

## Lösungsvorschlag zum 13. Übungsblatt

### Analysis I

Winter Semester 2025/2026

#### Aufgabe 1:

(a) Berechnen Sie die Ableitungen der folgenden Funktionen.

- (i)  $f: \mathbb{R} \rightarrow \mathbb{R}; f(x) = \cosh x := \frac{1}{2}(e^x + e^{-x})$ ,    (ii)  $f: (-\frac{\pi}{10}, \frac{\pi}{10}) \rightarrow \mathbb{R}; f(x) = (8-x)\tan(5x)$ ,  
 (iii)  $f: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}; f(x) = \ln|x|$ ,    (iv)  $f: \mathbb{R} \rightarrow \mathbb{R}; f(x) = \cos(x)\arctan(x)$ .

(b) Seien  $a, b, c \in \mathbb{R}$  und  $f: \mathbb{R} \rightarrow \mathbb{R}$  definiert durch

$$f(x) = \begin{cases} ax^2 + bx + c, & x \geq 0, \\ \cos x, & x < 0. \end{cases}$$

Für welche  $a, b, c \in \mathbb{R}$  ist  $f$  stetig; stetig differenzierbar; dreimal differenzierbar?

#### Lösungsvorschlag zu Aufgabe 1:

(a) (i) Nach Beispiel 5.13 a) ist für jedes  $a \in \mathbb{R}$  die Funktion  $g_a: \mathbb{R} \rightarrow \mathbb{R}, g_a(x) = e^{ax}$  differenzierbar und es gilt  $g'_a(x) = ae^{ax}$  für alle  $x \in \mathbb{R}$ . Aus der Linearität der Ableitung (s. Satz 5.5 a)) folgt daher für  $x \in \mathbb{R}$

$$f'(x) = (\frac{1}{2}g_1 + \frac{1}{2}g_{-1})'(x) = \frac{1}{2}g'_1(x) + \frac{1}{2}g'_{-1}(x) = \frac{1}{2}e^x - \frac{1}{2}e^{-x} = \sinh(x).$$

(ii) Nach Beispiel 5.13 b) und der Kettenregel (s. Satz 5.7) ist die Funktion  $g: (-\frac{\pi}{10}, \frac{\pi}{10}) \rightarrow \mathbb{R}, g(x) = \tan(5x)$  differenzierbar und für die Ableitung gilt  $g'(x) = 5(\cos(5x))^{-2}$  für alle  $x \in (-\frac{\pi}{10}, \frac{\pi}{10})$ . Weiter ist nach Beispiel 5.3 a) die Funktion  $h: \mathbb{R} \rightarrow \mathbb{R}, h(x) = 8-x$  differenzierbar und es gilt  $h'(x) = -1$  für alle  $x \in \mathbb{R}$ . Nach der Produktregel (s. Satz 5.5 b)) gilt also für  $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$

$$f'(x) = (h \cdot g)'(x) = h'(x)g(x) + h(x)g'(x) = -\tan(5x) + (8-x)5(\cos(5x))^{-2}$$

(iii) Nach Beispiel 5.13 c) ist  $\ln: \mathbb{R}_+ \rightarrow \mathbb{R}$  differenzierbar mit Ableitung  $\ln'(x) = \frac{1}{x}$  für alle  $x \in \mathbb{R}_+$ . Weiterhin ist die Funktion  $g: \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, g(x) = |x|$  differenzierbar und es gilt  $g'(x) = \operatorname{sgn}(x)$  für alle  $x \in \mathbb{R} \setminus \{0\}$ , wobei

$$\operatorname{sgn}(x) := \begin{cases} 1 & \text{für } x > 0, \\ -1 & \text{für } x < 0 \end{cases}$$

(man beachte, dass dies kein Widerspruch zu Beispiel 5.3 b) ist, da 0 *nicht* zum Definitionsbereich von  $g$  gehört). Nun folgt aus der Kettenregel (s. Satz 5.7) für  $x \in \mathbb{R} \setminus \{0\}$

$$f'(x) = (\ln \circ g)'(x) = \ln'(g(x))g'(x) = \frac{1}{|x|}\operatorname{sgn}(x) = \frac{1}{x}.$$

(iv) Aus der Produktregel (s. Satz 5.5 b)) und aus Beispielen 5.13 a), e) erhalten wir für  $x \in \mathbb{R}$

$$f'(x) = (\cos \cdot \arctan)'(x) = \cos'(x)\arctan(x) + \cos(x)\arctan'(x) = -\sin(x)\arctan(x) + \frac{\cos(x)}{1+x^2}.$$

(b) Die Funktion

$$f \text{ ist } \begin{cases} \text{stetig} & \text{genau dann, wenn } c = 1, \\ \text{stetig differenzierbar} & \text{genau dann, wenn } b = 0, c = 1, \\ \text{dreimal stetig differenzierbar} & \text{genau dann, wenn } a = -\frac{1}{2}, b = 0, c = 1. \end{cases}$$

