots, n$ und deshalb

$$\begin{aligned} s(Z_n) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot \underbrace{m_i}_{=0} = 0 \text{ für jede Zerlegung } Z_n \\ S(Z_n) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot \underbrace{M_i}_{=1} = \underbrace{x_n}_{=1} - \underbrace{x_0}_{=0} = 1 \text{ für jede Zerlegung } Z_n \\ \Rightarrow f(x) &\notin R[0, 1]. \end{aligned}$$

Beobachtung: Anscheinend sind nicht alle Funktionen Riemann-integrierbar! Können wir die Riemann-integrierbaren Funktionen charakterisieren?

Satz 8.9 (Riemannsches Integritätskriterium). Sei $f(x)$ auf $[a, b]$ beschränkt.

$$f(x) \in R[a, b] \Leftrightarrow \forall \varepsilon > 0 : \exists Z(\varepsilon) : S(Z) - s(Z) < \varepsilon$$

Beweis:

“ \Rightarrow ” Sei $f(x) \in R[a, b]$. Dann gilt $\sup_Z s(Z) = \inf_Z S(Z)$. Wähle $\varepsilon > 0$ beliebig. Man kann eine Zerlegung \tilde{Z} wählen, so dass $0 \leq \sup_Z s(Z) - s(\tilde{Z}) \leq \frac{\varepsilon}{2}$. Ebenso kann man eine Zerlegung \hat{Z} wählen, so dass $0 \leq S(\hat{Z}) - \inf_Z S(Z) \leq \frac{\varepsilon}{2}$. Sei nun $Z = \tilde{Z} + \hat{Z}$. Dann gilt:

$$\begin{aligned} &\left. \begin{aligned} 0 \leq \int_a^b f(x) dx - s(Z) &< \frac{\varepsilon}{2} \\ 0 \leq S(Z) - \int_a^b f(x) dx &< \frac{\varepsilon}{2} \end{aligned} \right\} + \\ \Rightarrow &\underline{0 \leq S(Z) - s(Z) < \varepsilon} \text{ für Zerlegung } Z = \tilde{Z} + \hat{Z} \end{aligned}$$

“ \Leftarrow ” $\forall \varepsilon > 0 : \exists Z_\varepsilon : S(Z_\varepsilon) - s(Z_\varepsilon) < \varepsilon$ bzw. $S(Z_\varepsilon) < s(Z_\varepsilon) + \varepsilon$.

$$\begin{aligned} \text{Also } s(Z_\varepsilon) \leq \sup_Z s(Z) &\stackrel{\text{Satz 8.7}}{\leq} \inf_Z S(Z) \leq S(Z_\varepsilon) \stackrel{\text{nach Vor.}}{\leq} s(Z_\varepsilon) + \varepsilon \quad \left| - \sup_Z s(Z) \right. \\ 0 \leq \inf_Z S(Z) - \sup_Z s(Z) &< \underbrace{s(Z_\varepsilon) - \sup_Z s(Z)}_{\leq 0} + \varepsilon \leq \varepsilon \text{ f\"ur alle } \varepsilon > 0. \end{aligned}$$

Mit $\varepsilon \rightarrow 0$ folgt also $\inf_Z S(Z) = \sup_Z s(Z)$ und $f \in R[a, b]$. □

Satz 8.10.

a) $f(x)$ stetig auf $[a, b] \Rightarrow f(x) \in R[a, b]$.

b) $f(x)$ auf $[a, b]$ monoton $\Rightarrow f(x) \in R[a, b]$.

Beweis:

a) Nach Satz 4.26 ist $f(x)$ sogar gleichmaig stetig auf $[a, b]$.

Also gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit $|f(x) - f(\tilde{x})| < \frac{\varepsilon}{b-a}$ fur alle $x, \tilde{x} \in [a, b]$ mit $|x - \tilde{x}| < \delta$.

Fur ein beliebiges $Z = (x_0, \dots, x_n)$ mit $|Z| < \delta$ ist $M_i - m_i < \frac{\varepsilon}{b-a}$ und

$$\begin{aligned} S(Z) - s(Z) &= \sum_{i=1}^n (x_i - x_{i-1})(M_i - m_i) \\ &< \frac{\varepsilon}{b-a} \cdot \underbrace{\sum_{i=1}^n (x_i - x_{i-1})}_{=x_n - x_0 = b-a} = \varepsilon. \end{aligned}$$

Laut Satz 8.9 ist damit $f(x) \in R[a, b]$.

b) Wahle eine aquidistante Zerlegung Z_n mit $x_i = a + \frac{b-a}{n} \cdot i$ fur $i = 0, 1, \dots, n$. Falls $f(x)$ monoton wachsend, dann $m_i = f(x_{i-1})$ und $M_i = f(x_i)$, sonst $M_i = f(x_{i-1})$ und $m_i = f(x_i)$. Folglich ist

$$\begin{aligned} S(Z_n) - s(Z_n) &= \sum_{i=1}^n (M_i - m_i) \underbrace{(x_i - x_{i-1})}_{\frac{b-a}{n}} \\ &= \frac{b-a}{n} \sum_{i=1}^n |f(x_i) - f(x_{i-1})| \\ &= \frac{b-a}{n} |f(b) - f(a)| \rightarrow 0 \quad \text{fur } n \rightarrow \infty \end{aligned}$$

Also $\inf_n S(Z_n) = \sup_n s(Z_n)$ und $f(x) \in R[a, b]$. □

Dieser Satz sichert uns Riemann-Integrierbarkeit fur eine groe Klasse von Funktionen zu. Wenn also die Integrierbarkeit von f auf $[a, b]$ gesichert ist, dann reicht es nur den Grenzwert der Unter- oder Obersumme zu bestimmen. Nachteilig bei dem Vorgehen bleibt jedoch der Umstand, dass Minima oder Maxima in den Teilintervallen berechnet werden mussen. Dieser Muhe wollen wir uns nun entledigen.

Definition 8.11. Sei $f(x)$ auf $[a, b]$ beschränkt, $Z = (x_0, x_1, \dots, x_n)$ eine Zerlegung von $[a, b]$ und $\tilde{x}_i \in [x_{i-1}, x_i]$ für $i = 1, \dots, n$. Man nennt

$$\sigma(Z, \tilde{x}) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\tilde{x}_i)$$

die (Riemannsche) Zwischensumme von f auf $[a, b]$.

Wegen $m_k \leq f(\tilde{x}_k) \leq M_k$ folgt aus der Definition sofort für alle Z

$$s(Z) \leq \sigma(Z, \tilde{x}) \leq S(Z).$$

Tatsächlich kann man durch geschickte Wahl von Zwischenpunkten $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ den Unter- und Obersummen beliebig nahe kommen:

Lemma 8.12. Sei $f(x)$ auf $[a, b]$ beschränkt. Für jede Zerlegung Z von $[a, b]$ und beliebiges $\varepsilon > 0$ gibt es Zwischenpunkte \tilde{x} und \hat{x} mit

a) $s(Z) \leq \sigma(Z, \tilde{x}) \leq s(Z) + \varepsilon$ und

b) $S(Z) - \varepsilon \leq \sigma(Z, \hat{x}) \leq S(Z)$

Beweis: Sei Z Zerlegung (x_0, \dots, x_n) von $[a, b]$ und $\varepsilon > 0$ gegeben.

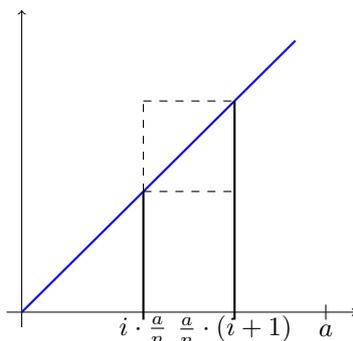
a) Für jedes $k = 1, \dots, n$ kann man $\tilde{x} \in [x_{k-1}, x_k]$ derart wählen, dass $m_k \leq f(\tilde{x}_k) < m_k + \frac{\varepsilon}{b-a}$ gilt. (je kleiner ε , desto näher schiebe \tilde{x}_k an Minimalstelle im Intervall). Mit diesem Satz von Zwischenpunkten ist dann

$$\begin{aligned} s(Z) &\stackrel{\text{per Def.}}{\leq} \sigma(Z, \tilde{x}) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\tilde{x}_i) \\ &\leq \sum_{i=1}^n (x_i - x_{i-1}) \cdot \left(m_k + \frac{\varepsilon}{b-a} \right) \\ &= s(Z) + \frac{\varepsilon}{b-a} \cdot \underbrace{\sum_{i=1}^n (x_i - x_{i-1})}_{=x_n - x_0 = b-a} \\ &= s(Z) + \varepsilon \end{aligned}$$

b) Analog für Obersummen.

□

Wenn also die Zerlegungen immer feiner werden, so dass $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$, dann rücken auch die Positionen von \tilde{x}_k und \hat{x}_k immer näher zusammen.



Satz 8.13. Sei f auf $[a, b]$ beschränkt, Z_n eine Folge von Zerlegungen mit $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$ mit passenden Zwischenpunkten $\tilde{x}^{(n)}$. Es gilt

$$f \in R[a, b] \Leftrightarrow \text{Jede Riemansche Zwischensumme konvergiert.}$$

In diesem Fall sind alle Grenzwerte gleich und sie haben den Wert

$$\lim_{n \rightarrow \infty} \sigma(Z_n, \tilde{x}^{(n)}) = \int_a^b f(x) dx =: I$$

Beweis:

\Rightarrow Sei $\sigma(Z_n, \tilde{x}^{(n)})$ eine beliebige Folge von Zwischensummen mit $|Z_n| \rightarrow 0$. Ist f integrierbar, so konvergieren sowohl $s(Z_n)$ als auch $S(Z_n)$ gegen I (Definition 8.6). Wegen $\underbrace{s(Z_n)}_{\rightarrow I} \leq \sigma(Z_n, \tilde{x}^{(n)}) \leq \underbrace{S(Z_n)}_{\rightarrow I}$ folgt sofort $\sigma(Z_n, \tilde{x}^{(n)}) \rightarrow I$ für $n \rightarrow \infty$ (Sandwich-Theorem).

\Leftarrow Sei Z_n eine Folge von Zerlegungen mit $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$ und jede zugehörige Folge von Summen $\sigma(Z_n, \hat{x}^{(n)})$ konvergent. Gemäß Lemma 8.12 gibt es zu jedem Z_n Zwischenpunkte $\tilde{x}^{(n)}$ und $\hat{x}^{(n)}$ mit

$$\underbrace{s(Z_n)}_{\rightarrow \sup_Z s(Z)} \leq \sigma(Z_n, \tilde{x}^{(n)}) < \underbrace{s(Z_n)}_{\rightarrow \sup_Z s(Z)} + \underbrace{\frac{1}{n}}_{\rightarrow 0}$$

und

$$\underbrace{S(Z_n)}_{\rightarrow \inf_Z S(Z)} - \underbrace{\frac{1}{n}}_{\rightarrow 0} < \sigma(Z_n, \hat{x}^{(n)}) \leq \underbrace{S(Z_n)}_{\rightarrow \inf_Z S(Z)}$$

damit $\sigma(Z_n, \tilde{x}^{(n)}) \rightarrow \sup_Z s(Z)$ und $\sigma(Z_n, \hat{x}^{(n)}) \rightarrow \inf_Z S(Z)$

Wäre $\tilde{x}^{(n)} = \hat{x}^{(n)}$, dann wären wir jetzt fertig, da Grenzwerte eindeutig sind und somit $\sup_Z s(Z) = \inf_Z S(Z)$ sein müsste. Mische nun beide Summenfolgen zu neuer Summenfolge $\sigma(Z_1, \hat{x}^{(1)})$, $\sigma(Z_2, \hat{x}^{(2)})$, $\sigma(Z_3, \hat{x}^{(3)})$, $\sigma(Z_4, \hat{x}^{(4)})$, ... Nach Konstruktion besitzt sie zwei konvergente Teilfolgen, die gegen das untere bzw. obere Riemann-Integral konvergieren. Die Mischfolge konvergiert aber nach

Voraussetzung auch. Deshalb müssen die Grenzwerte der Teilfolgen, also unteres und oberes Riemann-Integral, gleich sein. Damit ist $f \in R[a, b]$. \square

Praktischer Nutzen dieses Resultats: Steht Integrierbarkeit von f bereits fest (z. B. weil stetig oder monoton), dann kann man sich eine spezielle Wahl von Zerlegung und Zwischenpunkten aussuchen, um das Integral zu berechnen.

- Äquidistante Zerlegung: Für $f \in R[a, b]$ gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{i=1}^{n-1} hf(a + ih) \text{ mit } h = \frac{b-a}{n}$$

Beispiel: $f(x) = e^{\alpha x}$ für $\alpha \neq 0$ (monoton, stetig $\Rightarrow \in R[a, b]$)

$$\begin{aligned} \sum_{i=0}^{n-1} hf(a + ih) &= h \sum_{i=0}^{n-1} e^{\alpha(a+ih)} &&= he^{\alpha a} \sum_{i=0}^{n-1} (e^{\alpha h})^i \\ &= he^{\alpha a} \frac{(e^{\alpha h})^n - 1}{e^{\alpha h} - 1} &&= he^{\alpha a} \frac{e^{\alpha \frac{b-a}{n} n} - 1}{e^{\alpha h} - 1} \\ &= he^{\alpha a} \frac{e^{\alpha b - \alpha a} - 1}{e^{\alpha h} - 1} &&= (e^{\alpha b} - e^{\alpha a}) \frac{h}{e^{\alpha a} - 1} \frac{\alpha}{\alpha} \\ &= \frac{(e^{\alpha b} - e^{\alpha a})}{\alpha} \cdot \underbrace{\frac{\alpha h}{e^{\alpha h} - 1}}_{\rightarrow 1 \text{ für } h \rightarrow 0} \rightarrow \frac{e^{\alpha b} - e^{\alpha a}}{\alpha} \end{aligned}$$

weil $\lim_{h \rightarrow 0} \frac{\alpha \cdot h}{e^{\alpha h} - 1} \stackrel{\text{l'Hospital}}{=} \lim_{h \rightarrow 0} \frac{\alpha}{\alpha \cdot e^{\alpha h}} = 1$.

- Geometrische Progression (für $[a, b]$ mit $0 < a < b$).

$Z_n : x_i = a \cdot q^i$ für $i = 1, 2, \dots, n$ mit $q = \sqrt[n]{\frac{b}{a}} = \left(\frac{b}{a}\right)^{\frac{1}{n}} > 1$.

