Zentrale Formel der Differentialrechnung: Die Kettenregel

Seien E,F,G Banachräume, $U \subset E$ offen, $V \subset F$ offen $f: U \to F$, $g: V \to G$ Abbildungen mit $f(U) \subset V$, es sei $x_0 \in U$, $y_0 = f(x_0) \in V$, f sei in x_0 und g sei in $y_0 = f(x_0)$ differenzierbar; dabei sei $Df(x_0) = \varphi$ und $Dg(y_0) = \psi$.

Dann ist auch die Verknüpfung $g \circ f: U \to G$ in x_0 differenzierbar, und es gilt:

$$D(g \circ f)(x_0) = Dg\big(f(x_0)\big) \circ Df(x_0)$$

M.a.W.: die Ableitung der Verknüpfung ist gleich der Verknüpfung der Ableitungen. (Was sonst?)

Beweis:
Sei $Df(x_0) =: \varphi$, $Dg(y_0) =: \psi$. Lt. Voraussetzung gilt:
$\forall \epsilon > 0 \, \exists \delta_1 > 0 \, \forall h \in E: \|h\| < \delta_1 \;\to\; x_0 + h \in U$ und $\|f(x_0+h) - f(x_0) - \varphi(h)\| < \epsilon \|h\|$
$\forall \epsilon > 0 \, \exists \delta_2 > 0 \, \forall \eta \in F: \|\eta\| < \delta_2 \;\to\; y_0 + \eta \in V$ und $\|g(y_0+\eta) - g(y_0) - \psi(\eta)\| < \epsilon \|\eta\|$

Zu vorgegebenem $\epsilon > 0$ ist jetzt ein $\delta > 0$ zu konstruieren, so daß
$\|h\| < \delta \;\to\; x_0 + h \in U$ und $\|(g \circ f)(x_0+h) - (g \circ f)(x_0) - (\psi \circ \varphi)(h)\| < \epsilon \|h\|$

Klar, daß $x_0 + h \in U$ gelten wird, wenn man nur $\delta \leqslant \delta_1$ wählt.
Um herauszufinden, wie man δ wählen muß, damit auch die Abschätzung gilt, setzen wir
$\eta := f(x_0+h) - f(x_0)$, woraus $f(x_0+h) = y_0 + \eta$ folgt, und rechnen einfach los:

$$\|(g \circ f)(x_0+h) - (g \circ f)(x_0) - (\psi \circ \varphi)(h)\| = \|g\big(f(x_0+h)\big) - g\big(f(x_0)\big) - \psi\big(\varphi(h)\big)\| =$$
$$\|g(y_0+\eta) - g(y_0) - \psi(\eta) + \psi(\eta) - \psi\big(\varphi(h)\big)\| \leqslant \|g(y_0+\eta) - g(y_0) - \psi(\eta)\| + \|\psi(\eta) - \psi\big(\varphi(h)\big)\| =$$
$$\|g(y_0+\eta) - g(y_0) - \psi(\eta)\| + \|\psi\big(\eta - \varphi(h)\big)\| \leqslant \|g(y_0+\eta) - g(y_0) - \psi(\eta)\| + \|\psi\| \|\eta - \varphi(h)\|$$

Nun ist
$\eta - \varphi(h) = f(x_0+h) - f(x_0) - \varphi(h)$. Setzt man $\|h\| < \delta_1$ voraus, so erhält man
$\|\eta - \varphi(h)\| = \|f(x_0+h) - f(x_0) - \varphi(h)\| \leqslant \epsilon \|h\|$ und dann auch
$\|\eta\| \leqslant \|\eta - \varphi(h)\| + \|\varphi(h)\| = \|f(x_0+h) - f(x_0) - \varphi(h)\| + \|\varphi(h)\| \leqslant \epsilon \|h\| + \|\varphi\| \|h\| = (\epsilon + \|\varphi\|) \|h\|$

Wählt man nun $\|h\| < \min\left\{\delta_1, \frac{\delta_2}{\epsilon + \|\varphi\|}\right\}$, so ergibt sich $\|\eta\| < \delta_2$; wir können weiter abschätzen

$$\ldots \;\leqslant\; \epsilon \|\eta\| + \|\psi\| \|\eta - \varphi(h)\| \;\leqslant\; \epsilon(\epsilon + \|\varphi\|) \|h\| + \epsilon \|\psi\| \|h\| \;=\; \epsilon\big(\epsilon + \|\psi\| + \|\varphi\|\big) \|h\|$$

Damit sind wir im Prinzip fertig. Es wurde gezeigt:
$\forall \epsilon > 0 \, \exists \delta > 0 \, \forall h \in E: \|h\| < \delta \;\to\; x_0 + h \in U$ und
$$\|(g \circ f)(x_0+h) - (g \circ f)(x_0) - (\psi \circ \varphi)(h)\| \leqslant \epsilon\big(\epsilon + \|\psi\| + \|\varphi\|\big) \|h\|$$

Gibt man wiederum $\epsilon > 0$ vor, setzt $\tilde{\epsilon} := \min\left\{1, \frac{\epsilon}{1 + \|\psi\| + \|\varphi\|}\right\}$ und bestimmt δ zu $\tilde{\epsilon}$, so gilt für $\|h\| < \delta$ endlich $\|(g \circ f)(x_0+h) - (g \circ f)(x_0) - (\psi \circ \varphi)(h)\| \leqslant \tilde{\epsilon}\big(\tilde{\epsilon} + \|\psi\| + \|\varphi\|\big) \|h\| \leqslant \epsilon \|h\|$.