Nach Definition ist  $f|_{[0,\infty)} = p$  und  $f|_{(-\infty,0)} = q$ , wobei  $p: [0, \infty) \rightarrow \mathbb{R}, p(x) = ax^2 + bx + c$  und  $q: (-\infty, 0) \rightarrow \mathbb{R}, q(x) = \cos(x)$ . Durch eine induktive Anwendung von Beispielen 5.6 b) und 5.13 a) erhalten wir, dass  $p \in C^\infty(I)$  und  $q \in C^\infty(J)$ . Hieraus folgt, dass  $f \in C^\infty(\mathbb{R} \setminus \{0\})$ . Also hängt die  $k$ -malige (stetige) Differenzierbarkeit von  $f$  nur vom Punkt  $x_0 = 0$  ab. Mit Induktion zeigt man, dass für  $k \in \mathbb{N}$  die Funktion  $f$  genau dann  $k$ -mal stetig differenzierbar (d.h.  $f \in C^k(\mathbb{R})$ ), wenn

$$p^{(j)}(0) = q^{(j)}(0) \quad \text{für alle } j = 0, 1, \dots, k.$$

(mit der Konvention, dass  $p^{(0)} := p, q^{(0)} = q$  und  $C^0(\mathbb{R}) := C(\mathbb{R}, \mathbb{R})$ ). Nun gilt

$$\begin{aligned} p(x) &= ax^2 + bx + c, & p'(x) &= 2ax + b, & p''(x) &= 2a, & p'''(x) &= 0, \\ q(y) &= \cos(y), & q'(y) &= -\sin(y), & q''(y) &= -\cos(y), & q'''(y) &= \sin(y) \end{aligned}$$

für alle  $x \in I$  und  $y \in J$ . Also ist

$$f \text{ ist } \begin{cases} \text{stetig} & \Leftrightarrow p(0) = q(0) \Leftrightarrow c = 1, \\ \text{stetig differenzierbar} & \Leftrightarrow p'(0) = q'(0), c = 1 \Leftrightarrow b = 0, c = 1, \\ \text{zweimal stetig differenzierbar} & \Leftrightarrow p''(0) = q''(0), b = 0, c = 1 \Leftrightarrow a = -\frac{1}{2}, b = 0, c = 1. \end{cases}$$

Da  $p'''(0) = 0 = q'''(0)$  (unabhängig von  $a, b, c \in \mathbb{R}$ ) gilt, ist  $f$  für  $a = -\frac{1}{2}, b = 0, c = 1$  nicht nur zweimal, sondern sogar dreimal stetig differenzierbar.

### Aufgabe 2 (K):

(a) Zeigen Sie, dass die folgenden Funktionen differenzierbar sind und berechnen Sie deren Ableitung.

$$\begin{aligned} \text{(i)} \quad f &: (1, \infty) \rightarrow \mathbb{R}; f(x) = \ln(\ln x), & \text{(ii)} \quad f &: \mathbb{R} \rightarrow \mathbb{R}; f(x) = |x^2 - 4|^3, \\ \text{(iii)} \quad f &: \mathbb{R}_+ \rightarrow \mathbb{R}; f(x) = x^{(2^x)}, & \text{(iv)} \quad f &: \mathbb{R}_+ \rightarrow \mathbb{R}; f(x) = \frac{(x + \sin x \cos x + 3)^{5/3}}{e^x - \sin x}. \end{aligned}$$

(b) Berechnen Sie mit Hilfe der Umkehrregel (Satz 5.10) die Ableitung von

$$\begin{aligned} \text{(i)} \quad \arccos &: (-1, 1) \rightarrow \mathbb{R}, \text{ wobei } \arccos \text{ die Umkehrfunktion von } \cos : (0, \pi) \rightarrow \mathbb{R} \text{ ist,} \\ \text{(ii)} \quad \sinh^{-1} &: \mathbb{R} \rightarrow \mathbb{R}, \text{ wobei } \sinh^{-1} \text{ die Umkehrfunktion von } \sinh : \mathbb{R} \rightarrow \mathbb{R} \text{ ist.} \end{aligned}$$

Dabei sollen  $\arccos$  bzw.  $\sinh^{-1}$  in den Formeln für  $\arccos'$  bzw.  $(\sinh^{-1})'$  nicht vorkommen.

*Hinweis zu (b) (ii): Aufgabe 1 des 12. Übungsblattes kann hilfreich sein.*

### Lösungsvorschlag zu Aufgabe 2:

(a) (i) Zunächst einmal ist  $f$  wohldefiniert, da die Menge  $\ln((1, \infty)) = (0, \infty)$  im Definitionsbereich des  $\ln$  liegt. Nach Beispiel 5.13 c) ist  $\ln: \mathbb{R}_+ \rightarrow \mathbb{R}$  differenzierbar und für die Ableitung gilt  $\ln'(x) = \frac{1}{x}$  für alle  $x \in \mathbb{R}_+$ . Somit ist  $f$  nach der Kettenregel (s. Satz 5.7) differenzierbar und es gilt

$$f'(x) = (\ln \circ \ln)'(x) = \ln'(\ln(x)) \cdot \ln'(x) = \frac{1}{\ln(x)} \cdot \frac{1}{x} = \frac{1}{x \ln(x)} \quad \text{für alle } x \in (1, \infty).$$

(ii) Die Funktion  $g: \mathbb{R} \rightarrow \mathbb{R}$ ,

$$g(y) = |y|^3 = \begin{cases} y^3 & \text{für } y \geq 0, \\ -y^3 & \text{für } y < 0, \end{cases}$$

ist differenzierbar mit Ableitung  $g'(y) = 3y|y|$  für  $y \in \mathbb{R}$ : Für  $y \neq 0$  folgt dies direkt aus Beispiel 5.6 a) und für  $y = 0$  folgt ebenfalls mit Beispiel 5.6 a)

$$\frac{d^+g}{dy}(0) = 3y^2|y=0 = 0 \quad \text{und} \quad \frac{d^-g}{dy}(0) = -3y^2|y=0 = 0$$

und damit  $g'(0) = 0$ . Somit erhalten wir aus der Kettenregel, dass  $f$  differenzierbar ist und für die Ableitung gilt

$$f'(x) = \frac{d}{dx}(g(x^2 - 4)) = 3(x^2 - 4) \cdot |x^2 - 4| \cdot 2x = 6x \cdot (x^2 - 4) \cdot |x^2 - 4|.$$