Für $k = 1, 2, \dots, n$ gilt $x_k - x_{k-1} = a(q^k - q^{k-1}) = a \cdot q^{k-1}(q - 1)$.

$$\begin{aligned} |Z_n| &= \max_{k=1, \dots, n} \{a \cdot q^{k-1}(q - 1)\} = a \cdot q^{n-1}(q - 1) \\ &\leq \underbrace{a \cdot q^n}_{=b} (q - 1) &&= b \left(\underbrace{\sqrt[n]{\frac{b}{a}}}_{\rightarrow 1} - 1\right) \rightarrow 0 \end{aligned}$$

Also: $\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n a \cdot q^{k-1}(q - 1) \cdot f(a \cdot q^{k-1})$ mit $q = \sqrt[n]{\frac{b}{a}}$

Beispiel 8.2. $f(x) = x^\alpha$ für $\alpha \neq -1$

$$\begin{aligned} a(q - 1) \sum_{k=0}^{n-1} q^k \cdot f(a \cdot q^k) &= a(q - 1) \sum_{k=0}^{n-1} q^k \cdot a^\alpha \cdot q^{\alpha k} \\ &= a^{\alpha+1}(q - 1) \sum_{k=0}^{n-1} (q^{\alpha+1})^k = a^{\alpha+1}(q - 1) \frac{(q^{\alpha+1})^n - 1}{q^{\alpha+1} - 1} \\ &\stackrel{(*)}{=} a^{\alpha+1}(q - 1) \frac{\left(\frac{b}{a}\right)^{\alpha+1} - 1}{q^{\alpha+1} - 1} = (b^{\alpha+1} - a^{\alpha+1}) \cdot \underbrace{\frac{q - 1}{q^{\alpha+1} - 1}}_{\rightarrow \frac{1}{\alpha+1}} \rightarrow \frac{b^{\alpha+1} - a^{\alpha+1}}{\alpha + 1} \end{aligned}$$

$$\lim_{q \rightarrow 1} \frac{q-1}{q^{\alpha+1}-1} \stackrel{L'Hospital}{=} \lim_{q \rightarrow 1} \frac{1}{(\alpha+1) \cdot q^\alpha} = \frac{1}{\alpha+1}$$

$$(*) : (q^{\alpha+1})^n = \left(\frac{b}{a}\right)^{\frac{1}{n}(\alpha+1)n} = \left(\frac{b}{a}\right)^{\alpha+1}$$

Satz 8.14. Sind f und g auf $[a, b]$ integrierbar, so ist für beliebige Konstanten α und β auch die Funktion $\alpha f + \beta g$ auf $[a, b]$ integrierbar und es ist

$$\int_a^b (\alpha f(x) + \beta g(x)) \, dx = \alpha \cdot \int_a^b f(x) \, dx + \beta \cdot \int_a^b g(x) \, dx .$$

Beweis: Gegeben sei $(Z_n, \tilde{x}^{(n)})$ mit $|Z_n| \rightarrow 0$. Dann ist

$$\begin{aligned} \sigma(Z_n, \tilde{x}^{(n)}; \alpha f + \beta g) &= \sum_{k=1}^n (x_k - x_{k-1})(\alpha f(\tilde{x}_k) + \beta g(\tilde{x}_k)) \\ &= \alpha \cdot \sum_{k=1}^n (x_k - x_{k-1})f(\tilde{x}_k) + \beta \sum_{k=1}^n (x_k - x_{k-1})g(\tilde{x}_k) \\ &= \alpha \cdot \underbrace{\sigma(Z_n, \tilde{x}^{(n)}; f)}_{\substack{\text{Grenzwert existiert} \\ \text{nach Vor.} \\ \text{nämlich } \int f \, dx}} + \beta \cdot \underbrace{\sigma(Z_n, \tilde{x}^{(n)}; g)}_{\substack{\text{Grenzwert existiert} \\ \text{nach Vor.} \\ \text{nämlich } \int g \, dx}} \end{aligned}$$

$\Rightarrow \lim_{n \rightarrow \infty} \sigma(Z_n, \tilde{x}^{(n)}; \alpha f + \beta g)$ existiert und ist nach Satz 8.13 gerade $\int (\alpha f + \beta g) \, dx$. □

Satz 8.15. Seien $f(x)$ und $g(x)$ integrierbar auf $[a, b]$. Dann gilt:

a) $\max\{f(x), g(x)\}$, $\min\{f(x), g(x)\}$, $|f(x)|$ sind integrierbar auf $[a, b]$.

b) Falls $f(x) \leq g(x)$ für alle $x \in [a, b]$, dann $\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx$.

c) $\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx$

Beweis:

a) Sei $h(x) = \max\{f(x), g(x)\}$ mit $M_h = \sup\{h(x) \mid x \in [a, b]\}$ sowie $m_h = \inf\{h(x) \mid x \in [a, b]\}$ und M_f, M_g, m_f, m_g entsprechend. Es ist stets $m_h \geq m_f$ und $m_h \geq m_g$. Falls $M_f \geq M_g$, dann $M_h = M_f$ und $M_h - m_h = M_f - m_h \leq M_f - m_f \leq M_f - m_f + \underbrace{M_g - m_g}_{\geq 0}$

Falls $M_g \geq M_f$, dann $M_h = M_g$ und $M_h - m_h = M_g - m_h \leq M_g - m_g \leq M_g - m_g + \underbrace{M_f - m_f}_{\geq 0}$

Also ist für eine beliebige Zerlegung Z

$$S_h(Z) - s_h(Z) \leq \underbrace{S_f(Z) - s_f(Z)}_{\leq \frac{\varepsilon}{2} \text{ da } f \in R[a, b]} + \underbrace{S_g(Z) - s_g(Z)}_{\leq \frac{\varepsilon}{2} \text{ da } g \in R[a, b]} \leq \varepsilon$$

und damit $h \in R[a, b]$ nach Satz 8.10. Mit den Identitäten $\min\{f(x), g(x)\} = -\max\{-f(x), -g(x)\}$ und $|f(x)| = \max\{0, f(x)\} + \max\{0, -f(x)\}$ erhält man die übrigen Behauptungen in Teil a).

b) Sei Z_n eine Folge von Zerlegungen mit $|Z_n| \rightarrow 0$ für $n \rightarrow \infty$ und $\tilde{x}^{(n)}$ eine Folge von passenden Zwischenwerten. Es ist

$$\begin{array}{ccc} \sigma(Z_n, \tilde{x}^{(n)}; f) = \sum_{i=1}^n \underbrace{(x_i - x_{i-1})}_{\geq 0} f(\tilde{x}_i) & \leq & \sum_{i=1}^n \underbrace{(x_i - x_{i-1})}_{\geq 0} g(\tilde{x}_i) = \sigma(Z_n, \tilde{x}^{(n)}; g) \\ \downarrow n \rightarrow \infty & & \downarrow n \rightarrow \infty \\ \int_a^b f(x) dx & \leq & \int_a^b g(x) dx \end{array}$$

unter Verwendung von Satz 8.13.

c) Wegen $f(x) \leq |f(x)|$ und $-f(x) \leq |f(x)|$ ergibt sich aus b) sofort

$$\int_a^b f(x) dx \leq \int_a^b |f(x)| dx \quad \text{sowie} \quad -\int_a^b f(x) dx \leq \int_a^b |f(x)| dx$$

und damit die Behauptung. □

Satz 8.16 (Mittelwertsatz der Integralrechnung). *Sei $f \in R[a, b]$ und $m \leq f(x) \leq M$ für $x \in [a, b]$. Dann ist*

a) $m(b - a) \leq \int_a^b f(x) dx \leq M(b - a)$

b) *Ist $f(x)$ auf $[a, b]$ zudem stetig, dann existiert ein $\tilde{x} \in (a, b)$ mit*

$$\int_a^b f(x) dx = (b - a)f(\tilde{x})$$

Beweis:

a) Integration der Ungleichungen $m \leq f(x) \leq M$ gibt mit Satz 8.15b)

$$m(b - a) = \int_a^b m dx \leq \int_a^b f(x) dx \leq \int_a^b M dx = M(b - a)$$

b) Sei $m = \min\{f(x) \mid x \in [a, b]\}$ und $M = \max\{f(x) \mid x \in [a, b]\}$. Wegen a) ist

$\int_a^b f(x) dx = (b - a) \cdot v$ für ein $v \in [m, M]$. Nach Zwischenwertsatz nimmt $f(x)$ jeden Wert zwischen m und M an. Also existiert ein $\tilde{x} \in [a, b]$ mit $f(\tilde{x}) = v$. □

Satz 8.17 (Erweiterter Mittelwertsatz der Integralrechnung). *Sei $p(x) \geq 0$ für $x \in [a, b]$ und $p(x)$ sowie $f(x) \cdot p(x)$ integrierbar auf $[a, b]$.*

Wenn $m \leq f(x) \leq M$ auf $[a, b]$ ist, dann gilt

$$m \int_a^b p(x) dx \leq \int_a^b p(x)f(x) dx \leq M \int_a^b p(x) dx$$

Beweis: Analog zu 8.16a). □

Satz 8.18. Sei $f(x)$ auf $[a, b]$ beschränkt und $a < c < b$. Es gilt:

$$f \in R[a, b] \Leftrightarrow f \in R[a, c] \text{ und } f(x) \in R[c, b] .$$

In diesem Fall ist

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx .$$

Beweis: Sei \tilde{Z}_n eine Folge von Zerlegungen von $[a, c]$ und \hat{Z}_n eine Folge von Zerlegungen von $[c, b]$ mit $|\tilde{Z}_n| \rightarrow 0$ und $|\hat{Z}_n| \rightarrow 0$ für $n \rightarrow \infty$. Dann ist $Z_n = \tilde{Z}_n + \hat{Z}_n$ für jedes n eine Zerlegung von $[a, b]$ und es gilt $|Z_n| \leq \max\{|\tilde{Z}_n|, |\hat{Z}_n|\} \rightarrow 0$ für $n \rightarrow \infty$. Für alle $n \geq 0$ ist

$$\left. \begin{array}{l} S(\tilde{Z}_n) + S(\hat{Z}_n) = S(Z_n) \\ \text{sowie } s(\tilde{Z}_n) + s(\hat{Z}_n) = s(Z_n) \end{array} \right\} - (*)$$

$$\underbrace{S(\tilde{Z}_n) - s(\tilde{Z}_n)}_{\tilde{\delta}_n} + \underbrace{S(\hat{Z}_n) - s(\hat{Z}_n)}_{\hat{\delta}_n} = \underbrace{S(Z_n) - s(Z_n)}_{=: \delta_n}$$

Gemäß Satz 8.9 ist

$$\begin{aligned} f \in R[a, b] &\iff \delta_n \rightarrow 0 \\ &\iff \tilde{\delta}_n \rightarrow 0 \text{ und } \hat{\delta}_n \rightarrow 0 \\ &\stackrel{\text{Satz 8.9}}{\iff} f \in R[a, c] \text{ und } f \in R[c, b]. \end{aligned}$$

Aus (*) und Grenzübergang folgt Teil 2 der Behauptung. □

Definition 8.19. Für $a < b$ und $f \in R[a, b]$ wird $\int_a^b f(x) dx = - \int_b^a f(x) dx$ sowie $\int_c^c f(x) dx = 0$ für $c \in [a, b]$ festgelegt.

Wegen Satz 8.18 und Definition 8.19 ist für beliebiges $x \in [a, b]$ das Integral $F(x) := \int_a^x f(x) dx$ für $f \in R[a, b]$ definiert.

Satz 8.20. Sei $f \in R[a, b]$ und $[c, d] \subseteq [a, b]$. Dann gilt

$$\int_c^d f(x) dx = F(d) - F(c) .$$

Beweis: Mit Satz 8.18 gilt

$$\begin{aligned} \int_a^c f(x) dx + \int_c^d f(x) dx &= \int_a^d f(x) dx \\ \iff \int_c^d f(x) dx &= \int_a^d f(x) dx - \int_a^c f(x) dx = F(d) - F(c) \end{aligned}$$

□

8.2 Der Zusammenhang zwischen Differential- und Integralrechnung

Die Erkenntnis, dass Integration auch als Umkehrung der Differentiation gesehen werden kann, entstand in der 2. Hälfte des 17. Jahrhunderts.

Satz 8.21. Sei $f \in R[a, b]$ und in $\tilde{x} \in [a, b]$ stetig. Dann ist für ein $c \in [a, b]$ die Funktion $F(x) = \int_c^x f(t) dt$ in \tilde{x} differenzierbar und $F'(\tilde{x}) = f(\tilde{x})$. Ist also f stetig auf $[a, b]$, so ist F stetig differenzierbar und $F'(x) = f(x)$ für alle $x \in [a, b]$.

Beweis: Wegen Satz 8.20 ist $F(\tilde{x}+h) - F(\tilde{x}) = \int_{\tilde{x}}^{\tilde{x}+h} f(t) dt$ und $\frac{1}{h} \int_{\tilde{x}}^{\tilde{x}+h} \underbrace{f(\tilde{x})}_{=const.} dt =$

$$\frac{1}{h} f(\tilde{x}) \int_{\tilde{x}}^{\tilde{x}+h} dt = \frac{1}{h} f(\tilde{x}) \cdot (\tilde{x} + h - \tilde{x}) = f(\tilde{x}).$$

Also ist

$$\begin{aligned} \left| \frac{F(\tilde{x}+h) - F(\tilde{x})}{h} - f(\tilde{x}) \right| &= \left| \frac{1}{h} \int_{\tilde{x}}^{\tilde{x}+h} f(t) dt - \frac{1}{h} \int_{\tilde{x}}^{\tilde{x}+h} f(\tilde{x}) dt \right| \\ &= \left| \frac{1}{h} \int_{\tilde{x}}^{\tilde{x}+h} (f(t) - f(\tilde{x})) dt \right| \\ &\stackrel{\text{Satz 8.15c}}{\leq} \frac{1}{h} \int_{\tilde{x}}^{\tilde{x}+h} \underbrace{|f(t) - f(\tilde{x})|}_{\leq \varepsilon (*)} dt \\ &\leq \frac{\varepsilon}{h} \int_{\tilde{x}}^{\tilde{x}+h} dt \\ &= \frac{\varepsilon}{h} (\tilde{x} + h - \tilde{x}) = \varepsilon \end{aligned}$$

ε kann beliebig klein gemacht werden, indem $|h| \rightarrow 0$ (dann $|h| < \delta$). (*) da f in \tilde{x} stetig! $\forall \varepsilon > 0 : \exists \delta > 0 : |f(x) - f(\tilde{x})| < \varepsilon$ für $|x - \tilde{x}| < \delta$. \square

Definition 8.22. Sei die Funktion $f(x)$ in $[a, b]$ erklärt. Eine Funktion $F(x)$ mit der Eigenschaft $F'(x) = f(x)$ auf $[a, b]$ soll Stammfunktion oder unbestimmtes Integral von $f(x)$ heißen. Wir schreiben dafür $F(x) = \int f(x) dx$.

Offensichtlich ist mit $F(x)$ auch $F(x) + c$ für beliebiges $c \in \mathbb{R}$ eine Stammfunktion von $f(x)$, da die Ableitung einer Konstanten stets Null ergibt. Die Kenntnis der Stammfunktion von f erleichtert die Berechnung des bestimmten Integrals ungemein.

Satz 8.23. Ist $F(x)$ auf $[a, b]$ stetig differenzierbar, so gilt

$$F(b) - F(a) = \int_a^b F'(t) dt,$$

und für $x, c \in [a, b]$ entsprechend

$$F(x) = F(c) + \int_c^x F'(t) dt .$$

Beweis: Sei $G(x) := \int_a^x F'(t) dt$ und F' auf $[a, b]$ stetig. Dann ist G nach Satz 8.21 stetig differenzierbar und es gilt $G'(x) = F'(x)$ für $x \in [a, b]$. Für die Funktion $H(x) := F(x) - G(x)$ ist also die Ableitung $H'(x) = F'(x) - G'(x) \equiv 0$ für $x \in [a, b]$. Folglich muss $H(x)$ konstant sein auf $[a, b]$. Da aber $H(a) = F(a) - \underbrace{G(a)}_{=0, \text{ Def. 8.19}} = F(a)$ folgt somit $H(x) \equiv F(x)$ für $x \in [a, b]$. Also gilt für $x = b$ gerade $H(b) = F(b) - G(b)$ und $H(b) = F(a)$. Deshalb $F(b) - G(b) = F(a) \iff G(b) = F(b) - F(a) = \int_c^b F'(t) dt$. \square

Ist also $f \in R[a, b]$ und kennen wir die zugehörige Stammfunktion, so lässt sich das Integral ohne den langwierigen Riemannsches Grenzprozess bestimmen!

8.3 Die Technik des Integrierens

Die Technik des Integrierens basiert auf

- einer Liste von Stammfunktionen,
- der Methode der partiellen Integration,
- der Substitutionsregel und
- einem Fundus von Umformungen, so dass die vorgenannten Ansätze greifen.

