

## Übungen zur Analysis 2

*Aufgabe 12.1 ist die bereits angekündigte Fortsetzung von Aufgabe 11.6.*

**12.1 Optionspreise im Black-Scholes-Modell.** Eine “*europäische Call-Option*” ist ein Finanzderivat, das seinem Inhaber das Recht, aber nicht die Pflicht gibt, zu einem vereinbarten zukünftigen Zeitpunkt  $T$  eine Einheit eines vereinbarten “Basiswerts” (z.B. einer Aktie, einer Währung, eines Rohstoffs) zu einem vereinbarten Preis  $K$  zu kaufen. Die Option verfällt damit wertlos zur Zeit  $T$ , wenn der Marktpreis  $S_T$  des Basiswerts zur Zeit  $T$  kleiner oder gleich  $K$  ist; andernfalls hat sie dann den Wert  $S_T - K$ . Zur Zeit  $T$  besitzt die Call-Option also den Wert  $\max\{S_T - K, 0\}$ . Der Preis  $C$  einer solchen Call-Option zu einer früheren Zeit  $t < T$  wird vom Wert  $S$  des Basiswerts zu dieser Zeit abhängen:  $C = C(t, S)$ . Betrachten wir den mit dem Marktzins  $r$  diskontierten Optionspreis

$$L(t, x) := e^{-rt}C(t, S)$$

in Abhängigkeit vom logarithmierten diskontierten Basiswertpreis

$$x := \log(e^{-rt}S).$$

In einem berühmten finanzmathematischen Modell, dem *Black-Scholes Modell*, wird der diskontierte Preis  $L$  durch die *Rückwärts-Wärmeleitungsgleichung mit Drift*

$$\frac{\partial L}{\partial t} + \frac{\sigma^2}{2} \frac{\partial^2 L}{\partial x^2} - \frac{\sigma^2}{2} \frac{\partial L}{\partial x} = 0 \quad (1)$$

für  $t < T$  und  $x \in \mathbb{R}$  beschrieben. Hierbei ist  $\sigma > 0$  ein Modellparameter, die “*Volatilität*” des Basiswerts, der das Ausmaß der stochastischen Fluktuationen des Basiswerts quantifiziert. (Die zugrundeliegenden Modellannahmen und eine Herleitung der Gleichung (1) aus diesen Annahmen mit Methoden der Stochastik können Sie in Vorlesungen zur stochastischen Analysis und Finanzmathematik lernen.)

- (a) **Driftentfernung, Zeitskalierung und Zeitumkehr.** Zeigen Sie, dass mit der affin-linearen Transformation

$$f(\tilde{t}, \tilde{x}) := L(T - \sigma^{-2}\tilde{t}, \tilde{x} + \tilde{t}/2) \quad \text{mit } \tilde{t} > 0, \tilde{x} \in \mathbb{R}$$

die obige Rückwärts-Wärmeleitungsgleichung (1) äquivalent zur einfachen Wärmeleitungsgleichung

$$\frac{\partial f}{\partial \tilde{t}} = \frac{1}{2} \frac{\partial^2 f}{\partial \tilde{x}^2}$$

wird.

- (b) **Black-Scholes Differentialgleichung.** Folgern Sie aus der Rückwärts-Wärmeleitungsgleichung (1) die partielle Differentialgleichung

$$\frac{\partial C}{\partial t} + rS \frac{\partial C}{\partial S} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 C}{\partial S^2} - rC = 0.$$

Sie wird "Black-Scholes Differentialgleichung" genannt.