(iii) Gemäß Definition 4.46 gilt  $f(x) = x^{2^x} = e^{\ln(x)2^x} = e^{\ln(x)e^{\ln(2)x}}$  für alle  $x \in \mathbb{R}_+$ . Also ist  $f = \exp \circ g$  mit  $g: \mathbb{R}_+ \rightarrow \mathbb{R}$ ,  $g(x) = \ln(x)e^{\ln(2)x}$ . Nach Beispiel 5.13 und der Produktregel (s. Satz 5.5 b)) ist  $g$  differenzierbar. Nach der Kettenregel (s. Satz 5.7) ist damit auch  $f$  differenzierbar und wir erhalten für  $x \in \mathbb{R}_+$

$$\begin{aligned} f'(x) &= (\exp \circ g)'(x) = \exp(g(x))g'(x) = f(x) \left[ \frac{1}{x}e^{\ln(2)x} + \ln(x) \ln(2)e^{\ln(2)x} \right] \\ &= x^{2^x} 2^x \left( \frac{1}{x} + \ln(2) \ln(x) \right). \end{aligned}$$

(iv) Zunächst einmal ist  $f$  wohldefiniert, da für  $x \in \mathbb{R}_+$  einerseits  $x + \sin(x) \cos(x) + 3 \geq 0 - 1 + 3 = 2 \geq 0$  gilt (und damit gebrochene Potenzen hiervon definiert sind) und andererseits wegen  $e^x - \sin(x) > 1 - \sin(x) \geq 0$  die Nennerfunktion nullstellenfrei ist. Die Zählerfunktion  $g$  und die Nennerfunktion  $h$  sind nach Satz 5.5 a), b), Satz 5.7 und Beispielen 5.3 a), 5.13 a), d) differenzierbar. Nach der Quotientenregel (s. Satz 5.5 c)) ist damit auch  $f = \frac{g}{h}$  differenzierbar und für  $x \in \mathbb{R}$  gilt

$$f'(x) = \frac{\frac{5}{3}(x + \sin x \cos x + 3)^{2/3}(1 + \cos^2 x - \sin^2 x)[e^x - \sin x] - [x + \sin x \cos x + 3]^{5/3}[e^x - \cos(x)]}{(e^x - \sin x)^2}$$

(b) (i) Nach Beispiel 5.13 a) gilt  $\cos' = -\sin$ , und da  $\sin(y) \neq 0$  für alle  $y \in \arccos((-1, 1)) = (0, \pi)$ , können wir die Umkehrregel (s. Satz 5.10) anwenden und erhalten dann für alle  $x \in (-1, 1)$

$$\arccos'(x) = \frac{1}{\cos'(\arccos(x))} = -\frac{1}{\sin(\arccos(x))},$$

wobei sich die rechte Seite mit Hilfe des trigonometrischen Pythagoras (s. (3.8) des Skriptes) vereinfachen lässt: Für  $y \in (0, \pi)$  gilt  $\sin(y) > 0$  und daher können wir aus  $\cos^2(y) + \sin^2(y) = 1$  die Gleichung  $\sin(y) = \sqrt{1 - \cos^2(y)}$  folgern. Da  $\arccos((-1, 1)) = (0, \pi)$  gilt, erhalten wir daher

$$\arccos'(x) = -\frac{1}{\sin(\arccos(x))} = -\frac{1}{\sqrt{1 - \cos^2(\arccos(x))}} = -\frac{1}{\sqrt{1 - x^2}} \quad \text{für alle } x \in (-1, 1).$$

(ii) Zunächst einmal bestimmen wir die Ableitung von  $\sinh$ . Nach Beispiel 5.13 a) ist für jedes  $a \in \mathbb{R}$  die Funktion  $g_a: \mathbb{R} \rightarrow \mathbb{R}$ ,  $g_a(x) = e^{ax}$  differenzierbar und es gilt  $g'_a(x) = ae^{ax}$  für alle  $x \in \mathbb{R}$ . Damit erhalten wir für  $x \in \mathbb{R}$

$$\sinh(x) = \left( \frac{1}{2}g'_1 - \frac{1}{2}g'_{-1} \right)(x) = \left( \frac{1}{2}g_1 + \frac{1}{2}g_{-1} \right)(x) = \cosh(x).$$

Da  $\sinh$  nach dem letzten Übungsblatt bijektiv ist und  $\cosh(y) \neq 0$  für alle  $y \in \mathbb{R}$  ist, können wir die Umkehrregel anwenden. Dann erhalten wir für alle  $x \in \mathbb{R}$

$$(\sinh^{-1})'(x) = \frac{1}{(\sinh)'(\sinh^{-1}(x))} = \frac{1}{\cosh(\sinh^{-1}(x))}.$$

Den Term auf der rechten Seite können wir etwas vereinfachen: Nach Aufgabe 1 des 12. Übungsblatt gilt für  $x \in \mathbb{R}$

$$\cosh^2(x) - \sinh^2(x) = 1 \stackrel{\cosh > 0}{\iff} \cosh(x) = \sqrt{1 + \sinh^2(x)}.$$

Wir folgern für  $x \in \mathbb{R}$

$$(\sinh^{-1})'(x) = \frac{1}{\cosh(\sinh^{-1}(x))} = \frac{1}{\sqrt{1 + \sinh^2(\sinh^{-1}(x))}} = \frac{1}{\sqrt{1 + x^2}}.$$

### Aufgabe 3 (K):

Für welche  $\alpha \in \mathbb{R}$  ist die Funktion  $g : [0, \infty) \rightarrow \mathbb{R}$  gegeben durch

$$g(x) = \begin{cases} x^\alpha \sin\left(\frac{1}{x}\right), & x > 0 \\ 0, & x = 0 \end{cases}$$

stetig; differenzierbar; stetig differenzierbar?