Definition 8.24 (Notation).

$$F(x) \Big|_b^a := \left[F(x) \right]_b^a := F(b) - F(a).$$

Satz 8.25 (Partielle Integration). *Seien $f(x)$ und $g(x)$ auf $[a, b]$ stetig differenzierbar. Dann gilt*

$$\int_b^a f(x) \cdot g'(x) dx = f(x) \cdot g(x) \Big|_a^b - \int_b^a f'(x) \cdot g(x) dx .$$

Beweis: Die Produktregel der Differenziation lautet

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x) .$$

Integration auf beiden Seiten ergibt

$$\underbrace{\int_b^a (f(x) \cdot g(x))' dx}_{= f(x) \cdot g(x) \Big|_a^b} = \underbrace{\int_b^a f'(x) \cdot g(x) dx}_{(*)} + \int_b^a f(x) \cdot g'(x) dx$$

$f(x) = F'(x)$	$F(x) = \int F'(x) dx$	Definitionsbereich von f
$c (c \in \mathbb{R})$	cx	\mathbb{R}
x^α	$\frac{1}{\alpha+1} x^{\alpha+1}$	\mathbb{R} für $\alpha \in \mathbb{N}$ $\mathbb{R} \setminus \{0\}$ für $\alpha \in \mathbb{Z} \setminus \{-1\}$, $x > 0$ für $\alpha \neq 1$
$\frac{1}{x}$	$\log x $	$\mathbb{R} \setminus \{0\}$
e^x	e^x	\mathbb{R}
$a^x (a > 0, a \neq 1)$	$\frac{a^x}{\log a}$	\mathbb{R}
$\log x $	$x \cdot \log x - x$	$\mathbb{R} \setminus \{0\}$
$\frac{1}{1+x^2}$	$\arctan x$	\mathbb{R}
$\frac{1}{1-x^2}$	$\frac{1}{2} \log \left \frac{x+1}{x-1} \right $	$\mathbb{R} \setminus \{-1, 1\}$
$\frac{1}{\sqrt{1+x^2}}$	$\operatorname{Arsinh} x$	\mathbb{R}
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$	$(-1, 1)$
$\sin x$	$-\cos x$	\mathbb{R}
$\cos x$	$\sin x$	\mathbb{R}
$\tan x$	$-\log \cos x $	$\mathbb{R} \setminus \{x = \frac{\pi}{2} + k \cdot \pi \mid k \in \mathbb{Z}\}$

Tabelle 8.1: Funktionen und ihre Stammfunktionen (Auswahl)

Subtraktion von (*) auf beiden Seiten liefert die Behauptung □

Anmerkung: Die Formel für partielle Integration gilt auch für unbestimmte Integrale

$$\int f(x) \cdot g'(x) dx = f(x) \cdot g(x) - \int f'(x) \cdot g(x) dx,$$

wie man leicht durch Differentiation auf beiden Seiten verifiziert.

Entscheidend für eine erfolgreiche Anwendung der partiellen Integration ist die geeignete Wahl von f und g' , so dass sowohl f als auch $\int f'g dx$ auf der rechten Seite leicht gefunden werden können.

Beispiel 8.3.

$$a) \int \underbrace{x}_f \underbrace{e^x}_{g'} dx = \underbrace{x}_f \underbrace{e^x}_g - \int \underbrace{1}_{f'} \underbrace{e^x}_g dx = xe^x - e^x = (x-1)e^x$$

$$b) \int \underbrace{x^2}_f \underbrace{e^x}_{g'} dx = \underbrace{x^2}_f \underbrace{e^x}_g - \int \underbrace{2x}_{f'} \underbrace{e^x}_g dx = x^2e^x - 2(x-1)e^x = (x^2 - 2x + 2)e^x$$

$$c) \int \log x dx = \int \underbrace{\log x}_f \underbrace{1}_{g'} dx = \underbrace{\log x}_f \underbrace{x}_g - \int \underbrace{\frac{1}{x}}_{f'} \underbrace{x}_g dx = x \log x - x, x > 0$$

$$d) \int \sin^2 x dx = \int \underbrace{(\sin x)}_f \underbrace{(\sin x)}_{g'} dx = \underbrace{(\sin x)}_f \underbrace{(-\cos x)}_g - \int \underbrace{(\cos x)}_{f'} \underbrace{(-\cos x)}_g dx$$

$$= -\sin x \cos x + \int \underbrace{\cos^2 x}_{1-\sin^2 x} dx = -\sin x \cos x + \int \underbrace{1}_{=x} dx - \int \sin^2 x dx$$

$$\Rightarrow \int \sin^2 x dx = \frac{1}{2}(x - \sin x \cos x)$$

In der Regel setzt man $f(x)$ gleich dem Potenzfaktor, damit er durch mehrfache partielle Integration schließlich verschwindet. Manchmal führt aber auch $g'(x) = 1 \Rightarrow g(x) = x$ zum Ziel (wie in (c)).

Satz 8.26 (Substitutionsregel). Sei f stetig auf $\langle a, b \rangle$ und g stetig differenzierbar auf $\langle \alpha, \beta \rangle$, wobei $\langle a, b \rangle := [\min\{a, b\}, \max\{a, b\}]$ und

- $g(\langle \alpha, \beta \rangle) \subseteq \langle a, b \rangle$ sowie
- $g(\alpha) = a, g(\beta) = b$.

Dann ist $\int_a^b f(x) dx = \int_\alpha^\beta f(g(t))g'(t) dt$.

Beweis: Die Integrale existieren, weil f und g' sowie $f(g(t))$ stetig sind. Sei $F'(x) = f(x)$. Nach der Kettenregel der Differenzialrechnung gilt

$$\frac{d}{dt}F(g(t)) = F'(g(t))g'(t) = f(g(t))g'(t).$$

Integration auf beiden Seiten bezüglich t ergibt

$$F(g(t)) \Big|_\alpha^\beta = \int_\alpha^\beta f(g(t))g'(t) dt.$$

Da $F(g(t)) \Big|_\alpha^\beta = F(g(\beta)) - F(g(\alpha)) = F(b) - F(a) = \int_a^b F'(x) dx = \int_a^b f(x) dx$ folgt damit die Behauptung. \square

Anmerkung: Sei f stetig und g stetig differenzierbar. Dann gilt $\int f(g(x))g'(x) dx = F(g(x))$ wenn $F'(x) = f(x)$, also F Stammfunktion von f ist. Beweis durch Differenzieren auf beiden Seiten.

Entscheidend für eine erfolgreiche Anwendung der Substitutionsregel ist die geeignete Wahl der Substitution. Die eine führt schnell zum Ziel, eine andere nur weiter ins Dickicht. Hier ist Übung, Erfahrung, Intuition und auch Glück erforderlich, um den zielführenden Ansatz zu wählen.

Beispiel 8.4.

- Berechne $\int_1^2 (2t - 2)^9 dt$.
 $f(x) = x^9; g(t) = 2t - 2; g'(t)dt = 2dt = dx; \alpha = 1; \beta = 2$
 $g(\alpha) = g(1) = 0 = a; g(\beta) = g(2) = 2 = b$
 $\Rightarrow F(x) = \frac{1}{10}x^{10}$

Also:

$$\begin{aligned} \int_1^2 (2t - 2)^9 dt &= \frac{1}{2} \int_1^2 f(g(t))g'(t) dt = \frac{1}{2} F(x) \Big|_{a=0}^{b=2} \\ &= \frac{1}{2} \frac{1}{10} (2^{10} - 0) = \frac{512}{10} \end{aligned}$$

- Berechne $\int_0^1 \frac{1}{1+4t} dt$
 $f(x) = \frac{1}{x}; g(t) = 1 + 4t; g'(t)dt = 4dt = dx; \alpha = 0; \beta = 1$

$$g(\alpha) = g(0) = 1 = a; \quad g(\beta) = g(1) = 5 = b \\ \Rightarrow F(x) = \log |x|.$$

Also:

$$\int_0^1 \frac{1}{1+4t} dt = \frac{1}{4} \int_0^1 f(g(t))g'(t) dt = \frac{1}{4} F(x) \Big|_{a=1}^{b=5} = \frac{1}{4} (\log 5 - \underbrace{\log 1}_{=0}) = \frac{1}{4} \log 5.$$

Für das unbestimmte Integral ergibt sich

$$\int \frac{1}{1+4t} dt = \frac{1}{4} \int f(g(t))g'(t) dt = \frac{1}{4} F(g(t)) = \frac{1}{4} \log |1+4t|.$$

Wir betrachten einmal $\int_1^2 (2t-2)^9 dt$:

Man könnte auch sagen, dass wir den Ausdruck $2t-2$ durch eine neue Variable x ersetzen (substituieren!); also

$$x = 2t - 2.$$

Das Differential dx erhalten wir aus dem Produkt der Ableitung des substituierten Ausdrucks $(2t-2)'$ und dem Differential dt ; also

$$dx = (2t-2)' dt = 2dt \Leftrightarrow dt = \frac{1}{2} dx.$$

Die neuen Intervallgrenzen ergeben sich durch Einsetzen der ursprünglichen Intervallgrenzen in den substituierten Ausdruck; also $2t-2|_{t=1} = 0$ und $2t-2|_{t=2} = 2$. Insgesamt ergibt sich also durch Substitution $\int_1^2 (2t-2)^9 dt = \frac{1}{2} \int_0^2 x^9 dx$, was wir sofort mit Hilfe der Stammfunktionstabelle lösen können. Für das unbestimmte Integral machen wir den gleichen Ansatz, müssen aber schließlich rücksubstituieren ($x = 2t-2$), also

$$\int (2t-2)^9 dt = \frac{1}{2} \int x^9 dx = \frac{1}{20} x^{10} = \frac{1}{20} (2t-2)^{10}.$$

$$\text{Probe: } \left(\frac{1}{20} (2t-2)^{10}\right)' = \frac{1}{20} 10(2t-2)^9 \cdot 2 = (2t-2)^9 \checkmark$$

Satz 8.27. *Nützliche Substitutionsregeln:*

1. $\int f(ax+b) dx = \frac{1}{a} F(ax+b)$ für $F'(x) = f(x)$.
2. $\int f(x)f'(x) dx = \frac{1}{2} f^2(x)$.
3. $\int \frac{f'(x)}{f(x)} dx = \log |f(x)|$.

Beweis: 1) bis 3) lassen sich leicht durch Ableiten verifizieren. Wir wollen aber konstruktive Beweise geben:

$$1. \int f(ax+b) dx \underset{\substack{t=ax+b \\ dt=adx \\ \Leftrightarrow dx=\frac{1}{a}dt}}{=} \frac{1}{a} \int f(x) dt \underset{f=F'}{=} \frac{1}{a} F(t) = \frac{1}{a} F(ax+b)$$

$$2. \int f(x)f'(x) dx \underset{\substack{t=f(x) \\ dt=f'(x)dx}}{=} \int t dt = \frac{1}{2} t^2 = \frac{1}{2} f^2(x)$$

$$3. \int \frac{f'(x)}{f(x)} dx \underset{\substack{t=f(x) \\ dt=f'(x)dx}}{=} \int \frac{1}{t} dt = \log |t| = \log |f(x)|$$

□

Beispiel 8.5 (zu Satz 8.27).

$$1. \int \log(2x) dx = \frac{1}{2}[2x(\log(2x) - 1)] = x(\log(2x) - 1) \quad \text{für } F'(x) = \log x.$$

$$2. \int \underbrace{\sin x}_{f(x)} \underbrace{\cos x}_{f'(x)} dx = \frac{1}{2} \sin^2 x$$

$$3. \int \frac{dx}{x \log x} dx = \int \frac{\frac{1}{x}}{\log x} dx = \log(\log x), \quad x > 1$$

Durch Substitution lassen sich auch kompliziert anmutende Ausdrücke manchmal in Form von rationalen Funktionen darstellen. Ohne Beweis bemerken wir:

Satz 8.28. *Jede rationale Funktion ist elementar integrierbar.*

Beweis: Siehe z. B. [3, S. 278].

□

$$\text{Beispiel 8.6. } \int \frac{e^{3x}+3}{e^x+1} dx \underset{\substack{t=e^x \\ dt=e^x dx = t dx \\ \Leftrightarrow \frac{dt}{t} = dx}}{=} \int \underbrace{\frac{t^3+3}{t(t+1)}}_{\text{rationale Fkt.}} dt$$

$$\begin{aligned} \text{Polynomdivision: } (t^3 + 3) : (t^2 + t) &= t - 1 + (t + 3)/(t(t + 1)) \\ &\quad \underline{-(t^3 + t^2)} \\ &\quad \quad -t^2 + 3 \\ &\quad \quad \underline{-(-t^2 - t)} \\ &\quad \quad \quad t + 3 \end{aligned}$$

$\frac{t+3}{t(t+1)}$ ist echt gebrochen rational (Polynomgrad im Zähler < Polynomgrad im Nenner).

Partialbruchzerlegung:

$$\frac{t+3}{t(t+1)} = \frac{A}{t} + \frac{B}{t+1} = \frac{A(t+1) + Bt}{t(t+1)} = \frac{\overset{\substack{\stackrel{!}{=}1 \Rightarrow B=-2}{(A+B)}}{t} + \frac{\overset{\substack{!}{=}3}{A}}{t+1}$$

$$\begin{aligned} \int \frac{t^3+3}{t(t+1)} dt &= \int (t+1) dt + \int \frac{3}{t} dt - \int \frac{2}{t+1} dt \\ &= \frac{1}{2}t^2 + t + 3 \ln |t| - 2 \ln |t+1| \\ &\stackrel{\text{Rücksubst.}}{=} \frac{1}{2}e^{2x} + e^x + 3x - 2 \ln(e^x + 1) \end{aligned}$$

8.4 Uneigentliche Integrale

Voraussetzungen für das Riemann-Integral sind

- die Beschränktheit der zu integrierenden Funktion und
- die Beschränktheit des Integrationsintervalls.

Der Riemansche Integralbegriff soll nun ausgedehnt werden auf solche Fälle, bei denen man auf diese Einschränkungen verzichten kann.

Definition 8.29 (unbeschränkter Integrationsbereich). *Sei f auf $[a, \infty)$ erklärt und über jedes $[a, c]$ für $a < c < \infty$ integrierbar.*

Man legt fest:

$$\int_a^\infty f(x) dx := \lim_{c \rightarrow \infty} \int_a^c f(x) dx.$$

Wenn der Limes existiert, dann existiert das uneigentliche Integral und es wird konvergent genannt (andernfalls divergent). Entsprechend definieren wir für ein Intervall $(-\infty, a]$

$$\int_{-\infty}^a f(x) dx = \lim_{c \rightarrow -\infty} \int_c^a f(x) dx$$

und für ganz $\mathbb{R} = (-\infty, \infty)$

$$\int_{-\infty}^\infty f(x) dx := \int_{-\infty}^a f(x) dx + \int_a^\infty f(x) dx$$

für ein beliebiges $a \in \mathbb{R}$, wobei beide Integrale auf der rechten Seite existieren müssen.

Beispiel 8.7.

$$1. \int_0^\infty e^{-x} dx = \lim_{c \rightarrow \infty} \int_0^c e^{-x} dx = \lim_{c \rightarrow \infty} [-e^{-x}]_0^c = \lim_{c \rightarrow \infty} (1 - e^{-c}) = 1 \text{ (konvergent)}$$

$$2. \int_1^\infty \frac{1}{x^\alpha} dx = \lim_{c \rightarrow \infty} \int_1^c \frac{1}{x^\alpha} dx = \lim_{c \rightarrow \infty} \left[\frac{1}{1-\alpha} x^{1-\alpha} \right]_1^c = \lim_{c \rightarrow \infty} \frac{c^{1-\alpha} - 1}{1-\alpha}$$

$$= \begin{cases} \frac{1}{\alpha-1} & \text{für } \alpha > 1 \text{ (konvergent)} \\ \infty & \text{für } \alpha < 1 \text{ (divergent)} \end{cases}$$

$$3. \int_1^\infty \frac{1}{x} dx = \lim_{c \rightarrow \infty} \int_1^c \frac{1}{x} dx = \lim_{c \rightarrow \infty} [\log x]_1^c = \lim_{c \rightarrow \infty} \log c = \infty \text{ (divergent)}.$$

$$4. \int_{-\infty}^\infty \frac{dx}{1+x^2} = \lim_{c \rightarrow -\infty} \int_c^0 \frac{dx}{1+x^2} + \lim_{c \rightarrow \infty} \int_0^c \frac{dx}{1+x^2} = \lim_{c \rightarrow -\infty} \arctan x \Big|_c^0 + \lim_{c \rightarrow \infty} \arctan x \Big|_0^c =$$

$$- \lim_{c \rightarrow -\infty} \arctan c + \lim_{c \rightarrow \infty} \arctan c = -\left(-\frac{\pi}{2}\right) + \frac{\pi}{2} = \pi \text{ (konvergent)}.$$

$$5. \int_0^\infty \sin x dx = \lim_{c \rightarrow \infty} \int_0^c \sin x dx = \lim_{c \rightarrow \infty} -\cos x \Big|_0^c = \lim_{c \rightarrow \infty} (-\cos(c) + 1) \text{ (Limes existiert nicht } \Rightarrow \text{divergent)}.$$

Achtung: Im Allgemeinen gilt $\int_{-\infty}^\infty f(x) dx \neq \lim_{c \rightarrow \infty} \int_{-c}^c f(x) dx$, weil z. B. für $\int_{-\infty}^\infty x dx$ der Grenzwert $\lim_{c \rightarrow \infty} \int_{-c}^c x dx = \lim_{c \rightarrow \infty} \frac{1}{2} x^2 \Big|_{-c}^c = 0$, aber $\int_{-\infty}^\infty x dx = \int_{-\infty}^0 x dx + \int_0^\infty x dx$ existiert nicht, da jedes Teilintegral divergent.