- (c) **Black-Scholes-Preise europäischer Call-Optionen.** Finden Sie eine Lösung  $C(t, S)$ ,  $t < T$ ,  $S > 0$  der Black-Scholes Differentialgleichung mit der Randbedingung

$$C(t, S) \rightarrow \max\{S_T - K, 0\} \text{ für } (t, S) \rightarrow (T, S_T), t < T,$$

indem Sie die Lösung

$$f(\tilde{t}, \tilde{x}) = e^{\tilde{x} + \frac{\tilde{t}}{2}} \Phi(\tilde{t}^{-1/2}(\tilde{x} + \tilde{t} - \log K)) - K \Phi(\tilde{t}^{-1/2}(\tilde{x} - \log K))$$

der Wärmeleitungsgleichung aus Aufgabe 11.6 (a) transformieren. Überprüfen zur Probe mit einer direkten Rechnung, dass Ihr Ergebnis wirklich die Black-Scholes Differentialgleichung löst.

- (d) **Veranschaulichung des Ergebnisses.** Veranschaulichen Sie sich diese Optionspreise  $C(t, S)$  im Black-Scholes-Modell für den Fall  $r = 0$ ,  $T = 0$ ,  $\sigma = 1$  und  $K = 1$ , indem Sie den Graphen von  $S \mapsto C(t, S)$  für verschiedene  $t < 0$  in ein einziges  $S$ - $C$ -Diagramm skizzieren. (Eine qualitative Skizze genügt.) [10 Punkte]

## Lösung

- (a) Setzen wir  $k(\tilde{t}, \tilde{x}) = (T - \sigma^{-2}\tilde{t}, \tilde{x} + \tilde{t}/2)$ , dann folgt aus der Kettenregel

$$\frac{\partial f}{\partial \tilde{t}}(\tilde{t}, \tilde{x}) = \frac{\partial}{\partial \tilde{t}} L(k(\tilde{t}, \tilde{x})) = -\sigma^{-2} D_1 L(k(\tilde{t}, \tilde{x})) + \frac{1}{2} D_2 L(k(\tilde{t}, \tilde{x})), \quad (*)$$

sowie, wegen  $D_2 k = (0, 1)$ ,

$$\frac{\partial f}{\partial \tilde{x}}(\tilde{t}, \tilde{x}) = \frac{\partial}{\partial \tilde{x}} L(k(\tilde{t}, \tilde{x})) = D_2 L(k(\tilde{t}, \tilde{x}))$$

$$\frac{\partial^2 f}{\partial \tilde{x}^2}(\tilde{t}, \tilde{x}) = \frac{\partial}{\partial \tilde{x}} D_2 L(k(\tilde{t}, \tilde{x})) = D_2^2 L(k(\tilde{t}, \tilde{x})). \quad (**)$$

Angenommen  $L$  ist eine Lösung der Rückwärts-Wärmeleitungsgleichung mit Drift

$$D_1 L = \frac{\sigma^2}{2} D_2 L - \frac{\sigma^2}{2} D_2^2 L, \quad (1)$$

dann

$$\frac{\partial f}{\partial \tilde{t}} \stackrel{(*),(1)}{=} \frac{1}{2} (D_2^2 L \circ k - D_2 L \circ k + D_2 L \circ k) = \frac{1}{2} D_2^2 L \circ k \stackrel{(**)}{=} \frac{1}{2} \frac{\partial^2 f}{\partial \tilde{x}^2}.$$

(b) Wegen  $C(t, S) = e^{rt} L(t, x(t, S))$  mit  $x(t, S) = -rt + \log S$ , gilt

$$D_1 C(t, S) = r e^{rt} L(t, x(t, S)) + e^{rt} D_1 L(t, x(t, S)) - r e^{rt} D_2 L(t, x(t, S)) \quad (2)$$

$$D_2 C(t, S) = e^{rt} D_2 L(t, x(t, S)) D_2 x(t, S) = e^{rt} S^{-1} D_2 L(t, x(t, S)) \quad (3)$$

$$D_2^2 C(t, S) = e^{rt} (-S^{-2} D_2 L(t, x(t, S)) + S^{-2} D_2^2 L(t, x(t, S))) . \quad (4)$$

Und damit

$$\begin{aligned} D_1 C &\stackrel{(2),(3)}{=} rC + e^{rt} D_1 L - rS D_2 C \stackrel{(1)}{=} rC - rS D_2 C + e^{rt} \left( \frac{\sigma^2}{2} D_2 L - \frac{\sigma^2}{2} D_2^2 L \right) \\ &\stackrel{(4)}{=} rC - rS D_2 C - \frac{\sigma^2}{2} S^2 D_2^2 C , \end{aligned}$$

was die gesuchte Black-Scholes Differentialgleichung ist.