### Lösungsvorschlag zu Aufgabe 3:

Die Funktion

$$g \text{ ist } \begin{cases} \text{stetig} & \iff \alpha > 0, \\ \text{differenzierbar} & \iff \alpha > 1, \\ \text{stetig differenzierbar} & \iff \alpha > 2. \end{cases} \quad (1)$$

Für den Beweis von (1) benutzen wir folgendes

**Lemma:** Sei  $\beta \in \mathbb{R}$ . Dann gelten:

$$(a) \lim_{x \rightarrow 0^+} x^\beta \sin\left(\frac{1}{x}\right) \text{ existiert} \iff \lim_{x \rightarrow 0^+} x^\beta \cos\left(\frac{1}{x}\right) \text{ existiert} \iff \beta > 0.$$

Obige Grenzwerte sind, sofern sie existieren, null.

(b) Die Funktionen  $h_{1,\beta} : [0, \infty) \rightarrow \mathbb{R}$  und  $h_{2,\beta} : [0, \infty) \rightarrow \mathbb{R}$  definiert durch

$$h_{1,\beta}(x) = \begin{cases} x^\beta \sin\left(\frac{1}{x}\right), & x > 0, \\ 0, & x = 0 \end{cases} \quad \text{und} \quad h_{2,\beta}(x) = \begin{cases} x^\beta \cos\left(\frac{1}{x}\right), & x > 0, \\ 0, & x = 0 \end{cases}$$

sind genau dann stetig, falls  $\beta > 0$ .

**Beweis:** (a) Wir zeigen nur

$$\lim_{x \rightarrow 0^+} x^\beta \sin\left(\frac{1}{x}\right) \text{ existiert} \iff \beta > 0,$$

der Beweis der Äquivalenz „ $\lim_{x \rightarrow 0^+} x^\beta \cos\left(\frac{1}{x}\right) \text{ existiert} \iff \beta > 0$ “ geht analog.

„ $\Leftarrow$ “: Sei  $\beta > 0$ . Da die Funktion  $\mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto x^\beta$  stetig ist, haben wir für  $x \rightarrow 0^+$

$$0 \leq |x^\beta \sin\left(\frac{1}{x}\right)| = |x^\beta| |\sin\left(\frac{1}{x}\right)| \leq x^\beta \rightarrow 0,$$

wobei wir ausgenutzt haben, dass  $|\sin(y)| \leq 1$  für alle  $y \in \mathbb{R}$ . Nach dem Sandwichkriterium folgt damit  $\lim_{x \rightarrow 0^+} x^\beta \sin\left(\frac{1}{x}\right) = 0$ .

„ $\Rightarrow$ “: Angenommen,  $\lim_{x \rightarrow 0^+} x^\beta \sin\left(\frac{1}{x}\right)$  existiert. Nach Beispiel 4.52 ist dann  $\beta \neq 0$ . Angenommen, es gälte  $\beta < 0$ . Dann gälte für die Folge  $(x_n) = \left(\left(\frac{\pi}{2} + 2\pi n\right)^{-1}\right)$ , dass  $x_n \searrow 0$  für  $n \rightarrow \infty$ , aber

$$x_n^\beta \sin\left(\frac{1}{x_n}\right) = \left(\frac{\pi}{2} + 2\pi n\right)^{-\beta} \cdot 1 = e^{-\ln\left(\frac{\pi}{2} + 2\pi n\right)\beta} \rightarrow \infty \quad (n \rightarrow \infty),$$

da  $-\beta > 0$ ; dies stünde allerdings im Widerspruch zur Voraussetzung, dass  $\lim_{x \rightarrow 0^+} x^\beta \sin\left(\frac{1}{x}\right)$  existiert. Also muss  $\beta > 0$  sein.

(b) Die Funktionen  $h_{1,\beta}$  und  $h_{2,\beta}$  sind als Verknüpfung stetiger Funktionen stetig auf  $(0, \infty)$  (unabhängig von  $\beta \in \mathbb{R}$ ), also sind die Funktionen genau dann auf  $[0, \infty)$  stetig, wenn sie im Punkt  $x_0 = 0$  stetig sind. Letzteres ist nach (a) aber äquivalent dazu, dass  $\beta > 0$  ist. □

Sei nun  $g$  so wie in der Aufgabenstellung und  $\alpha \in \mathbb{R}$ . Mit Hilfe des obigen Lemmas können wir nun die Behauptung (1) leicht beweisen:

- *Stetigkeit:* Da  $g = h_{1,\alpha}$  gilt, ist  $g$  nach obigem Lemma genau dann stetig, wenn  $\alpha > 0$  ist.
- *Differenzierbarkeit:* Die Funktion  $(0, \infty) \rightarrow \mathbb{R}, x \mapsto x^\alpha$  ist nach Beispiel 5.13 d) für jedes  $\alpha \in \mathbb{R}$  differenzierbar, die Funktion  $(0, \infty) \rightarrow \mathbb{R}, x \mapsto \sin\left(\frac{1}{x}\right)$  ist nach Beispiel 5.13 a), d) und der Kettenregel differenzierbar, sodass auch  $g$  nach der Produktregel auf  $(0, \infty)$  differenzierbar ist. Für  $x > 0$  gilt hierbei

$$g'(x) = \alpha x^{\alpha-1} \sin\left(\frac{1}{x}\right) + x^\alpha \cdot \left(-\frac{1}{x^2}\right) \cos\left(\frac{1}{x}\right) = \alpha h_{1,\alpha-1}(x) - h_{2,\alpha-2}(x).$$