Definition 8.30 (unbeschränkte Funktionen).

1. Sei $f \in R[c, b]$ für jedes c mit $a < c < b$ und $\lim_{x \rightarrow a^+} |f(x)| = \infty$. Man definiert $\int_a^b f(x) dx := \lim_{c \rightarrow a^+} \int_c^b f(x) dx$, falls der Limes existiert.
2. Sei $f \in R[a, c]$ für jedes c mit $a < c < b$ und $\lim_{x \rightarrow b^-} |f(x)| = \infty$. Man definiert $\int_a^b f(x) dx := \lim_{c \rightarrow b^-} \int_a^c f(x) dx$, falls der Limes existiert.
3. Sei $f \in R[c, d]$ für alle c, d mit $a < c < d < b$ und $\lim_{x \rightarrow a^+} |f(x)| = \infty$ sowie $\lim_{x \rightarrow b^-} |f(x)| = \infty$. Man definiert unter Rückführung auf die Fälle 1.) und 2.) $\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx$, ($a < c < b$) falls beide uneigentlichen Integrale auf der rechten Seite existieren.

Beispiel 8.8.

1. $\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \lim_{c \rightarrow 1^-} \int_0^c \frac{dx}{\sqrt{1-x^2}} = \lim_{c \rightarrow 1^-} [\arcsin x]_0^c = \lim_{c \rightarrow 1^-} \arcsin c = \frac{\pi}{2}$
2. $\int_0^1 \log x dx = \lim_{c \rightarrow 0^+} \int_c^1 \log x dx = \lim_{c \rightarrow 0^+} [x \log x - x]_c^1 = -1 - \lim_{c \rightarrow 0^+} c \log c = -1$
 da $\lim_{c \rightarrow 0^+} c \log c = \lim_{c \rightarrow 0^+} c \rightarrow 0^+ \frac{\log c}{c^{-1}} \stackrel{l'Hospital}{=} \lim_{c \rightarrow 0^+} \frac{c^{-1}}{-c^{-2}} = - \lim_{c \rightarrow 0^+} c = 0$
3. wie 1.) mit $[-1, 1]$

Anmerkung: Falls sowohl Integrationsbereich als auch zu integrierende Funktion unbeschränkt sind, dann spaltet man in mehrere uneigentliche Integrale auf und fordert, dass *alle* uneigentlichen Teilintegrale existieren müssen.

Beispiel 8.9. Sei $f(x) = \min\{\log x, \frac{\log 16}{x^2}\}$

$$= \begin{cases} \log x, & \text{für } x \leq 2 \\ \frac{\log 16}{x^2}, & \text{für } x > 2 \end{cases} \quad \text{mit } x > 0 \text{ (stetig!)}$$

$$\begin{aligned} \int_0^\infty f(x) dx &= \int_0^2 f(x) dx + \int_2^\infty f(x) dx = \lim_{a \rightarrow 0^+} \int_a^2 \log x dx + \lim_{b \rightarrow \infty} \int_2^b \frac{\log 16}{x^2} dx = \\ & \lim_{a \rightarrow 0^+} [x \log x - x]_a^2 + \lim_{b \rightarrow \infty} [-\frac{\log 16}{x}]_2^b \\ & 2 \log 2 - 2 - \underbrace{\lim_{a \rightarrow 0^+} a \log a}_{=0} + \underbrace{(-\lim_{b \rightarrow \infty} \frac{\log 16}{b} + \frac{\log 16}{2})}_{=0} = 4 \log 2 - 2, \text{ da } \log 16 = \log 2^4 = \\ & 4 \log 2. \end{aligned}$$

8.5 Anwendungen

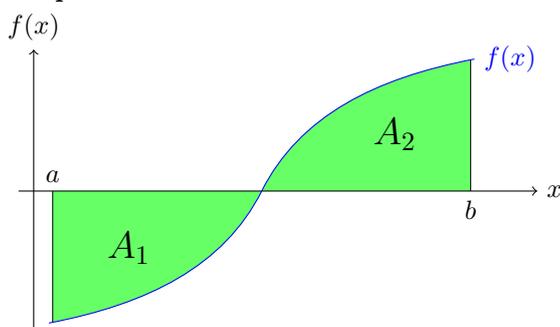
Prinzipiell ergeben sich die Anwendungen der Integralrechnung aus drei Sichtweisen:

1. Integrieren heißt Summieren (\rightarrow Flächenberechnung)
2. Integrieren heißt Mitteln (Mittelwertsatz)
3. Integrieren heißt Rekonstruktion (von Funktionen aus ihrer Änderungsrate).

8.5.1 Summieren

Flächenberechnung als Grenzwert der Summe infinitesimal schmaler Rechtecke. Aber im Allgemeinen ist der Wert des bestimmten Integrals ungleich der Flächen zwischen x -Achse und Graph der Funktion.

Beispiel 8.10.



A_1 : Fläche zwischen $f(x)$, x -Achse und Gerade $x = a$
 A_2 : Fläche zwischen $f(x)$, x -Achse und Gerade $x = b$

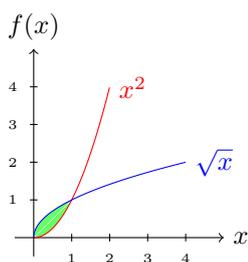
$$\int_a^b f(x) dx \stackrel{?}{=} \begin{cases} A_1 + A_2 & = \int_a^b |f(x)| dx \\ A_1 - A_2 \\ A_2 - A_1 & \checkmark \\ |A_1 - A_2| \\ \frac{1}{2}(A_1 + A_2) \end{cases}$$

\Rightarrow Das Integral bilanziert Flächen!

Die analytische Bestimmung von gänzlich nichtlinear berandeten Flächen ist möglich:

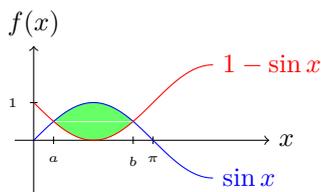
Beispiel 8.11.

- Gesucht ist der Flächeninhalt zwischen \sqrt{x} und x^2 im Bereich $[0, 1]$



$$\int_0^1 \sqrt{x} dx - \int_0^1 x^2 dx = \frac{2}{3}x^{\frac{3}{2}} \Big|_0^1 = \frac{1}{3}$$

- Gesucht ist der Flächeninhalt zwischen $\sin x$ und $1 - \sin x$ im Bereich $[a, b]$



$$\sin x \stackrel{!}{=} 1 - \sin x \Leftrightarrow \sin x = \frac{1}{2}$$

$$\Rightarrow a = x_1 = \frac{\pi}{6} \text{ und } b = x_2 = \frac{5}{6}\pi$$

$$\int_a^b \sin x \, dx - \int_a^b (1 - \sin x) \, dx = 2 \int_a^b \sin x \, dx - \int_a^b dx = -2 \cos x \Big|_a^b - (b - a)$$

$$= -2 \left(\cos \frac{5}{6}\pi - \cos \frac{1}{6}\pi \right) - \frac{2}{3}\pi = -2 \left(-\frac{\sqrt{3}}{2} - \frac{\sqrt{3}}{2} \right) - \frac{2}{3}\pi = 2\sqrt{3} - \frac{2}{3}\pi \approx 1,3$$

Die Bestimmung eines Integrals in geschlossener Form (also Darstellung mit Hilfe von elementaren Funktionen) ist nicht immer möglich.

Beispiel 8.12.

1. Integralsinus $\text{Si}(x) := \int_0^x \frac{\sin(t)}{t} \, dt$
2. Fehlerfunktion $\text{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt$
3. $\int \sqrt{x^3 + 1} \, dx$ oder $\int (x^2 + 1)^{\frac{1}{3}} \, dx$

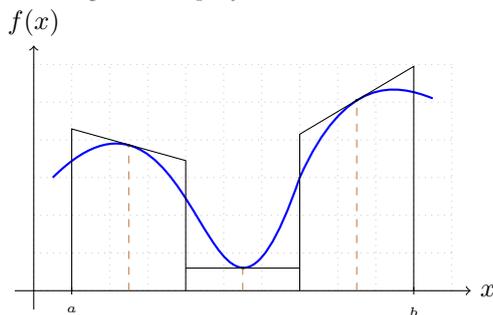
In solchen Fällen kommt die numerische Integration zum Einsatz.

Ansatz: Riemannsche Zwischensummen

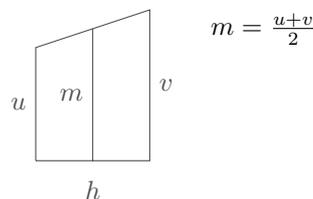
$$\sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\tilde{x}_i) \text{ für } \tilde{x}_i \in [x_{i-1}, x_i] \text{ „passend“}$$

z. B. $\tilde{x}_i = x_{i-1} + \frac{x_i - x_{i-1}}{2} = \frac{x_{i-1} + x_i}{2}$ (Intervallmitte)

→ Tangententrapezformel

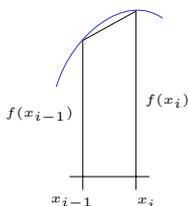


Flächeninhalt eines Trapezes
= $m \cdot h$



Satz 8.13 sichert Konvergenz zum Integralwert für $n \rightarrow \infty$. Je größer n , desto besser die Approximation.

Auch implizite Wahl von \tilde{x}_i möglich → Sehnentrapezregel.

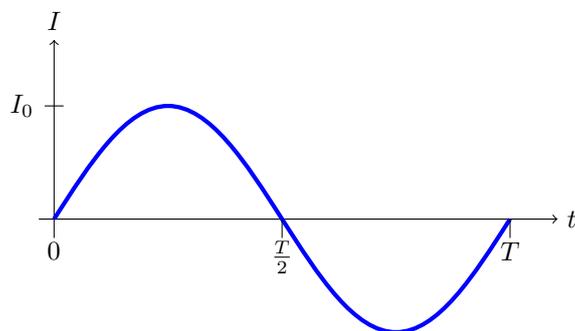


$\frac{f(x_{i+1}) + f(x_i)}{2}$ liegt zwischen $f(x_{i+1})$ und $f(x_i)$. Falls f stetig ist, sichert der Zwischenwertsatz die Existenz eines \tilde{x}_i mit $f(\tilde{x}_i) = \frac{1}{2} (f(x_{i+1}) + f(x_i))$

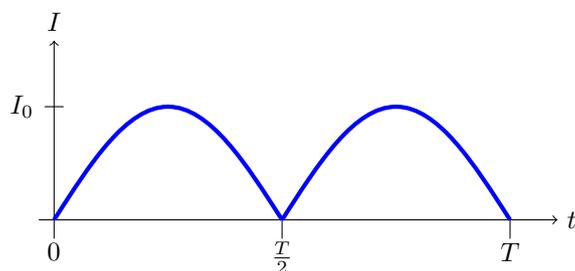
8.5.2 Mitteln

Mitteln mit Hilfe des Mittelwertsatzes der Integralrechnung.

Beispiel 8.13. Zweiweggleichrichter erzeugt aus Sinuswechselstrom
 $I(t) = I_0 \cdot \sin(\omega t)$, $\omega = \frac{2\pi}{T}$



diesen Verlauf



Wie groß ist der (lineare) Mittelwert des Stroms?

$$\begin{aligned}
 \frac{1}{\frac{T}{2} - 0} \int_0^{T/2} I(t) dt &= \frac{2}{T} \cdot I_0 \int_0^{T/2} \sin(\omega t) dt \\
 &= \frac{2 \cdot I_0}{T} \cdot \frac{1}{\omega} \left[-\cos(\omega t) \right]_0^{T/2} \\
 &= \frac{2 \cdot I_0}{T} \cdot \frac{T}{2\pi} \left(\underbrace{-\cos\left(\frac{2\pi}{T} \cdot \frac{T}{2}\right)}_{\cos \pi = -1} + \underbrace{\cos(0)}_0 \right) = \frac{2}{\pi} \cdot I_0
 \end{aligned}$$

8.5.3 Rekonstruieren

Rekonstruieren mit dem zweiten Hauptsatz (Satz 8.23)

$$F(x) = F(a) + \int_a^x F'(t) dt \text{ für } x \in [a, b] \text{ für stetiges } F'$$

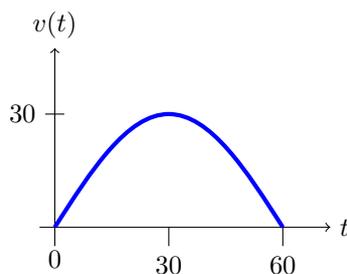
Man kann aus der Änderungsrate der Funktion die Funktion selbst rekonstruieren.

Beispiel 8.14. Fahrtenschreiber: *Geschwindigkeit im Zeitverlauf* $v(t)$. *Rekonstruktion des zurückgelegten Weges aus Geschwindigkeit als momentane Änderungsrate des Weges:* $x(t)$ sei gefahrene Strecke zum Zeitpunkt t . Dann ist

$$\frac{x(t+h) - x(t)}{(t+h) - t}$$

die mittlere Geschwindigkeit in h Zeiteinheiten. Grenzprozess $h \rightarrow 0$ ergibt $x'(t)$ als momentane Geschwindigkeit $v(t)$. Wir kennen $v(t) = x'(t)$ vom Fahrtenschreiber und wollen daraus die Funktion $x(t)$ rekonstruieren. Offensichtlich ist $x(0) = 0$. Also $x(t) = x(0) + \int_0^t x'(s) ds$.

Angenommen, der Fahrtenschreiber liefert die Daten für das Diagramm



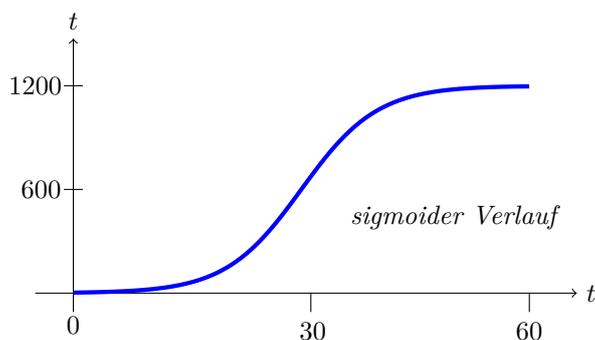
mit $v(t) = \frac{1}{30}t(60 - t)$ für $t \in [0, 60]$ mit Einheit $[\text{m s}^{-1}]$. Dann ergibt sich

$$x(t) = 0 + \int_0^t \frac{1}{30}s(60 - s) ds \quad (8.5.1)$$

$$= \frac{1}{30} \int_0^t \frac{1}{30}s(60 - s^2) ds \quad (8.5.2)$$

$$= \frac{1}{30} \left[30s^2 - \frac{1}{3}s^3 \right]_0^t = t^2 - \frac{1}{90}t^3. \quad (8.5.3)$$

Für $t = 30$ ergibt sich $x(30) = 900 - \frac{1}{90} \cdot 30 \cdot 900 = 600 [\text{m}]$.