(c) Da  $f$  die einfache Wärmeleitungsgleichung löst, ist wegen Teil (a) und (b),

$$C(t, S) = e^{rt} L(t, x(t, S)) = e^{rt} f\left(\sigma^2(T-t), x(t, S) - \frac{\sigma^2}{2}(T-t)\right)$$

eine Lösung der Black-Scholes Differentialgleichung, wobei wir die Umkehrung der Transformation  $(\tilde{t}, \tilde{x}) \mapsto (t, x) = (T - \sigma^{-2}\tilde{t}, \tilde{x} + \tilde{t}/2)$  gebildet haben, um  $f(\tilde{t}, \tilde{x})$  in den Variablen  $t$  und  $x$  auszudrücken. Außerdem gilt für  $(t, S) \rightarrow (T, S_T)$ ,

$$\sigma^2(T-t) \rightarrow 0 , \quad x(t, S) - \frac{\sigma^2}{2}(T-t) \rightarrow x_0$$

wobei  $x_0 := x(T, S_T) = -rT + \log S_T$ . Da  $f(t, x) \rightarrow g(x_0) = \max\{e^{x_0} - \tilde{K}, 0\}$ , (siehe Aufgabe 11.6) falls  $(t, x) \rightarrow (0, x_0)$  mit  $t > 0$ , so gilt demnach

$$C(t, S) \rightarrow e^{rT} \max\{e^{x_0} - \tilde{K}, 0\} = \max\{S_T - K, 0\} , \quad \text{falls } (t, S) \rightarrow (T, S_T) ,$$

wobei wir  $e^{x(T, S_T)} = e^{-rT} S_T$  und  $\tilde{K} = e^{-rT} K$  benutzt haben. Explizit gilt

$$\begin{aligned} C(t, S) &= e^{rt+x(t,S)} \Phi\left(\frac{x(t, S) + \frac{\sigma^2}{2}(T-t) + rT - \log K}{\sigma(T-t)^{\frac{1}{2}}}\right) \\ &\quad - e^{rt} \tilde{K} \Phi\left(\frac{x(t, S) - \frac{\sigma^2}{2}(T-t) + rT - \log K}{\sigma(T-t)^{\frac{1}{2}}}\right) \\ &= S \Phi\left(\frac{(r + \frac{\sigma^2}{2})(T-t) + \log(S/K)}{\sigma(T-t)^{\frac{1}{2}}}\right) \\ &\quad - e^{-r(T-t)} K \Phi\left(\frac{(r - \frac{\sigma^2}{2})(T-t) + \log(S/K)}{\sigma(T-t)^{\frac{1}{2}}}\right) \\ &= S \Phi(g_+(t, S)) - K e^{-r(T-t)} \Phi(g_-(t, S)) , \end{aligned} \quad (5)$$

wobei  $\Phi$  die Verteilungsfunktion der Standardnormalverteilung ist, d.h.

$$\Phi(x) = \mathcal{N}_{0,1}((-\infty, x]) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du ,$$

und

$$g_{\pm}(t, S) := \frac{(r \pm \frac{\sigma^2}{2})(T-t) + \log(S/K)}{\sigma(T-t)^{\frac{1}{2}}} = (r\sigma^{-1} \pm \frac{\sigma}{2})(T-t)^{\frac{1}{2}} + \sigma^{-1} \log(S/K)(T-t)^{-\frac{1}{2}} .$$