Daher ist die Funktion  $g$  genau dann auf  $[0, \infty)$  differenzierbar, wenn  $g$  im Punkt  $x_0 = 0$  differenzierbar ist. Letzteres ist genau dann der Fall, wenn der Grenzwert

$$\lim_{x \rightarrow 0^+} \frac{g(x) - g(0)}{x} = \lim_{x \rightarrow 0^+} \frac{g(x)}{x} = \lim_{x \rightarrow 0^+} x^{\alpha-1} \sin\left(\frac{1}{x}\right)$$

existiert. Nach obigem Lemma ist dies aber genau dann der Fall, wenn  $\alpha - 1 > 0$ , also  $\alpha > 1$  ist. In diesem Fall gilt zusammengefasst

$$g'(x) = \alpha h_{1,\alpha-1}(x) - h_{2,\alpha-2}(x) \quad \text{für alle } x \in [0, \infty) \quad (2)$$

- *Stetige Differenzierbarkeit:* Es gilt

$$g \text{ ist stetig differenzierbar} \iff \alpha > 2.$$

„ $\Leftarrow$ “: Ist  $\alpha > 2$ , so ist  $g$  nach dem bereits Bewiesenen differenzierbar und für die Ableitung gilt (2). Nach obigem Lemma sind wegen  $\alpha > 2$  die Funktionen  $h_{1,\alpha-1}$  und  $h_{2,\alpha-2}$  stetig, und nach (2) damit auch  $g'$ . Also ist  $g$  stetig differenzierbar.

„ $\Rightarrow$ “: Sei  $g$  stetig differenzierbar. Dann ist  $g$  insbesondere differenzierbar und nach dem bereits Bewiesenen gilt  $\alpha > 1$ . An (2) erkennen wir  $h_{2,\alpha-2} = g' - \alpha h_{1,\alpha-1}$ . Da die rechte Seite der Gleichung stetig ist ( $g'$  ist stetig, da  $g$  nach Voraussetzung stetig differenzierbar ist und  $h_{1,\alpha-1}$  ist nach obigem Lemma stetig wegen  $\alpha > 1$ ), muss es auch die linke Seite sein. Aus obigem Lemma folgt dann sofort  $\alpha > 2$ .

Bemerkung: Insbesondere folgt für  $\alpha = 2$ , dass  $g$  zwar differenzierbar, aber *nicht* stetig differenzierbar ist. Also folgt aus Differenzierbarkeit keine stetige Differenzierbarkeit!

#### Aufgabe 4:

(a) Sei  $f: [0, 2] \rightarrow \mathbb{R}, f(x) = x^3 - 3x + 1$ . Zeigen Sie, dass  $f$  Maximum und Minimum annimmt und bestimmen Sie diese.

(b) Seien  $a, b \in \mathbb{R}$  mit  $a < b$  und  $f: [a, b] \rightarrow \mathbb{R}$  differenzierbar in  $a$ . Die Funktion  $f$  besitze in  $a$  ein lokales Minimum. Zeigen Sie  $\frac{d^+}{dx} f(a) \geq 0$ . Was passiert für ein lokales Maximum in  $a$ , und was für  $\frac{d^-}{dx}$  bei  $b$ ?

(c) Seien  $I, J \subseteq \mathbb{R}$  Intervalle,  $f: I \rightarrow \mathbb{R}$  und  $g: J \rightarrow \mathbb{R}$  mit  $f(I) \subseteq J$ , sowie  $x_0 \in I$  kein Randpunkt. Formulieren und beweisen Sie eine Kettenregel für  $\frac{d^+}{dx}(g \circ f)(x_0)$ .

#### Lösungsvorschlag zu Aufgabe 4:

(a) Zunächst einmal ist  $f$  als Polynom stetig und das Intervall  $[0, 2]$  abgeschlossen und beschränkt. Also

nimmt  $f$  nach Theorem 4.24 Maximum und Minimum an. Sei  $x_0 \in [0, 2]$  die Stelle des Minimums und  $x_1 \in [0, 2]$  die Stelle des Maximums. Wir wollen nun  $x_0$  und  $x_1$  mit Hilfe von Satz 5.17 bestimmen. Die Funktion  $f$  ist als Polynom nicht nur stetig sondern nach Beispiel 5.6 b) auch differenzierbar. Für  $x \in [0, 2]$  gilt

$$f'(x) = 0 \iff 3x^2 - 3 = 0 \iff x^2 = 1 \stackrel{x \geq 0}{\iff} x = 1.$$

Weiter gilt  $f(1) = 1 - 3 + 1 = -1$ . Dies bedeutet: Wenn  $f$  in  $(0, 2)$  ein lokales Extremum hat, so muss dieses  $-1$  sein. Für die Funktionswerte am Rand erhalten wir  $f(0) = 1$  und  $f(2) = 3$ . Damit ist  $f(x_0) \leq f(1) < f(0)$  und  $f(x_0) \leq f(1) < f(2)$ , es muss also  $x_0 \in (0, 2)$  sein. Insbesondere muss  $f(x_0)$  ein lokales Minimum sein, und da es nur ein lokales Extremum gibt, muss  $x_0 = 1$ , also  $f(x_0) = f(1) = -1$  das Minimum von  $f$  sein. Da  $f$  nichtkonstant ist und es nur ein lokales Extremum gibt, muss  $f$  das Maximum auf dem Rand des Intervalls  $[0, 2]$  annehmen (anderenfalls wäre  $x_1$  ein lokales Maximum, und es müsste  $f'(x_1) = 0$  und nach obiger Äquivalenz damit  $x_0 = x_1 = 1$  gelten; dies ginge aber nur dann, wenn  $f$  konstant wäre). Da  $f(0) < f(2)$  gilt, folgern wir  $x_1 = 2$ . Damit ist  $f(2) = 3$  das Maximum von  $f$ .