Weiteres Beispiel zur Rekonstruktion:

Ausbreitungsgeschwindigkeit einer Epidemie $\hat{=}$ momentane Änderungsrate der Zahl der Infizierten

Sei $x(t)$ die Anzahl der Infizierten zum Zeitpunkt $t \geq 0$ und $x(0) = 1$. Die Änderungsrate beschreibt den Zuwachs an Infizierten; er wird proportional zur Anzahl der Infizierten angenommen

$$\underbrace{x'(t)} = c \cdot x(t) \quad , \text{ für } c > 0.$$

Welche Funktionen erfüllen die Gleichung überhaupt?

Wir raten: $x(t) = e^{c \cdot t}$ mit $x'(t) = c \cdot e^{c \cdot t}$. Eigentlich sind wir schon fertig: $x(t) = e^{ct}$ ist die Lösung!
 „Probe“ mit Satz 8.23:

$$x(t) = x(0) + \int_0^t x'(s) \, ds = 1 + c \cdot \int_0^t e^{cs} \, ds = 1 + c \cdot \frac{1}{c} \left[e^{cs} \right]_0^t = 1 + (e^{ct} - 1) = e^{ct}$$

Stimmt!

Wie kommt man ohne „gutes Raten“ aus? Theorie der \rightarrow Differentialgleichungen!

8.6 Exkurs: Differentialgleichungen

Da in Anwendungen die unabhängige Variable meist die Zeit (time) ist, verwendet man als Argument meist t statt x . Ableitungen oder Funktionen $x(t), y(t), \dots$ werden auch in Newtonscher Schreibweise mit $\dot{x}(t) := \frac{dx(t)}{dt}$, $\ddot{x}(t) := \frac{d^2x(t)}{dt^2}$ und kürzer mit $\dot{x} := \frac{dx}{dt}$ bzw. $\ddot{x} := \frac{d^2x}{dt^2}$ bezeichnet.

Definition 8.31. Eine Gleichung der Form $f(t, x, \dot{x}, \ddot{x}, \dots) = 0$ heißt gewöhnliche Differentialgleichung (DGL). Die Ordnung der höchsten auftretenden Ableitung heißt die Ordnung der DGL.

Anmerkung: Treten in der Gleichung auch sogenannte *partielle Ableitungen* (\rightarrow nächstes Kapitel) auf, so spricht man von *partiellen DGL*. Hier behandeln wir nur gewöhnliche DGL!

Zurück zum Beispiel der Ausbreitung einer Epidemie:

Die Änderungsrate der Infizierten war $x'(t) = c \cdot x(t)$ für $c > 0$ (bzw. $\dot{x} = c \cdot x$)
 Gesucht ist $x(t)$ für $x(0) = 1$, und $t \geq 0$.

$$\begin{aligned} \frac{dx(t)}{dt} = c \cdot x(t) \quad \text{bzw. kürzer} \quad \frac{dx}{dt} = c \cdot dt \quad \Big| \cdot \frac{dt}{x} \\ \Leftrightarrow \frac{dx}{x} = c \cdot dt \quad \Big| \text{Integration} \quad \int_{x_0}^x \frac{d\tilde{x}}{\tilde{x}} = c \int_{t_0}^t d\tilde{t} \\ \log \tilde{x} \Big|_{x_0}^x = c \cdot \tilde{t} \Big|_{t_0}^t \\ \Leftrightarrow \log \frac{x}{x_0} = c(t - t_0) \quad \Big| e^{(\cdot)} \end{aligned}$$

$$\begin{aligned} x &= x_0 \cdot e^{c \cdot (t-t_0)} \\ x(t) &= x(0) \cdot e^{c \cdot (t-t_0)}, \quad x(0) = 1 \text{ Anfangsbedingung} \\ \text{Lösung: } x(t) &= e^{c \cdot t} \end{aligned}$$

8.6.1 Lineare DGL 1. Ordnung

$x'(t) = a(t) \cdot x(t) + b(t)$ wobei $a(t)$ und $b(t)$ gegebene stetige Funktion sind. Falls $b(t) \equiv 0$, dann heißt die DGL *homogen*, sonst *inhomogen*.

• Lösung homogener DGL durch *Trennung der Variablen*:

$$\begin{aligned} x'(t) = a(t) \cdot x(t) \quad \Leftrightarrow \quad \frac{dx}{dt} = a(t) \cdot x(t) \quad \Big| \cdot \frac{dt}{x(t)} \\ \Rightarrow \frac{dx}{x(t)} = a(t) \cdot dt \end{aligned}$$

Integration auf beiden Seiten liefert

$$\int_{x_0}^x \frac{d\tilde{x}}{\tilde{x}} = \int_{t_0}^t a(\tilde{t}) d\tilde{t} \quad \Leftrightarrow \log \tilde{x} \Big|_{x_0}^x = \int_{t_0}^t a(\tilde{t}) d\tilde{t} \quad \text{Sei } A'(t) = a(t)$$

$$\Leftrightarrow \log \frac{x}{x_0} = A(t) - A(t_0) \quad \Leftrightarrow x(t) = x_0 \cdot e^{A(t)-A(t_0)} \quad \text{bzw. } x_0 \cdot \exp\left(\int_{t_0}^t a(\tilde{t}) d\tilde{t}\right)$$

• Lösung inhomogener DGL durch *Variation der Konstanten*

$$x'(t) = a(t)t \cdot x(t) + b(t)$$

Idee: in homogener Lösung Konstante x_0 zur Funktion $x_0(t)$ machen und behaupten, dass $\underbrace{x(t) = x_0(t) \cdot \varphi(t)}_{(*)}$ mit

$$\varphi(t) \stackrel{\varphi'(t)=a(t) \cdot \varphi(t)}{=} \exp\left(\int_{t_0}^t a(\tilde{t}) d\tilde{t}\right) = e^{A(t)-A(t_0)}$$

Produktregel:

$$x'(t) = x'_0 \cdot \varphi(t) + x_0(t) \cdot \varphi'(t) = x'_0(t) \cdot \varphi(t) + \underbrace{x_0(t) \cdot a(t) \cdot \varphi(t)}_{(*)=a(t) \cdot x(t)} \stackrel{!}{=} a(t)x(t) + b(t)$$

$$\Rightarrow x'_0(t) \stackrel{!}{=} \frac{b(t)}{\varphi(t)} \quad \left| \text{ integrieren!} \right.$$

$$\Rightarrow \int_{t_0}^t x'_0(\tilde{t}) d\tilde{t} = x_0(t) - \underbrace{x_0(t_0)}_{=:c_0} \stackrel{!}{=} \int_{t_0}^t \frac{b(\tilde{t})}{\varphi(\tilde{t})} d\tilde{t} \quad \left| \text{ einsetzen in } (*) \right.$$

$$x(t) = \varphi(t) \cdot \left[c_0 + \int_{t_0}^t \frac{b(\tilde{t})}{\varphi(\tilde{t})} d\tilde{t} \right], \quad c_0 := x_0(t_0) \stackrel{!}{=} x_0$$

Beweis durch Ableiten:

$$x(t) = \varphi(t) \cdot \left[c_0 + \int_{t_0}^t \frac{b(s)}{\varphi(s)} ds \right] \quad (**)$$

$$x'(t) = c_0 \cdot \varphi'(t) + \varphi'(t) \cdot \int_{t_0}^t \frac{b(s)}{\varphi(s)} ds + \varphi(t) \cdot \frac{b(t)}{\varphi(t)}$$

$$= \varphi'(t) \cdot \left[c_0 + \int_{t_0}^t \frac{b(s)}{\varphi(s)} ds \right] + b(t) \quad \varphi'(t) = a(t) \cdot \varphi(t)$$

$$= a(t) \cdot \varphi(t) \cdot \underbrace{\left[c_0 + \int_{t_0}^t \frac{b(s)}{\varphi(s)} ds \right]}_{x(t) \text{ wegen } (**)} + b(t)$$

$$= a(t) \cdot x(t) + b(t)$$

$$x'(t) = \underbrace{2t}_{a(t)} \cdot x(t) + \underbrace{t^3}_{b(t)} \quad \text{mit } x(0) = x_0$$

1. Lösung der homogenen DGL

$$\varphi(t) = \exp\left(\int_0^t 2s ds\right) = \exp\left([s^2]_0^t\right) = e^{t^2}$$

2. Lösung der inhomogenen DGL

$$x(t) = \varphi(t) \cdot \left[x_0 + \int_{t_0}^t \frac{b(s)}{\varphi(s)} ds \right] = e^{t^2} \cdot \left[x_0 + \int_{t_0}^t \frac{s^3}{e^{s^2}} ds \right]$$

$$\begin{aligned} \int_{t_0}^t \frac{s^3}{e^{s^2}} ds &= \frac{1}{2} \cdot \int_{t_0}^t u \cdot e^{-u} du \text{ mit } u = s^2, du = 2s \cdot ds \Rightarrow u \cdot du = 2s^2 ds \\ &\stackrel{\text{part.}}{=} \frac{1}{2} [-ue^{-u} + \int e^{-u} du] = \frac{1}{2} [-ue^{-u} - e^{-u}] \\ &\stackrel{\text{Int.}}{=} -\frac{1}{2}e^{-u}(u+1) = -\frac{1}{2}e^{-s^2}(s^2+1) \end{aligned}$$

$$\begin{aligned} \Rightarrow \int_{t_0}^t \frac{s^3}{e^{s^2}} ds &= \left[-\frac{1}{2}e^{-s^2}(s^2+1) \right]_0^t = \frac{1}{2} - \frac{1}{2}e^{-t}(t^2+1) \\ \Rightarrow x(t) &= e^{t^2} \cdot \left[x_0 + \frac{1}{2} - \frac{1}{2}e^{-t}(t^2+1) \right] \end{aligned}$$

8.6.2 Nichtlineare DGL

$x'(t) = a(t)g(x(t))$, g stetig und nichtlinear.

Methode: Trennung der Variablen

Ansatz: $\frac{dx}{dt} = a(t)g(x) \mid \cdot \frac{dt}{g(x)}$

$\frac{dx}{g(x)} = a(t)dt \mid$ Integration auf beiden Seiten

$$G(x) := \int_{x_0}^x \frac{dy}{g(y)} = \int_{t_0}^t a(s) ds$$

Dann $G(x)$ nach x auflösen (falls es gelingt).

Beispiel 8.15. $x'(t) = e^{x(t)} \cdot \cos(t)$ mit $x(0) = x_0$ und $t_0 = 0$

$\frac{dx}{dt} = \underbrace{\cos(t)}_{a(t)} \cdot \underbrace{e^x}_{g(x)}$, also $g(x) = e^x$

$$\begin{aligned} \int_{x_0}^x \frac{dy}{e^y} &= \int_0^t \cos(s) ds \\ \Leftrightarrow [-e^{-y}]_{x_0}^x &= [\sin(s)]_0^t \\ \Leftrightarrow -e^{-x} + e^{-x_0} &= \sin(t) - \underbrace{\sin(0)}_{=0} \\ \Leftrightarrow e^{-x} &= e^{-x_0} - \sin(t) \quad | \ln \\ \Leftrightarrow -x &= \ln(e^{-x_0} - \sin(t)) \\ \Leftrightarrow x(t) &= -\ln(e^{-x_0} - \sin(t)) \end{aligned}$$

Methode: Numerische Lösung

$x'(t) = f(t, x(t))$ Ebene der Punkte (t, x)

für jeden Punkt (t, x) wird die Ableitung von $x(t)$ durch $f(t, x(t))$ gegeben.

Beispiel 8.16. $x'(t) = -x(t) + 1$

Richtungsfeld

Kennt man den Startwert (Anfangswert) $x(t_0)$, dann kennen wir den Punkt

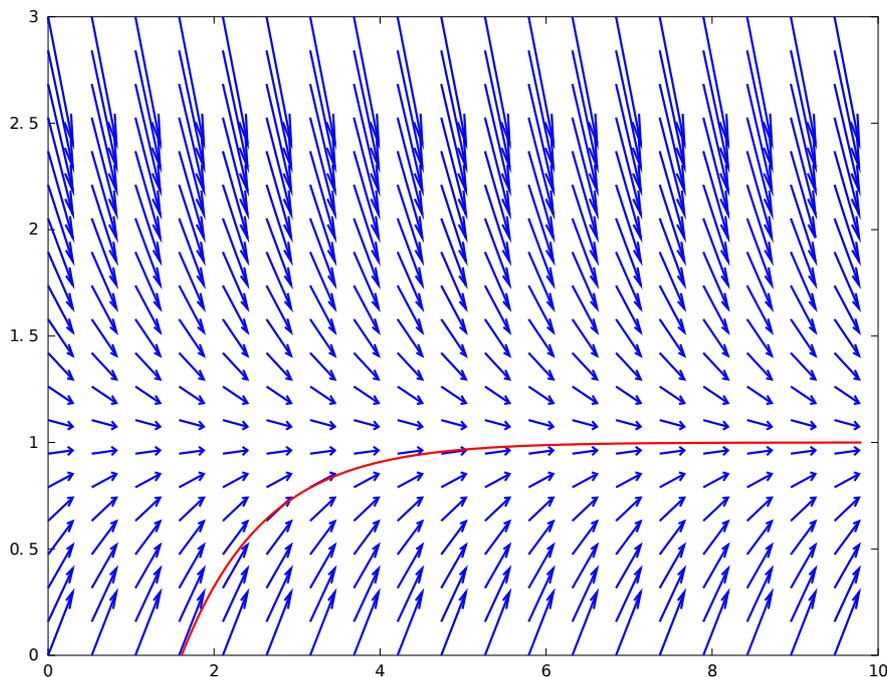


Abbildung 8.5: Richtungsfeld mit Isokline

$(t_0, x(t_0))$ und die Steigung $f(t_0, x(t_0))$, sodass für $t + dt$ der nächste Punkt $x(t + dt)$ bekannt ist.

Streckenzugverfahren nach Euler

$$x'(t) = f(t, x(t)) \text{ mit } x(t_0) = x_0$$

numerisch für Zeiten $t_0 \leq t \leq T$

Intervall $[t_0, T] \rightarrow$ Zerlegung: $t_i = t_0 + i \underbrace{dt}_{=:h} = t_0 + i \cdot \frac{T - t_0}{N}$ $i = 0, \dots, N$

$$x(t_1) = x(t_0 + h) \underset{\text{Taylor}}{\approx} x(t_0) + \underbrace{x'(x_0)}_{f(t_0, x_0)} \cdot h$$

Wir suchen: $x_1 = x_0 + f(t_0, x_0) \cdot h$

$$x_{i+1} = x_i + f(t_i, x_i) \cdot h \quad i = 0, \dots, N - 1$$

Kapitel 9

Differentialrechnung im \mathbb{R}^n

9.1 Grundlagen des \mathbb{R}^n

Definition 9.1. Seien A_1, A_2, \dots, A_n beliebige Mengen. Die Menge $A_1 \times A_2 \times \dots \times A_n = \{(a_1, a_2, \dots, a_n) \mid a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}$ wird kartesisches Produkt genannt. Ihre Elemente (a_1, a_2, \dots, a_n) heißen n -Tupel und jeder Eintrag a_i ($i = 1, \dots, n$) Komponente. Zwei n -Tupel a und \tilde{a} sind gleich, also $a = \tilde{a}$, wenn $a_i = \tilde{a}_i$ für alle $i = 1, \dots, n$

Notation: Sind alle Mengen A_i gleich, also etwa $A_i = A$ für $i = 1, \dots, n$, so schreibt man kurz A^n für das kartesische Produkt.