Es folgt

$$\begin{aligned}\partial_t C(t, S) &= S \Phi'(g_+(t, S)) \partial_t g_+(t, S) - K r e^{-r(T-t)} \Phi(g_-(t, S)) \\ &\quad - K e^{-r(T-t)} \Phi'(g_-(t, S)) \partial_t g_-(t, S) \\ \partial_S C(t, S) &= \Phi(g_+(t, S)) + S \Phi'(g_+(t, S)) \partial_S g_+(t, S) \\ &\quad - K e^{-r(T-t)} \Phi'(g_-(t, S)) \partial_S g_-(t, S)\end{aligned}$$

Für die partiellen Ableitungen der Funktionen  $g_{\pm}$  erhalten wir

$$\begin{aligned}\partial_t g_{\pm}(t, S) &= -\frac{1}{2}(r\sigma^{-1} \pm \frac{\sigma}{2})(T-t)^{-\frac{1}{2}} + \frac{\sigma^{-1}}{2} \log(S/K)(T-t)^{-\frac{3}{2}} \\ \partial_S g_{\pm}(t, S) &= \sigma^{-1} S^{-1}(T-t)^{-\frac{1}{2}}\end{aligned}$$

Wegen  $\partial_S g_+ = \partial_S g_-$  und<sup>1</sup>  $K e^{-r(T-t)} \Phi'(g_-(t, S)) = S \Phi'(g_+(t, S))$  ergibt sich

$$\begin{aligned}\partial_t C(t, S) &= -S \Phi'(g_+(t, S)) \frac{\sigma}{2}(T-t)^{-\frac{1}{2}} - K r e^{-r(T-t)} \Phi(g_-(t, S)) \\ \partial_S C(t, S) &= \Phi(g_+(t, S)) \\ \partial_S^2 C(t, S) &= \Phi'(g_+(t, S)) \partial_S g_+(t, S).\end{aligned}$$

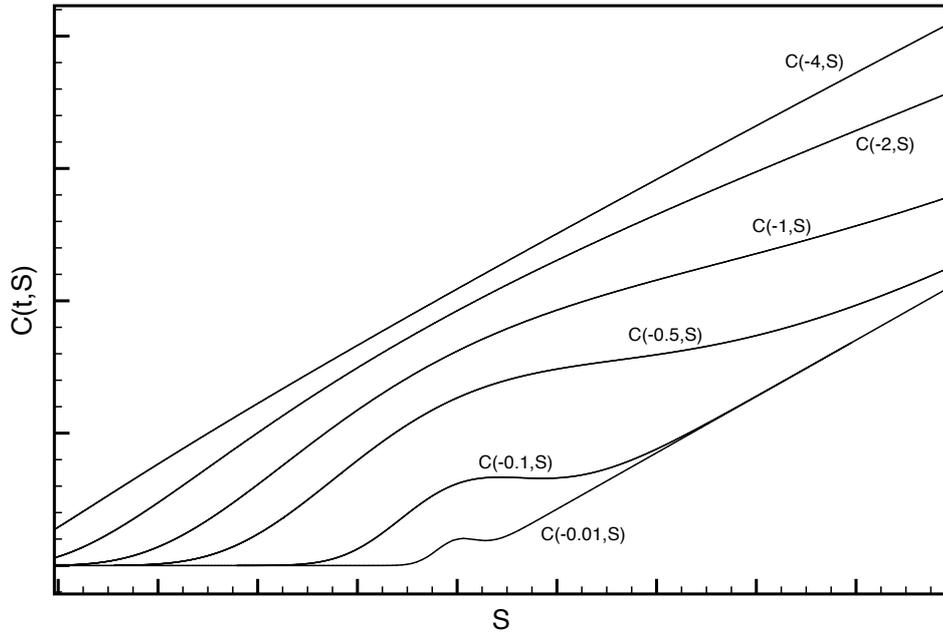
Es gilt also

$$\begin{aligned}rS \partial_S C(t, S) + \frac{\sigma^2}{2} S^2 \partial_S^2 C(t, S) &= rS \Phi(g_+(t, S)) + S \Phi'(g_+(t, S)) \frac{\sigma}{2}(T-t)^{-\frac{1}{2}} \\ rC(t, S) - \partial_t C(t, S) &= rS \Phi(g_+(t, S)) + S \Phi'(g_+(t, S)) \frac{\sigma}{2}(T-t)^{-\frac{1}{2}},\end{aligned}$$

d.h.  $C$  aus (5) löst in der Tat die Black-Scholes Differentialgleichung.