(b) Da  $f$  im Punkt  $a$  ein lokales Minimum besitzt, existiert ein  $\delta > 0$  so, dass für alle  $x \in (a, a + \delta)$  gilt  $f(x) \geq f(a)$ . D.h.  $f(x) - f(a) \geq 0$ . Da wir  $x$  aus einer rechtsseitigen Umgebung von  $a$  wählen, gilt weiter  $x - a > 0$ . Der Differenzenquotient ist somit als Quotient zweier nicht-negativer Werte selbst nicht-negativ

$$\frac{f(x) - f(a)}{x - a} \geq 0 \quad \text{für alle } x \in (a, a + \delta).$$

Da  $f$  in  $a$  differenzierbar ist, existiert der Grenzwert für  $x \rightarrow a^+$ . Folglich

$$\frac{d^+}{dx} f(a) = \lim_{x \rightarrow a^+} \frac{f(x) - f(a)}{x - a} \geq 0.$$

(c) Eine Möglichkeit wäre folgende:

Seien  $I, J \subseteq \mathbb{R}$  Intervalle,  $f: I \rightarrow \mathbb{R}$  und  $g: J \rightarrow \mathbb{R}$  mit  $f(I) \subseteq J$ , sowie  $x_0 \in I$  kein Randpunkt. Es gelte  $f(I \cap [x_0, \infty)) \subseteq J \cap [f(x_0), \infty)$ . Es existiere  $\frac{d^+}{dx} f(x_0)$  und  $\frac{d^+}{dx} g(f(x_0))$ . Dann existiert auch  $\frac{d^+}{dx} g \circ f(x_0)$  und es gilt

$$\frac{d^+}{dx} g \circ f(x_0) = \frac{d^+}{dx} g(f(x_0)) \cdot \frac{d^+}{dx} f(x_0).$$

(Andere Regeln sind auch korrekt. Die Aufgabe ist relativ offen gestellt.)

*Beweis.* Sei  $(x_n)_{n \in \mathbb{N}}$  eine Folge in  $I$  mit  $x_n > x_0$  und  $\lim_{n \rightarrow \infty} x_n = x_0$ . Wir untersuchen den Differenzenquotienten

$$Q_n := \frac{g(f(x_n)) - g(f(x_0))}{x_n - x_0}$$

für  $n \rightarrow \infty$ .

Nach Voraussetzung gilt  $f(I \cap [x_0, \infty)) \subseteq [f(x_0), \infty)$ . Da  $x_n > x_0$  ist, folgt auch  $f(x_n) \geq f(x_0)$ . Da  $f$  in  $x_0$  rechtsseitig differenzierbar ist, ist  $f$  in  $x_0$  auch rechtsseitig stetig. Es gilt also  $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$ .

Wir unterscheiden zwei Fälle für die Glieder der Folge  $(x_n)_n$ :

**Fall 1:**  $f(x_n) = f(x_0)$ . Dann gilt  $g(f(x_n)) - g(f(x_0)) = 0$ , folglich ist  $Q_n = 0$ . Tritt dieser Fall für unendlich viele  $n$  auf (d.h. für eine Teilfolge), so folgt aus der Existenz des Limes

$$\frac{d^+}{dx} f(x_0) = \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = 0.$$

Der Grenzwert von  $Q_n$  ist in diesem Fall 0, was mit der oben formulierten Kettenregel übereinstimmt

$$\frac{d^+}{dx} g(f(x_0)) \cdot \underbrace{\frac{d^+}{dx} f(x_0)}_{=0} = 0.$$

**Fall 2:**  $f(x_n) \neq f(x_0)$ . Wegen der eingangs gezeigten Ungleichung muss dann  $f(x_n) > f(x_0)$  gelten. Wir erweitern den Differenzenquotient wie im Beweis von Satz 5.7

$$Q_n = \frac{g(f(x_n)) - g(f(x_0))}{f(x_n) - f(x_0)} \cdot \frac{f(x_n) - f(x_0)}{x_n - x_0}.$$

Der zweite Faktor konvergiert für  $n \rightarrow \infty$  gegen  $\frac{d^+}{dx} f(x_0)$ . Für den ersten Faktor betrachten wir die Folge  $y_n := f(x_n)$ . Es gilt  $y_n \rightarrow f(x_0)$  mit  $y_n > f(x_0)$ . Aufgrund der Existenz der rechtsseitigen Ableitung von  $g$  in  $f(x_0)$  gilt

$$\lim_{n \rightarrow \infty} \frac{g(f(x_n)) - g(f(x_0))}{f(x_n) - f(x_0)} = \frac{d^+}{dx} g(f(x_0)).$$

Also konvergiert  $Q_n$  gegen das Produkt  $\frac{d^+}{dx} g(f(x_0)) \cdot \frac{d^+}{dx} f(x_0)$ . □

**Aufgabe 5 (Zusatz: Anwendungsbeispiel NN):**

Die folgende Aufgabe wird in der Übung besprochen.

Wir modellieren ein vereinfachtes neuronales Netz als Verkettung von  $L$  Schichten, wobei jede Schicht aus nur einem einzigen Neuron besteht. (Mehrere Neuronen lassen sich mit Techniken der Analysis 2 untersuchen.)