Definition 9.2. Die Menge $\mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R} \text{ für } i = 1, \dots, n\}$ heißt n -dimensionaler Euklidischer Raum. Die n -Tupel $x \in \mathbb{R}^n$ werden auch als Vektoren bezeichnet.

Anmerkung: Typischerweise werden Vektoren als Spaltenvektoren aufgefasst.

Also $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^T$, wobei das hochgestellte T kennzeichnet, dass der Zeilenvektor transponiert und damit ein Spaltenvektor ist.

Definition 9.3. Seien $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$.

1. $x + y = (x_1 + y_1, \dots, x_n + y_n)^T$ Addition
2. $\lambda x = (\lambda x_1, \dots, \lambda x_n)^T$ Skalarmultiplikation
3. $x^T y = x_1 y_1 + \dots + x_n y_n$ Skalarprodukt
4. $\|x\| = \sqrt{x^T x}$ Euklidische Norm
5. $d(x, y) = \|x - y\|$ Abstand von x und y
6. $x \leq y \Leftrightarrow \forall i = 1, \dots, n : x_i \leq y_i$ kleiner gleich
7. $x < y \Leftrightarrow \forall i = 1, \dots, n : x_i < y_i$ kleiner

Notation: Seien $a, b \in \mathbb{R}^n$. Dann bezeichnet $[a, b] = \{x \in \mathbb{R}^n \mid a \leq x \leq b\} = [a_1, b_1] \times \dots \times [a_n, b_n]$ und $(a, b) = \{x \in \mathbb{R}^n \mid a < x < b\} = (a_1, b_1) \times \dots \times (a_n, b_n)$ sowie entsprechend für $[a, b)$ und $(a, b]$.

Definition 9.4. Sei $\varepsilon > 0$ und $x \in \mathbb{R}^n$. Dann wird

$$\mathcal{U}_\varepsilon(x) = \{\tilde{x} \in \mathbb{R}^n \mid \|\tilde{x} - x\| < \varepsilon\}$$

eine ε -Umgebung von x genannt. Für eine Menge $M \subseteq \mathbb{R}^n$ heißt ein Element $x \in \mathbb{R}^n$

- innerer Punkt von M , wenn ein $\varepsilon > 0$ mit $\mathcal{U}_\varepsilon(x) \subset M$ existiert;
- Randpunkt von M , wenn für jedes $\varepsilon > 0$ Elemente $\tilde{x}, \hat{x} \in \mathcal{U}_\varepsilon(x)$ existieren mit $\tilde{x} \in M$ und $\hat{x} \notin M$;
- isolierter Punkt von M , wenn ein $\varepsilon > 0$ mit $\mathcal{U}_\varepsilon(x) \cap M = \{x\}$ existiert;
- Häufungspunkt von M , wenn für jedes $\varepsilon > 0$ ein Element $\tilde{x} \in \mathcal{U}_\varepsilon(x) \cap M$ mit $\tilde{x} \neq x$ existiert.

Die Menge aller inneren Punkte von M wird mit $\overset{\circ}{M}$ oder $\text{int}(M)$, die Menge aller Randpunkte mit ∂M bezeichnet. Eine Menge $M \subseteq \mathbb{R}^n$ heißt offen, wenn jeder Punkt von M ein innerer Punkt ist, und sie heißt abgeschlossen, wenn sie alle ihre Häufungspunkte enthält.

Beispiel 9.1. $(a, b) \subseteq \mathbb{R}^n$ und $\mathcal{U}_\varepsilon(x)$ sind offene Mengen, während $[a, b]$ und $\{\tilde{x} \in \mathbb{R}^n \mid \|\tilde{x} - x\| \leq \varepsilon\}$ abgeschlossene Mengen sind.

$M \setminus \partial M$ ist offen.

$M \cup \partial M$ ist abgeschlossen.

Definition 9.5 (Normkonvergenz). Eine Folge (x_k) mit $x_k \in \mathbb{R}^n$ für $k \geq 0$ konvergiert gegen $a \in \mathbb{R}^n$ für $k \rightarrow \infty$ wenn $\|x_k - a\| \rightarrow 0$.

Satz 9.6. Im \mathbb{R}^n ist Normkonvergenz gleichbedeutend mit komponentenweiser Konvergenz, also

$$\|x^{(k)} - \tilde{x}\| \rightarrow 0 \Leftrightarrow \forall i = 1, \dots, n : x_i^{(k)} \rightarrow \tilde{x}_i \text{ für } k \rightarrow \infty .$$

Für den Beweis dieses Satzes ist folgendes Resultat hilfreich:

Lemma 9.7. Sei $x \in \mathbb{R}^n$ und $\|\cdot\|$ die Euklidische Norm. Es gilt

$$0 \leq \max\{|x_1|, \dots, |x_n|\} \leq \|x\| \leq \sqrt{n} \max\{|x_1|, \dots, |x_n|\} .$$

Beweis: Quadrieren der Ungleichungen liefert

$$0 \leq \max\{x_1^2, \dots, x_n^2\} \leq \sum_{i=1}^n x_i^2 \leq n \max\{x_1^2, \dots, x_n^2\}$$

und damit die Behauptung. □

Beweis: (von Satz 9.6)

„ \Rightarrow “ Da wegen Lemma 9.7 gilt, dass

$$\begin{aligned} \|x^{(k)} - \tilde{x}\| &= \sqrt{(x_1^{(k)} - \tilde{x}_1)^2 + \dots + (x_n^{(k)} - \tilde{x}_n)^2} \\ &\geq \max\{|x_i^{(k)} - \tilde{x}_i| : i = 1, \dots, n\} \geq 0 \end{aligned}$$

und $\max\{|x_i^{(k)} - \tilde{x}_i| : i = 1, \dots, n\} \rightarrow 0 \Leftrightarrow \forall i = 1, \dots, n : x_i^{(k)} \rightarrow \tilde{x}_i$ für $k \rightarrow \infty$, folgt komponentenweise Konvergenz aus Normkonvergenz (s. Satz 3.10, Sandwich-Theorem).

„ \Leftarrow “ Da $\max\{|x_i^{(k)} - \tilde{x}_i| : i = 1, \dots, n\} \rightarrow 0 \Leftrightarrow \forall i = 1, \dots, n : x_i^{(k)} \rightarrow \tilde{x}_i$ für $k \rightarrow \infty$ und $\max\{|x_i^{(k)} - \tilde{x}_i| : i = 1, \dots, n\} \geq \frac{1}{\sqrt{n}} \|x^{(k)} - \tilde{x}\| \geq 0$ wegen Lemma 9.7 folgt Normkonvergenz aus komponentenweiser Konvergenz (s. Satz 3.10, Sandwich-Theorem). \square

$x_k \rightarrow a$ für $x_k \in \mathbb{R}^n$ und $k \rightarrow \infty$ bedeutet also, dass der Abstand $d(x_k, a) = \|x_k - a\| \rightarrow 0$ gegen Null geht. Dies ist gleichbedeutend damit, dass x_k komponentenweise gegen a konvergiert. Also $x^{(k)} = (2 + \frac{1}{k}, \frac{2}{k})^T \rightarrow (2, 0)$ für $k \rightarrow \infty$.

9.2 Stetigkeit im \mathbb{R}^n

Definition 9.8. Sei $f : M \rightarrow \mathbb{R}$ mit $M \subseteq \mathbb{R}^n$ eine Funktion und a ein Häufungspunkt von M . Dann heißt $\lim_{x \rightarrow a} f(x) = b \Leftrightarrow \forall \varepsilon > 0 : \exists \delta > 0 : |f(x) - b| < \varepsilon$ für alle $x \in M \setminus \{a\}$ mit $\|x - a\| < \delta$.

Definition 9.9. Eine Funktion $f : M \rightarrow \mathbb{R}$ mit $M \subseteq \mathbb{R}^n$ heißt stetig in $\tilde{x} \in M \Leftrightarrow$ für jede Folge (x_k) mit $x_k \in \mathbb{R}^n$, die gegen \tilde{x} strebt, konvergiert $f(x_k)$ gegen $f(\tilde{x})$. Ist f in jedem $x \in M$ stetig, so heißt f stetig auf M .

Beispiel 9.2.

1. $f(x, y) = \frac{xy}{1+x^2+y^2}$. Seien $x_k \rightarrow \tilde{x}$ und $y_k \rightarrow \tilde{y}$. Dann ist
 $f(x_k, y_k) = \frac{x_k y_k}{1+x_k^2+y_k^2} \rightarrow \frac{\tilde{x}\tilde{y}}{1+\tilde{x}^2+\tilde{y}^2} = f(\tilde{x}, \tilde{y})$ für alle $(\tilde{x}, \tilde{y})^T \in \mathbb{R}^2$.
 \Rightarrow stetig auf \mathbb{R}^2

2. $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } (x, y)^T \neq (0, 0)^T \\ 0 & \text{sonst} \end{cases}$

Seien $x_k \rightarrow \tilde{x}$ und $y_k \rightarrow \tilde{y}$. Für $(\tilde{x}, \tilde{y})^T \neq (0, 0)^T$ und $(x_k, y_k)^T \neq (0, 0)^T$ für $k \rightarrow \infty$ gilt $f(x_k, y_k) \rightarrow f(\tilde{x}, \tilde{y})$ und damit Stetigkeit für $\mathbb{R}^2 \setminus \{(0, 0)^T\}$.

Seien x_k und y_k nun Nullfolgen. Dann gilt

$$f(x_k, y_k) = \frac{x_k y_k}{x_k^2 + y_k^2} \stackrel{x_k = y_k}{=} \frac{x_k^2}{2x_k^2} = \frac{1}{2} \rightarrow \frac{1}{2} \neq f(0, 0) = 0$$

\Rightarrow unstetig in $(0, 0)$.

Wählt man $x_k = \frac{1}{k}$ und $y_k = \frac{2}{k}$, dann $f(x_k, y_k) \rightarrow \frac{2}{3}$.

9.3 Partielle Ableitungen

Definition 9.10. Seien $f : M \rightarrow \mathbb{R}$ mit $M \subseteq \mathbb{R}^n$ und $e^{(k)}$ der k -te Einheitsvektor für $k = 1, \dots, n$ im \mathbb{R}^n mit $e_k^{(k)} = 1$ und $e_i^{(k)} = 0$ für $i \neq k$. Der Limes

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{f(x + h \cdot e^{(i)}) - f(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_n)}{h} \\ &=: \frac{\partial f(x)}{\partial x_i} =: f_{x_i}(x) =: \mathcal{D}_i f(x) \end{aligned}$$

heißt partielle Ableitung (1. Ordnung) von f nach x_i an der Stelle $x \in M$, sofern er existiert. Der aus den partiellen Ableitungen 1. Ordnung gebildete Vektor $\nabla f(x) := \text{grad } f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \dots, \frac{\partial f(x)}{\partial x_n} \right)^T$ wird Gradient genannt. Sind diese partiellen Ableitungen stetig, dann ist f in x stetig partiell differenzierbar und wir schreiben $f \in C^1$.

Handwerkliches: Man erhält die partielle Ableitung von f nach x_i , indem man alle Variablen außer x_i als Konstanten auffasst und die dann nur noch von einer Variablen, nämlich x_i , abhängige Funktion in gewohnter Weise ableitet.

Beispiel 9.3.

$$\begin{aligned} a) \quad & f(x, y, z) = 2x^2 + 3xy + z \\ & f_x(x, y, z) = \frac{\partial f(x, y, z)}{\partial x} = 4x + 3y \\ & f_y(x, y, z) = \frac{\partial f(x, y, z)}{\partial y} = 3x \\ & f_z(x, y, z) = \frac{\partial f(x, y, z)}{\partial z} = 1 \\ & \nabla f(x, y, z) = (4x + 3y, 3x, 1)^T \end{aligned}$$

$$b) \quad f(x, y) = xy \cdot \sin(xy)$$

Produktregel:

$$\begin{aligned} u &= xy & v &= \sin(xy) \\ u_x &= y & v_x &= y \cos(xy) \\ u_y &= x & v_y &= x \cos(xy) \\ f_x &= u_x \cdot v + v_x \cdot u = y \sin(xy) + xy^2 \cos(xy) \\ f_y &= u_y \cdot v + v_y \cdot u = x \sin(xy) + x^2y \cos(xy) \\ \nabla f(x, y) &= (f_x, f_y)^T \end{aligned}$$

Definition 9.11. Sei $f : M \rightarrow \mathbb{R}$ mit $M \in \mathbb{R}^n$. Wenn die n partiellen Ableitungen 1. Ordnung existieren und stetig sind, dann werden für $i, j = 1, \dots, n$ die Ausdrücke

$$\frac{\partial^2 f(x)}{\partial x_j \partial x_i} := \frac{\partial f_{x_i}(x)}{\partial x_j} := f_{x_i x_j}(x)$$

partielle Ableitungen 2. Ordnung von f nach x_i und x_j an der Stelle $x \in M$ genannt. Die aus den partiellen Ableitungen 2. Ordnung gebildete quadratische Matrix

$$\nabla^2 f(x) := H(x) := \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{pmatrix}$$

heißt Hesse-Matrix. Sind die partiellen Ableitungen stetig, so schreiben wir $f \in C^2$.

Handwerkliches: Man erhält die partielle Ableitung 2. Ordnung von f nach x_i und x_j , wenn man zunächst partiell nach x_i ableitet und anschließend das Resultat partiell nach x_j ableitet.

Beispiel 9.4. $f(x, y) = 2xy + 3x + 4x^2y$

$$\begin{aligned} f_x &= 2y + 3 + 8xy & f_{xx} &= 8y & f_{xy} &= 2 + 8x \\ f_y &= 2x + 4x^2 & f_{yx} &= 2 + 8x & f_{yy} &= 0 \end{aligned}$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 8y & 2 + 8x \\ 2 + 8x & 0 \end{pmatrix}$$

Satz 9.12 (Satz von Schwarz). *(Ohne Beweis)*

Ist $f : M \rightarrow \mathbb{R}$ mit $M \subseteq \mathbb{R}^n$ in C^2 , dann $f_{x_i x_j} = f_{x_j x_i} \quad \forall i, j = 1, \dots, n$

9.4 Minima und Maxima

Definition 9.13. Sei $f : M \rightarrow \mathbb{R}$ mit $M \subseteq \mathbb{R}^n$

a) $x^* \in M$ heißt globale Minimalstelle von f , falls $\forall x \in M : f(x^*) \leq f(x)$.
Der Wert $f(x^*)$ wird dann globales Minimum genannt.

b) $x^* \in M$ heißt lokale Minimalstelle von f , falls $\exists \varepsilon > 0 : \forall x \in \mathcal{U}_\varepsilon(x^*) \cap M : f(x^*) \leq f(x)$.

Der Wert $f(x^*)$ wird dann lokales Minimum genannt.

c) Entsprechend spricht man von Maximalstellen und Maxima, wenn die Ungleichungen umgekehrt werden.

Offensichtlich ist jede globale Minimal- bzw. Maximalstelle auch eine lokale Minimal- bzw. Maximalstelle. Die Umkehrung ist i.A. falsch.

Satz 9.14 (Notwendiges Kriterium). Sei $f : M \rightarrow \mathbb{R}$ für offenes $M \subseteq \mathbb{R}^n$. Ist $x^* \in M$ lokale Extremalstelle von f und ist f in x^* partiell differenzierbar, so ist $\nabla f(x^*) = (0, 0, \dots, 0)^T$.

Beweis: Ist x^* lokale Extremalstelle von f , so auch für alle

$f_i(x_i) := f(x_1^*, \dots, x_{i-1}^*, x_i, x_{i+1}^*, \dots, x_n^*)$ mit $x_i = x_i^*$. Folglich gilt laut Satz 5.14 notwendig $f'_i(x_i^*) = 0$. Da $f'_i(x_i^*) = \frac{\partial f(x^*)}{\partial x_i}$, folgt die Behauptung. \square

Handwerkliches: Nach Nullsetzen der n partiellen Ableitungen 1. Ordnung muss man im Allgemeinen ein nichtlineares Gleichungssystem lösen. Alle Lösungen dieses Gleichungssystems sind kritische Punkte oder stationäre Lösungen oder Lösungskandidaten für die Extremalaufgabe. Ob Extrema an diesen Stellen vorliegen und wenn ja, ob Minima oder Maxima, bedarf weiterer Untersuchung.

