Die Kettenregel

Seien $U \subseteq \mathbb{R}^m$ und $V \subseteq \mathbb{R}^n$ offen,
$g : U \to V$ differenzierbar in $x \in U$, und
$f : V \to \mathbb{R}^l$ differenzierbar in $g(x) \in V$.
Dann ist auch $f \circ g$ differenzierbar in x, und es gilt:

$$\boxed{D(f \circ g)(x) = Df(g(x)) \cdot Dg(x)}$$

Beweisstruktur:

1. Darstellung von f und g durch Linearisierung inklusive Fehlertermen

2. Einsetzen dieser Darstellungen ineinander

3. Kontrolle der auftretenden Fehlerterme

Beweis der Kettenregel: Wir versehen $\mathbb{R}^m$, $\mathbb{R}^n$ und $\mathbb{R}^l$ mit je einer Norm $\|\cdot\|$. Weil g differenzierbar in x ist, gilt

$$g(y) = g(x) + \underbrace{Dg(x)(y-x)}_{\text{lineare Näherung}} + \underbrace{\|y-x\|r(y)}_{\text{Fehlerterm}}$$

für $y \in U$, wobei $r(y) \xrightarrow{y \to x} 0$. Ebenso gilt, weil f differenzierbar in $g(x)$ ist:

$$f(v) = f(g(x)) + \underbrace{Df(g(x))(v-g(x))}_{\text{lineare Näherung}} + \underbrace{\|v-g(x)\|s(v)}_{\text{Fehlerterm}}$$

für $v \in V$, wobei $s(v) \xrightarrow{v \to g(x)} 0$. Durch Einsetzen von $v = g(y)$ erhalten wir:

$$\begin{aligned} f(g(y)) = f(g(x)) &+ Df(g(x))[\underbrace{Dg(x)(y-x)}_{\text{lineare Näherung}} + \underbrace{\|y-x\|r(y)}_{\text{Fehlerterm}}] \\ &+ \underbrace{\|g(y)-g(x)\|s(g(y))}_{\text{Fehlerterm}} \\ &= f(g(x)) + \underbrace{Df(g(x))\,Dg(x)\,(y-x)}_{\text{lineare Näherung}} + \underbrace{h(y)}_{\text{Fehlerterm}}, \end{aligned}$$

mit dem Fehlerterm

$$h(y) = \|y-x\|Df(g(x))r(y) + \|g(y)-g(x)\|s(g(y)).$$

Nun gilt

$$Df(g(x))\,r(y) \xrightarrow{y \to x} 0$$

wegen $r(y) \xrightarrow{y \to x} 0$, und

$$\begin{aligned} \|g(y)-g(x)\| &= \Big\| Dg(x)(y-x) + \|y-x\|r(y) \Big\| \\ &\le \|Dg(x)(y-x)\| + \|y-x\| \cdot \|r(y)\| = O(\|y-x\|) \quad \text{für } y \to x, \end{aligned}$$

sowie

$$s(g(y)) \xrightarrow{y \to x} 0,$$

weil $g(y) \xrightarrow{y \to x} g(x)$ und $s(v) \xrightarrow{v \to g(x)} 0$. Es folgt

$$\begin{aligned} \|h(y)\| &\le \|y-x\| \cdot \|Df(g(x))\,r(y)\| \\ &\quad + \Big[\|Dg(x)(y-x)\| + \|y-x\| \cdot \|r(y)\|\Big]\,\|s(g(y))\| \\ &= o(\|y-x\|) \quad \text{für } y \to x, \end{aligned}$$

also die Behauptung.