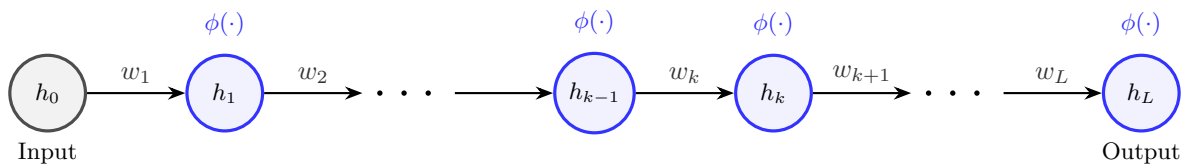
Das Netzwerk sei rekursiv definiert durch

$$\begin{aligned} h_0 &= x \quad (\text{Eingabe}), \\ h_k &= \phi(w_k \cdot h_{k-1}) \quad \text{für } k = 1, \dots, L. \end{aligned}$$

Hierbei sind  $w_k \in \mathbb{R}$  Gewichte und  $\phi: \mathbb{R} \rightarrow \mathbb{R}$  ist eine Aktivierungsfunktion. Ziel ist es, den Fehler

$$E(x, w_1, \dots, w_L) = \frac{1}{2} (h_L(x, w_1, \dots, w_L) - y)^2$$

für eine gegebene Menge an (Eingabe, Ausgabe)-Paaren  $\{(x, y)\} \subseteq \mathbb{R} \times \{0, 1\}$  zu minimieren, indem die Gewichte  $\{w_k\}_{k=1, \dots, L}$  angepasst werden. (Dies wäre ein *binäres Klassifikationsproblem*.)



Eine weit verbreitete Aktivierungsfunktion ist die *logistische Sigmoidfunktion*

$$\phi: \mathbb{R} \rightarrow (0, 1); \quad \phi(s) = \sigma(s) = \frac{1}{1 + e^{-s}},$$

welche wir in dieser Aufgabe untersuchen.

- (a) Berechnen Sie die Ableitung  $\sigma'(s)$ . Zeigen Sie, dass  $\sigma'(s) = \sigma(s)(1 - \sigma(s))$ .
- (b) Bestimmen Sie das Maximum der Ableitung  $\sigma'(s)$ .
- (c) Seien  $x, y, w_2, \dots, w_L \in \mathbb{R}$  gegeben. Sei  $E^{(x,y,w_2,\dots,w_L)}(w_1) := E(x, y, w_1, w_2, \dots, w_L)$  und analog  $h_k^{(x,y,w_2,\dots,w_k)}(w_1) := h_k(x, y, w_1, w_2, \dots, w_k)$ . Wir frieren quasi alle Argumente außer dem Gewicht  $w_1$  ein. Zeigen Sie

$$\frac{d}{dw_1} E^{(x,y,w_2,\dots,w_L)}(w_1) = \left( h_L^{(x,y,w_2,\dots,w_L)}(w_1) - y \right) \cdot \left[ \prod_{k=2}^L w_k \sigma'(w_k h_{k-1}^{(x,y,w_2,\dots,w_{k-1})}(w_1)) \right] \cdot h_0 \sigma'(w_1 h_0).$$

- (d) Nehmen Sie nun an, dass alle Gewichte beschränkt sind durch  $|w_k| \leq 1$ , die Eingabe  $|x| = |h_0| \leq 1$  und der Fehlerterm  $|h_L - y| \leq 1$  ist. Zeigen Sie

$$\left| \frac{d}{dw_1} E^{x,y,w_2,\dots,w_L}(w_1) \right| \leq \frac{1}{4L}.$$

*Bemerkung:* Dies bedeutet, dass das Ergebnis bei tiefen Netzen ( $L$  groß) nur unwesentlich von einer Änderung des Gewichts  $w_1$  abhängt. Daher ist die Sigmoidfunktion für tiefe Netze gegebenenfalls ungeeignet. ( $\rightarrow$  *Vanishing Gradient Problem*) Es bietet sich an eine andere Aktivierungsfunktion zu verwenden. Beispielsweise findet die „*Rectified Linear Unit*“-Funktion

$$\phi(s) = \text{ReLU}(s) = \max(0, s)$$

Anwendung.

### Lösungsvorschlag zu Aufgabe 5:

- (a) Wir betrachten die Funktion  $\sigma(s) = (1 + e^{-s})^{-1}$ . Unter Anwendung der Kettenregel, Satz 5.7, gilt

$$\begin{aligned} \sigma'(s) &= \frac{d}{ds} (1 + e^{-s})^{-1} = -(1 + e^{-s})^{-2} \cdot \frac{d}{ds} (1 + e^{-s}) \\ &= -(1 + e^{-s})^{-2} \cdot (-e^{-s}) = \frac{e^{-s}}{(1 + e^{-s})^2}. \end{aligned}$$

Weiter formen wir um

$$\begin{aligned} \sigma'(s) &= \frac{1}{1 + e^{-s}} \cdot \frac{e^{-s}}{1 + e^{-s}} = \frac{1}{1 + e^{-s}} \cdot \frac{(1 + e^{-s}) - 1}{1 + e^{-s}} \\ &= \sigma(s) \cdot \left( \frac{1 + e^{-s}}{1 + e^{-s}} - \frac{1}{1 + e^{-s}} \right) = \sigma(s) \cdot (1 - \sigma(s)). \end{aligned}$$

- (b) Gesucht ist das Maximum der Funktion  $f(s) = \sigma'(s) = \sigma(s)(1 - \sigma(s))$ .

Mit Hilfe des Zwischenwertsatz sieht man z.B. leicht ein, dass  $\sigma$  surjektiv nach  $(0, 1)$  abbildet (Machen Sie sich das klar!). Daher substituieren wir  $u = \sigma(s)$  mit  $u \in (0, 1)$ . D.h das Maximum von  $f$  ist gleich dem Maximum von  $g: (0, 1) \rightarrow \mathbb{R}; g(u) = u(1 - u) = u - u^2$ .