Satz 9.15. *(Ohne Beweis)*

Sei $f : M \rightarrow \mathbb{R}$ mit offenem $M \subseteq \mathbb{R}^n$. Wenn die partiellen Ableitungen zweiter Ordnung existieren und stetig sind und außerdem $\nabla f(x^*) = (0, 0, \dots, 0)^T$ für ein $x^* \in M$ gilt, dann ist x^* eine

a) lokale Minimalstelle, wenn $x^T \cdot \nabla^2 f(x^*) x > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$

b) lokale Maximalstelle, wenn $x^T \cdot \nabla^2 f(x^*) x < 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$

Ist $x^T \nabla^2 f(x^*) x$ für mindestens ein x negativ und ein x positiv, so liegt kein Extremum vor.

Spezialfall $n = 2$

$$\nabla^2 f(x) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix}$$

$$\Delta_1 = f_{xx}(x^*) \quad \Delta_2 = f_{xx}(x^*) \cdot f_{yy}(x^*) - f_{xy}^2(x^*)$$

$$x^T \nabla^2 f(x^*) x > 0 \Leftrightarrow \overbrace{f_{xx}(x^*)}^{\Delta_1} > 0 \text{ und } \overbrace{f_{xx}(x^*) \cdot f_{yy}(x^*) - f_{xy}^2(x^*)}^{\Delta_2} > 0$$

$$x^T \nabla^2 f(x^*) x < 0 \Leftrightarrow f_{xx}(x^*) < 0 \text{ und } f_{xx}(x^*) \cdot f_{yy}(x^*) - f_{xy}^2(x^*) > 0$$

Ist $f_{xx}(x^*) \cdot f_{yy}(x^*) - f_{xy}^2(x^*) < 0 \Rightarrow$ kein Extremum in x^*

Ist $f_{xx}(x^*) \cdot f_{yy}(x^*) - f_{xy}^2(x^*) = 0 \Rightarrow$ keine Aussage möglich.

Beispiel 9.5.

$$\begin{aligned}
 a) \quad & f(x, y) = x^3 + y^3 - 3xy \\
 & \left. \begin{aligned} f_x = 3x^2 - 3y &\stackrel{!}{=} 0 \Leftrightarrow x^2 \stackrel{!}{=} y \\ f_y = 3y^2 - 3x &\stackrel{!}{=} 0 \Leftrightarrow x \stackrel{!}{=} y^2 \end{aligned} \right\} \Rightarrow \text{kritische Punkte } (0,0)^T \text{ und } (1,1)^T
 \end{aligned}$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}$$

$$\nabla^2 f(0, 0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix} \text{ ist indefinit } (\Delta_2 < 0) \Rightarrow (0, 0)^T \text{ keine Extremalstelle}$$

$$\nabla^2 f(1, 1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

$$\left. \begin{aligned} \Delta_1 = 6 > 0 \\ \Delta_2 = 6 \cdot 6 - (-3)(-3) = 27 > 0 \end{aligned} \right\} \Rightarrow \text{lokales Minimum: } f(1, 1) = -1$$

Globales Minimum?

$f(x, x) \rightarrow +\infty$ für $x \rightarrow \infty$ und $f(x, x) \rightarrow -\infty$ für $x \rightarrow -\infty \Rightarrow f$ unbeschränkt
 \Rightarrow kein globales Minimum.

$$\begin{aligned}
 b) \quad & f(x, y) = x^2 + y^2 - 2xy + 1 \\
 & \left. \begin{aligned} f_x = 2x - 2y &\stackrel{!}{=} 0 \Leftrightarrow x \stackrel{!}{=} y \\ f_y = 2y - 2x &\stackrel{!}{=} 0 \Leftrightarrow x \stackrel{!}{=} y \end{aligned} \right\} \Rightarrow \text{kritische Punkte } (x, y)^T \text{ mit } x = y
 \end{aligned}$$

$$\nabla^2 f(x, y) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}$$

$\Delta_1 > 0$ und $\Delta_2 = 0 \rightarrow$ so keine Entscheidung möglich

Analyse durch Nachdenken und Abschätzung der Funktion

$$f(x, y) = x^2 + y^2 - 2xy + 1 = \underbrace{(x - y)^2}_{\geq 0} + 1 \geq 1 \text{ wird angenommen, wenn } x = y;$$

also für alle kritischen Punkte \Rightarrow alle Elemente in $\{(x, y)^T \in \mathbb{R}^2 \mid x = y\}$ sind lokale und globale Minimalstellen.

$$\begin{aligned}
 c) \quad & f(x, y) = \sqrt{x^2 + y^2} \\
 & \text{Offensichtlich: globales Minimum in } (0, 0)^T
 \end{aligned}$$

Analyse durch partielle Ableitung

$$f_x = 2x \cdot \frac{1}{2\sqrt{x^2 + y^2}} \quad f_y = 2y \cdot \frac{1}{2\sqrt{x^2 + y^2}}$$

$$\left. \begin{aligned} f_x(0, y) = 0 \text{ für } y \neq 0 \\ f_y(x, 0) = 0 \text{ für } x \neq 0 \end{aligned} \right\} \text{ nicht simultan erfüllbar}$$

$f_x(0, 0)$ und $f_y(0, 0)$ nicht definiert $\Rightarrow \nexists$ partiellen Ableitungen in $(0, 0)^T$

d) Gegeben seien N Punkte $x^{(1)}, \dots, x^{(N)} \in \mathbb{R}^n$. Für welches $x \in \mathbb{R}^n$ wird die Summe der quadratischen Abstände zwischen $x^{(i)}$ und x minimal?

$$\text{Formal: } \sum_{i=1}^N \|x^{(i)} - x\|^2 \rightarrow \min!$$

$$f(x) = \sum_{i=1}^N \sum_{j=1}^n \left(\underbrace{x_j^{(i)}}_{\text{bekannt}} - \underbrace{x_j}_{\text{Variable}} \right)^2$$

alle Variablen mit Index $j \neq k$ sind Konstanten!

$$\begin{aligned}
\frac{\partial f(x)}{\partial x_k} &= \sum_{i=1}^N \sum_{j=1}^n \underbrace{\frac{\partial}{\partial x_k} (x_j^{(i)} - x_j)^2}_{=0 \text{ für } j \neq k} \\
&= \sum_{i=1}^N \frac{\partial}{\partial x_k} \left(\underbrace{x_k^{(i)}}_{\text{konstant}} - \underbrace{x_k}_{\text{Variable}} \right)^2 \\
&= \sum_{i=1}^N 2(x_k^{(i)} - x_k)(-1) \\
&= 2 \sum_{i=1}^N (x_k - x_k^{(i)}) \\
&= 2 \left(\sum_{i=1}^N x_k - \sum_{i=1}^N x_k^{(i)} \right) \\
&= 2Nx_k - 2 \sum_{i=1}^N x_k^{(i)} \stackrel{!}{=} 0
\end{aligned}$$

für $k = 1, \dots, n$

$$\Leftrightarrow Nx_k \stackrel{!}{=} \sum_{i=1}^N x_k^{(i)} \Leftrightarrow x_k \stackrel{!}{=} \frac{1}{N} \sum_{i=1}^N x_k^{(i)}$$

also: $x^* = \frac{1}{N} \sum_{i=1}^N x^{(i)}$ ist Lösungskandidat

$$\frac{\partial^2 f(x)}{\partial x_k^2} = 2N \text{ und } \frac{\partial^2 f(x)}{\partial x_k \partial x_l} = 0 \text{ für } k \neq l \Rightarrow \nabla^2 f(x) = \begin{pmatrix} 2N & & 0 \\ & \ddots & \\ 0 & & 2N \end{pmatrix}$$

Offensichtlich ist $x^T \nabla^2 f(x) x = 2N \sum_{j=1}^n x_j^2 = 2N \|x\|^2 > 0$ für $x \neq 0$. Lokales Minimum!

Kapitel 10

Abzählende Kombinatorik

Die abzählende Kombinatorik ist ein Teilgebiet der diskreten Mathematik und bietet Grundlagen

- zur Behandlung von kombinatorischen Problemen der Informatik und
- zur Einführung der Wahrscheinlichkeitstheorie mit endlichem Grundraum.

10.1 Grundlagen

Zur Erinnerung:

$|A|$ bezeichnet die *Kardinalität* (Anzahl der Elemente) der Menge A .

Satz 10.1. *Seien A_1, \dots, A_n endliche Mengen. Dann gilt*

$$a) \left| \bigcup_{i=1}^n A_i \right| = \sum_{i=1}^n |A_i| \text{ falls } A_i \cap A_j = \emptyset \text{ für alle } i \neq j.$$

$$b) \left| \prod_{i=1}^n A_i \right| = \prod_{i=1}^n |A_i|.$$

Beweis:

ad a) Da die Mengen paarweise disjunkt sind, können in der Vereinigung keine Duplikate wegfallen.

ad b) Vollständige Induktion:

$$\text{Für } n = 1 \text{ ist die Aussage wahr, da } \left| \prod_{i=1}^1 A_i \right| = |A_1| = \prod_{i=1}^1 |A_i|.$$

Sei die Aussage wahr für $n > 1$. Dann ist

$$\begin{aligned}
 \left| \prod_{i=1}^{n+1} A_i \right| &= \left| \left(\prod_{i=1}^n A_i \right) \times A_{n+1} \right| \\
 &= \left| \bigcup_{a \in A_{n+1}} \left\{ (x, a) : x \in \prod_{i=1}^n A_i \right\} \right| \\
 &\stackrel{(a)}{=} \sum_{a \in A_{n+1}} \left| \left\{ (x, a) : x \in \prod_{i=1}^n A_i \right\} \right| \\
 &= \sum_{a \in A_{n+1}} \left| \prod_{i=1}^n A_i \right| \\
 &= |A_{n+1}| \cdot \left| \prod_{i=1}^n A_i \right| \\
 &\stackrel{(*)}{=} |A_{n+1}| \cdot \prod_{i=1}^n |A_i| = \prod_{i=1}^{n+1} |A_i|
 \end{aligned}$$

unter Verwendung der Induktionsannahme in Gleichung (*).

□

10.2 Prinzip des doppelten Abzählens

Wenn man die Elemente einer Menge auf zwei verschiedene Weisen abzählt, dann müssen die Ergebnisse gleich sein. Die Präzisierung dieser Aussage gibt das folgende Resultat.

Satz 10.2 (Prinzip des doppelten Abzählens). *Sei $R \subseteq A \times B$ eine Relation über endliche Mengen A und B . Für $x \in A$ bezeichne $a(x)$ die Anzahl der mit x in Relation stehenden $y \in B$ und $b(y)$ die mit y in Relation stehenden $x \in A$. Dann gilt*

$$\sum_{x \in A} a(x) = \sum_{y \in B} b(y).$$

Beweis: Sei $I_R(x, y) = \begin{cases} 1 & , \text{ falls } (x, y) \in R \\ 0 & , \text{ sonst} \end{cases}$

die *Indikatorfunktion* oder *charakteristische Funktion* der Relation R . Dann gilt per Definition

$$\begin{aligned}
 a(x) &= \sum_{y \in B} I_R(x, y) \quad \text{für } x \in A \text{ und} \\
 b(y) &= \sum_{x \in A} I_R(x, y) \quad \text{für } y \in B.
 \end{aligned}$$

Folglich ist

$$\begin{aligned} \sum_{x \in A} a(x) &= \sum_{x \in A} \sum_{y \in B} I_R(x, y) = \underbrace{\sum_{y \in B} \sum_{x \in A} I_R(x, y)}_{= \sum_{(x, y) \in A \times B} I_R(x, y) = |R|} = \sum_{y \in B} b(y). \quad \square \end{aligned}$$

Dieses Resultat scheint eine offensichtliche Tatsache zu sein (und ist, wie gesehen, auch leicht zu begründen); es liefert aber oft überraschende Vereinfachungen.

Beispiel 10.1. In einem sozialen Netzwerk beschreibe der Beliebtheitsgrad $b(p) \in \mathbb{N}_0$ einer Person $p \in P$ im Netzwerk, wieviele Freunde man im Netzwerk hat. Freundschaftsbeziehungen sind symmetrisch. Für jedes soziale Netzwerk dieser Art ist die Summe der Beliebtheitsgrade über aller Personen im Netzwerk eine gerade Zahl. Warum?

Wir verwenden das Prinzip des doppelten Abzählens.

Sei $F \subseteq P \times P$ die Menge aller Freundschaftsbeziehungen zwischen 2 Personen im Netzwerk und die Indikatorfunktion $I(p, f) = 1$, wenn die Person $p \in P$ an der Freundschaftsbeziehung $f \in F$ beteiligt ist, ansonsten ist $I(p, f) = 0$.

Wir zählen nun auf zwei verschiedene Weisen:

1. Die Summe aller Freundschaftsbeziehungen, an denen p beteiligt ist, ist gerade $b(p)$. Also

$$\sum_{p \in P} \underbrace{\left(\sum_{f \in F} I(p, f) \right)}_{b(p)} = \sum_{p \in P} b(p).$$

2. Die Summe aller Personen, die an einer Freundschaftsbeziehung $f \in F$ beteiligt sind, ist genau 2. Also

$$\sum_{f \in F} \underbrace{\left(\sum_{p \in P} I(p, f) \right)}_2 = 2 \cdot |F|.$$

Folglich ist

$$\sum_{p \in P} b(p) = 2 \cdot |F|$$

und damit die Summe aller Beliebtheitsgrade eine gerade Zahl. \square

10.3 Schubfachprinzip

Das nächste Resultat dient der Bereitstellung von nicht-konstruktiven Existenzargumenten. Wir betrachten zunächst eine mathematisch präzise Formulierung, bevor die etwas anschaulichere Darstellung erörtert wird.

Satz 10.3 (Schubfachprinzip; engl. pigeonhole principle). Sei $f : X \rightarrow Y$ eine Abbildung zwischen endlichen Mengen X und Y mit $|X| > |Y|$. Dann existiert ein $y \in Y$ mit $|f^{-1}(y)| \geq 2$.

Wenn man also n Elemente auf k Fächer aufteilt und $n > k$ ist, dann landen zwangsläufig in irgendeinem Fach mindestens 2 Elemente. Wenn also $n = 400$ Studierende im Hörsaal sind, dann haben mindestens 2 von ihnen am selben Tag Geburtstag ($k = 366$). Die Existenz eines solchen Paares von Studierenden ist somit gesichert! Es wird aber nichts darüber ausgesagt, wer diese Personen sind. Zum Beweis wird also keine Lösung konstruiert.

Beispiel 10.2.

a) In einer Umzugskiste sind 10 nicht unterscheidbare weiße Socken und 10 nicht unterscheidbare schwarze Socken. Wie oft muss ich höchstens in der Dunkelheit Socken aus der Kiste ziehen, um sicher zu gehen, dass ich ein gleichfarbiges Paar habe?

Die Anzahl der Fächer ist $k = 2$ (weiße Socken, schwarze Socken). Nach Satz 10.3 muss $n > k$ sein, so dass $n = k + 1 = 3$ Züge reichen.

b) Gegeben sei ein Kartenspiel (z.B. für Skat) mit 32 Karten. Wie oft muss ich höchstens vom Stapel ziehen, um sicher zwei Karten mit der gleichen Spielfarbe in der Hand zu haben?

Die Anzahl der Fächer ist $k = 4$, nämlich die Anzahl der Spielfarben (Kreuz, Pik, Herz, Karo). Nach Satz 10.3 muss $n > k$ sein, so dass $n = k + 1 = 5$ Züge reichen.

c) Wieviele Personen müssen im Hörsaal sein, dass sicher 3 von ihnen am selben Tag Geburtstag haben?