Entsprechend Satz 5.18, bestimmen wir die Nullstelle der Ableitung von  $g$ .

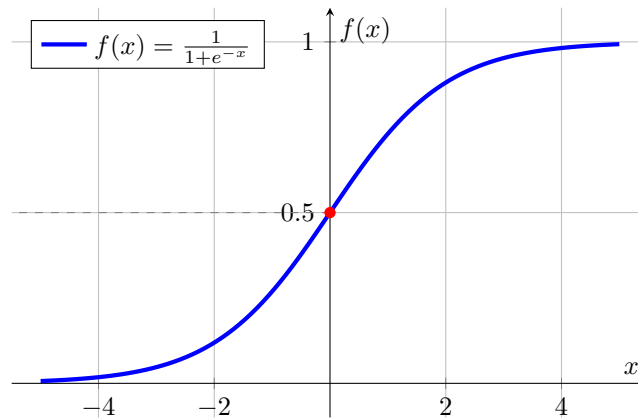
$$g'(u) = 1 - 2u \stackrel{!}{=} 0 \implies u = \frac{1}{2}.$$

Da  $g''(u) = -2 < 0$ , liegt ein (zunächst lokales) Maximum vor mit  $g(\frac{1}{2}) = \frac{1}{4}$ . Untersuchen wir noch kurz den Rand  $\{0, 1\}$  des Definitionsbereichs von  $g$ , so sehen wir  $\lim_{u \rightarrow 0} g = \lim_{u \rightarrow 1} g = 0$ . Daher gilt  $\max_{u \in (0,1)} g(u) = \frac{1}{4}$ . Der maximale Wert von  $\sigma'$  ist somit

$$\max_{s \in \mathbb{R}} \sigma'(s) = \frac{1}{2} \left( 1 - \frac{1}{2} \right) = \frac{1}{4}.$$

Dieser wird z.B. für  $s = 0$  angenommen mit  $\sigma(0) = \frac{1}{2}$ . (Wenn Sie sich den Funktionsgraphen von  $\sigma$  anschauen, sollte Sie das nicht überraschen.)

### Logistic Sigmoid Function



(c) Wir betrachten alle Größen im Netzwerk als Funktionen von  $w_1$ . Im folgenden lassen wir häufig die Argumente der Funktionen und die Abhängigkeit von  $(x, y, w_2, \dots, w_k)$  weg um die Notation zu vereinfachen. Wenn wir im Folgenden also  $E$  schreiben meinen wir  $E = E^{(x, y, w_2, \dots, w_L)}(w_1)$ , analog für  $h_k$ . Idee ist es die Kettenregel, Satz 5.7  $L$ -mal anzuwenden.

Wir beginnen mit der Fehlerfunktion  $E(w_1) = \frac{1}{2}(h_L(w_1) - y)^2$ . Nach der Kettenregel gilt

$$\frac{d}{dw_1} E = (h_L - y) \cdot \frac{d}{dw_1} h_L.$$

Für jede Schicht  $k$  mit  $L \geq k \geq 2$  ist  $h_k = h_k^{(x, y, w_2, \dots, w_k)} = \sigma(w_k \cdot h_{k-1}^{(x, y, w_2, \dots, w_{k-1})}(w_1))$ . Da  $w_k$  hierbei eine Konstante ist, gilt nach Satz 5.7

$$\frac{d}{dw_1} h_k = \sigma'(w_k h_{k-1}) \cdot w_k \cdot \frac{d}{dw_1} h_{k-1}.$$

Rekursiv folgt

$$\begin{aligned} \frac{d}{dw_1} h_L &= w_L \sigma'(s_L) \cdot \frac{d}{dw_1} h_{L-1} = w_L \sigma'(s_L) \cdot \left( w_{L-1} \sigma'(s_{L-1}) \cdot \frac{d}{dw_1} h_{L-2} \right) \\ &= \dots = \left( \prod_{k=2}^L w_k \sigma'(w_k h_{k-1}) \right) \cdot \frac{d}{dw_1} h_1 \end{aligned}$$

Für die erste Schicht gilt  $h_1 = \sigma(w_1 \cdot h_0)$ . Hier ist  $w_1$  die Variable, nach der abgeleitet wird, und  $h_0$  ist konstant, also

$$\frac{d}{dw_1} h_1 = \sigma'(w_1 h_0) \cdot h_0.$$

Insgesamt erhalten wir

$$\frac{d}{dw_1} E = (h_L - y) \cdot \underbrace{\left[ \prod_{k=2}^L w_k \sigma'(w_k h_{k-1}) \right]}_{k\text{-te Schicht}} \cdot \underbrace{h_0 \sigma'(w_1 h_0)}_{\text{erste Schicht}}.$$

(d) Aus Teil c) folgt

$$\left| \frac{dE}{dw_1} \right| = |h_L - y| \cdot \left| \prod_{k=2}^L w_k \sigma'(w_k h_{k-1}) \right| \cdot |h_0 \sigma'(w_1 h_0)|.$$

Setzen wir also die gegebenen Schranken, sowie Teil b) ( $|\sigma'(s)| \leq \frac{1}{4}$ ) ein, so ergibt sich

$$\left| \frac{dE}{dw_1} \right| \leq 1 \cdot \left[ \prod_{k=2}^L \left( 1 \cdot \frac{1}{4} \right) \right] \cdot \left( 1 \cdot \frac{1}{4} \right) = \left( \frac{1}{4} \right)^{L-1} \cdot \frac{1}{4} = \frac{1}{4^L}.$$

Der Tikz-Code der Grafiken wurde mithilfe von *Gemini 3 Pro* erstellt.