Dies erfordert eine leichte Verallgemeinerung von Satz 10.3. Wir haben $k = 366$ Fächer. Wir können maximal $2k$ viele Personen in die Fächer „legen“, ohne dass in einem Fach 3 Personen sind (in jedem Fach sind 2 Personen). Kommt noch eine Person hinzu, dann sind sicher in einem Fach 3 Personen. Also muss hier $n > 2k$ gelten, so dass $n = 2k + 1 = 733$ Personen ausreichen. Im vollen Audimax trifft diese Situation also zu.

Eine Verallgemeinerung von Satz 10.3 lautet also:

Gibt es mehr als m mal so viele Elemente wie Fächer, dann muss ein Fach mindestens $m + 1$ Elemente enthalten. \square

10.4 Auswahl von k aus n Elementen

Definition 10.4. Für $n \in \mathbb{N}_0$ und $k \in \mathbb{N}$ wird der Ausdruck

$$\binom{n}{k} := \frac{n}{1} \cdot \frac{n-1}{2} \cdot \dots \cdot \frac{n-k+1}{k} \quad (\text{sprich: „}n \text{ über } k\text{“})$$

als Binomialkoeffizient bezeichnet. Für $k = 0$ wird

$$\binom{n}{0} := 1$$

für alle $n \in \mathbb{N}_0$ festgelegt. \square

Eigenschaften:

$$\text{a) } \binom{n}{k} = \begin{cases} \frac{n!}{k!(n-k)!} & , \text{ für } n \geq k \\ 0 & , \text{ für } n < k \end{cases}$$

$$\text{b) } \binom{n}{0} = \binom{n}{n} = 1 \text{ für } n \in \mathbb{N}_0.$$

$$\text{c) } \text{Aus a) folgt } \binom{n}{k} = \binom{n}{n-k} \text{ für } n \geq k \geq 0.$$

Satz 10.5 (Rekursionsformel für Binomialkoeffizienten). *Es gilt* $\binom{n+1}{k+1} = \binom{n}{k+1} + \binom{n}{k}$ *für* $k, n \in \mathbb{N}_0$.

Beweis: mit Fallunterscheidung (keine Induktion).

1. Fall: Sei $k = 0$ und $n \in \mathbb{N}_0$. Dann ist die Formel korrekt:

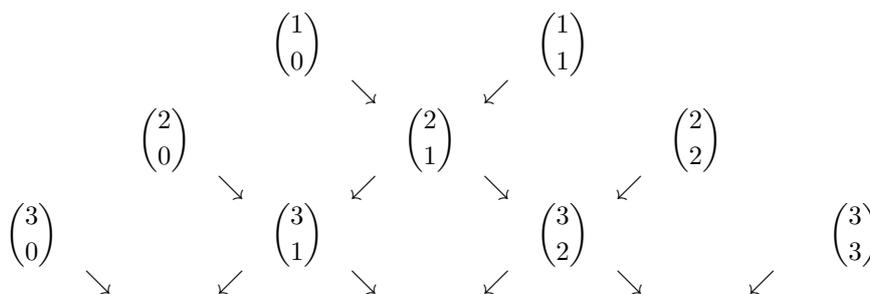
$$\binom{n+1}{1} = \frac{n+1}{1} \stackrel{!}{=} \binom{n}{1} + \binom{n}{0} = \frac{n}{1} + 1.$$

2. Fall: Sei $k > 0$ und $n \in \mathbb{N}_0$. Dann ist

$$\begin{aligned} \binom{n}{k+1} + \binom{n}{k} &= \frac{n!}{k!(n-k)!} + \frac{n!}{(k+1)!(n-(k+1))!} \\ &= \frac{n!}{k!(n-k)!} + \frac{n!(n-k)}{(k+1) \cdot k!(n-k) \cdot (n-(k+1))!} \\ &= \frac{n!}{k!(n-k)!} + \frac{n!}{k!(n-k)!} \cdot \frac{n-k}{k+1} \\ &= \frac{n!}{k!(n-k)!} \cdot \left(1 + \frac{n-k}{k+1}\right) \\ &= \frac{n!}{k!(n-k)!} \cdot \frac{n+1}{k+1} \\ &= \frac{(n+1)!}{(k+1)!(n+1-(k+1))!} = \binom{n+1}{k+1}. \end{aligned}$$

□

Die Illustration dieser Rekursionsformel erfolgt am sogenannten *Pascalschen Dreieck*:



Der binomische Lehrsatz ist Namensgeber für die Binomialkoeffizienten:

Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$.

Durch geschickte Wahl von a und b ergeben sich leicht folgende Identitäten:

- Es gilt $\sum_{k=0}^n \binom{n}{k} = 2^n$ für $a = b = 1$ und $n \in \mathbb{N}_0$.
- Es gilt $\sum_{k=0}^n \binom{n}{k} (-1)^k = 0$ für $a = -1, b = 1$ und $n \in \mathbb{N}$.

Satz 10.6. Die Anzahl der Möglichkeiten für die Auswahl von k Elementen aus einer n -elementigen Menge ist

$$\binom{n}{k}$$

für $k, n \in \mathbb{N}_0$.

Beweis: Bezeichne $a(n, k)$ die Anzahl von k -elementigen Teilmengen einer n -elementigen Menge. Unser Ziel ist die Bestimmung von $a(n, k)$. Erneut benötigen wir eine Fallunterscheidung:

Fall 1: Sei $|M| = n + 1$ und $A \subseteq M$ mit $|A| = k + 1$ für $k, n \geq 1$.

Gesucht ist nun eine Formel für $a(n + 1, k + 1)$.

Wähle dazu ein beliebiges $x \in M$. Entweder ist $x \notin A$ oder $x \in A$.

Gilt $x \notin A$, dann ist A eine $(k + 1)$ -elementige Teilmenge der n -elementigen Menge $M \setminus \{x\}$. Davon gibt es $a(n, k + 1)$ Stück.

Gilt $x \in A$, so konstruieren wir alle diese Teilmengen und zählen sie dann ab. Dazu platzieren wir zuerst x in die leere Menge. Anschließend müssen noch k weitere Elemente aus $M \setminus \{x\}$ platziert werden. Das sind gerade alle k -elementigen Teilmengen aus $M \setminus \{x\}$. Davon gibt es $a(n, k)$ Stück.

Insgesamt gibt es also $a(n, k + 1) + a(n, k)$ Stück. Damit haben wir eine Rekursionsformel gefunden:

$$a(n + 1, k + 1) = a(n, k) + a(n, k + 1).$$

Fall 2: Sei $n = 0$ und $k > 0$. Es gibt keine nichtleeren Teilmengen der leeren Menge. Also ist $a(0, k) = 0$.

Fall 3: Sei $n \in \mathbb{N}_0$ und $k = 0$. Es gibt nur eine leere Menge, die man ziehen könnte. Also ist $a(n, 0) = 1$.

Damit sind alle Fälle untersucht und wir stellen fest, dass sowohl die Rekursionsformel aus Fall 1 als auch die Anfangsbedingungen in den beiden anderen Fällen genau mit denen des Binomialkoeffizienten (Satz 10.5) übereinstimmen. Folglich müssen die so erzeugten Werte gleich sein, so dass

$$a(n, k) = \binom{n}{k}$$

gilt.

□

Beispiel 10.3.

- a) Lotto 6 aus 49 (ohne Superzahl): $\binom{49}{6}$ Möglichkeiten.
- b) Lotto 6 aus 49 (mit Superzahl): $\binom{49}{6} \cdot \binom{10}{1}$ Möglichkeiten.
- c) Eurojackpot: $\binom{50}{5} \cdot \binom{8}{2}$ Möglichkeiten.

Satz 10.7. Die Anzahl der Möglichkeiten für die Auswahl von k Elementen einer n -elementigen Menge für $k, n \in \mathbb{N}_0$ ist:

	Anordnung im k -Tupel	keine Anordnung
mit Zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

Beweis: Zunächst sollen die gezogenen Elemente von links nach rechts in einem k -Tupel abgelegt werden.

Wenn die Elemente nach jedem Zug wieder in die Menge zurückgelegt werden können, dann hat man

$$\left. \begin{array}{l} \text{bei der 1. Komponente} : n \text{ Möglichkeiten} \\ \text{bei der 2. Komponente} : n \text{ Möglichkeiten} \\ \vdots \\ \text{bei der } k. \text{ Komponente} : n \text{ Möglichkeiten} \end{array} \right\} \underbrace{n \cdot n \cdot \dots \cdot n}_{k \text{ mal}} = n^k$$

Wenn die Elemente nach jedem Zug nicht wieder in die Menge zurückgelegt werden können, dann hat man

$$\left. \begin{array}{l} \text{bei der 1. Komponente} : n \text{ Möglichkeiten} \\ \text{bei der 2. Komponente} : n - 1 \text{ Möglichkeiten} \\ \vdots \\ \text{bei der } k. \text{ Komponente} : n - (k - 1) \text{ Möglichkeiten} \end{array} \right\} = \frac{n!}{(n - k)!}$$

Ab jetzt soll die Reihenfolge der gezogenen Elemente keine Rolle mehr spielen. Wenn die Elemente nach jedem Zug nicht wieder in die Menge zurückgelegt werden können, dann hat man die gleiche Situation wie bei der Auswahl einer k -elementigen Menge aus einer n -elementigen Menge in Satz 10.6, also

$$\binom{n}{k} \text{ Möglichkeiten.}$$

Wenn die Elemente nach jedem Zug wieder in die Menge zurückgelegt werden können, dann erhalten wir die Anzahl der Möglichkeiten durch folgende Überlegung: Gegeben seien n Urnen U_1, \dots, U_n , wobei in Urne U_i mindestens k Kugeln

mit Nummer i (und nur solche) für $i = 1, \dots, n$ sind. Eine k -elementige Menge ziehen wir, indem wir bei Urne U_1 starten und für jede Urne i entscheiden, ob wir entweder eine Kugel ziehen (Z) oder zur nächsten Urne weitergehen (W). Man kann also aus einer Urne mehrfach ziehen, bevor man weitergeht. Insgesamt darf aber nur k mal gezogen werden und wir müssen $n - 1$ mal weitergehen, um aus jeder Urne möglicherweise ziehen zu können. Offensichtlich beschreibt eine Sequenz mit k mal Z und $n - 1$ mal W eindeutig eine mögliche Auswahl von k Elementen aus einer n -elementigen Menge und jede Auswahl ist auch derart zu realisieren.

Insgesamt hat die Sequenz $k + n - 1$ viele Einträge mit Z oder W und jede Sequenz beschreibt eine mögliche Auswahl. Wieviele verschiedene Sequenzen sind möglich? Also: Auf wieviele Möglichkeiten kann ich k Einträge von Typ Z in der Sequenz anordnen? Das sind gerade

$$\binom{n+k-1}{k} \text{ Möglichkeiten.}$$

Man könnte auch fragen, auf wieviele verschiedenen Möglichkeiten man die $n - 1$ Einträge vom Typ W anordnen kann. Das sind gerade

$$\binom{n+k-1}{n-1} \text{ Möglichkeiten.}$$

Beide Ausdrücke müssen gleich sein wie wir sofort aus dem Prinzip des doppelten Abzählens folgern können. Alternativ liefert das Einsetzen in die „Fakultätsformel“ für den Binomialkoeffizienten die Gültigkeit der Identität. \square

10.5 Prinzip der Inklusion / Exklusion

Gegeben seien die endlichen Mengen A, B und C . Wir wollen wissen, wieviele Elemente in der Vereinigung dieser Mengen liegen, wenn sie nicht wie in Satz 10.1 a) als disjunkt vorausgesetzt werden.

Satz 10.8. *Seien A_1, \dots, A_n endliche Mengen. Dann gilt*

$$\left| \bigcup_{i=1}^n A_i \right| = \sum_{k=1}^n (-1)^{k+1} \cdot \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}|.$$

Beweis: Laut Behauptung gilt

$$\begin{aligned} & \left| \bigcup_{i=1}^n A_i \right| \\ = & |A_1| + |A_2| + \dots + |A_n| && \text{Ebene 1} \\ - & (|A_1 \cap A_2| + |A_1 \cap A_3| + \dots + |A_1 \cap A_n| + \dots + |A_{n-1} \cap A_n|) && \text{Ebene 2} \\ + & (|A_1 \cap A_2 \cap A_3| + |A_1 \cap A_2 \cap A_4| + \dots + |A_{n-2} \cap A_{n-1} \cap A_n|) && \text{Ebene 3} \\ & \vdots \\ \pm & |A_1 \cap A_2 \cap \dots \cap A_n| && \text{Ebene } n \end{aligned}$$

Die Formel ist richtig, wenn jedes Element in der Vereinigungsmenge exakt ein Mal gezählt wird. Sei ein Element in k der Mengen A_1, \dots, A_n enthalten ($1 \leq k \leq n$). Wie oft wird das Element gezählt?

auf Ebene 1: $\binom{k}{1}$ Mal hinzu
 auf Ebene 2: $\binom{k}{2}$ Mal abziehen
 auf Ebene 3: $\binom{k}{3}$ Mal hinzu
 \vdots
 auf Ebene k : $\binom{k}{k}$ Mal hinzu (k ungerade) / abziehen (k gerade)

Also ist die Häufigkeit, mit der das Element gezählt wurde, gerade

$$\begin{aligned} \sum_{i=1}^k \binom{k}{i} (-1)^{i+1} &= (-1) \cdot \sum_{i=1}^k \binom{k}{i} (-1)^i \\ &= (-1) \cdot \left[\sum_{i=0}^k \binom{k}{i} (-1)^i - \binom{k}{0} \right] \\ &= \underbrace{\binom{k}{0}}_{=1} - \underbrace{\sum_{i=0}^k \binom{k}{i} (-1)^i}_{=(-1+1)^k=0^k=0} = 1 \end{aligned}$$

und die Behauptung damit bewiesen. □

Beispiel 10.4.

a) *Wieviele Wörter mit 4 Buchstaben beginnen oder enden mit einem Vokal?*

$$V*** + ***V - V**V = 5 \cdot 26^3 + 26^3 \cdot 5 - 5 \cdot 26^2 \cdot 5.$$

b) *Wieviele Wörter mit 10 Buchstaben enthalten nicht alle Vokale?*

Sei W_Z die Menge aller Worte der Länge 10 ohne Buchstabe Z . Wir suchen $|W_A \cup W_E \cup W_I \cup W_O \cup W_U| =$

$$\begin{aligned} &|W_A| + |W_E| + |W_I| + |W_O| + |W_U| && \binom{5}{1} \cdot 25^{10} \\ - &(|W_A \cap W_E| + |W_A \cap W_I| + \dots + |W_O \cap W_U|) && - \binom{5}{2} \cdot 24^{10} \\ + &(|W_A \cap W_E \cap W_I| + \dots) && + \binom{5}{3} \cdot 23^{10} \\ - &\dots && - \binom{5}{4} \cdot 22^{10} \\ + &|W_A \cap W_E \cap W_I \cap W_O \cap W_U| && + \binom{5}{5} \cdot 21^{10} \\ &= \sum_{k=1}^5 (-1)^{k+1} \binom{5}{k} (26-k)^{10}. \end{aligned}$$

□

Literaturverzeichnis

- [1] O. Forster *Analysis 1*. Vieweg + Teubner Studium, 11. Auflage, 2011.
- [2] W. Beekmann *Analysis 1*. Skript, Fernuniversität Hagen, 1998.
- [3] W. Walter *Analysis 1*. Springer Verlag, 7. Auflage, 2004.
- [4] K. Königsberger *Analysis*. Springer Verlag, 6. Auflage, 2003.
- [5] M. Skutella. Mathematik für Informatiker I und II. Technical report, TU Dortmund, 2006.
- [6] G. Teschl, S. Teschl. *Mathematik für Informatiker - Band 2 Analysis und Statistik*. Springer Verlag, 2. Auflage, 2007.
- [7] M. Oberguggenberger, A. Ostermann. *Analysis for Computer Scientists*. Springer Verlag, 2011.