(d) Einsetzen von  $r = 0$ ,  $T = 0$ ,  $\sigma = 1$  und  $K = 1$  in (5) ergibt

$$C(t, S) = S \Phi\left(\frac{1}{2}|t|^{\frac{1}{2}} + \log(S) |t|^{-\frac{1}{2}}\right) - \Phi\left(-\frac{1}{2}|t|^{\frac{1}{2}} + \log(S) |t|^{-\frac{1}{2}}\right)$$



<sup>1</sup>Dies folgt direkt durch Einsetzen von  $g_+$  und  $g_-$  in  $\Phi'(x) = (2\pi)^{-1/2} e^{-x^2/2}$ .

## 12.2 Newtonverfahren und vereinfachtes Newtonverfahren.

Lösen Sie das Gleichungssystem

$$r \cos \phi = x$$

$$r \sin \phi = y$$

für die rechte Seite  $(x, y) = (1.1, 0.1)$  näherungsweise numerisch mit dem Taschenrechner (oder, noch besser, mit einem Computer)

- (a) mit dem vereinfachten Newtonverfahren zum Start- und Referenzpunkt  $(r_0, \phi_0) = (1, 0)$ ,  
 (b) mit dem Newtonverfahren zum gleichen Startpunkt.

Führen Sie jeweils einige Iterationen, z.B. drei, aus. Beobachten Sie, wie schnell sich die berechneten Näherungen an die exakte Lösung  $(r, \phi) = (\sqrt{x^2 + y^2}, \arctan(y/x))$  annähern, aber auch, welcher Rechenaufwand jeweils nötig ist. [10 Punkte]

### Lösung

- (a) Nach Bemerkung 2.63 gilt es  $(r, \phi)^{(n+1)} = (r, \phi)^{(n)} - df_{(r, \phi)_0}^{-1}(f((r, \phi)^{(n)}) - (1, 1; 0, 1))$ ,  $n \in \mathbb{N}_0$  zu berechnen mit  $r_0, \phi_0$  s. d.  $f(r_0, \phi_0)$  genügend nahe bei  $(1, 1; 0, 1)$ .<sup>2</sup> Wir wissen bereits (z. B. aus Aufgabe 10.4), dass

$$df_{(r, \phi)} = \begin{pmatrix} \cos(\phi) & -r \sin(\phi) \\ \sin(\phi) & r \cos(\phi) \end{pmatrix} \Rightarrow \det(df) = r, df_{(r, \phi)}^{-1} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -r^{-1} \sin(\phi) & r^{-1} \cos(\phi) \end{pmatrix}.$$

Wähle  $\begin{pmatrix} r \\ \phi \end{pmatrix}^{(0)} = \begin{pmatrix} r \\ \phi \end{pmatrix}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{(1)}$ . Es folgt

$$\begin{pmatrix} r \\ \phi \end{pmatrix}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1^{-1} \end{pmatrix} \begin{pmatrix} 1 - 1, 1 \\ 0 - 0, 1 \end{pmatrix} = \begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix},$$

$$\begin{pmatrix} r \\ \phi \end{pmatrix}^{(2)} = \begin{pmatrix} 1, 1055 \\ 0, 0902 \end{pmatrix}, \begin{pmatrix} r \\ \phi \end{pmatrix}^{(3)} = \begin{pmatrix} 1, 1045 \\ 0, 0906 \end{pmatrix}.$$

Dieser gerundete Wert entspricht fast dem gerundeten Werte der exakten Lösung  $(r, \phi) = (\sqrt{x^2 + y^2}, \arctan(y/x)) \sim (1, 1045; 0, 0907)$ . Allgemein gilt für den Fehler

$$\begin{aligned} \varepsilon^{(n+1)} &:= \begin{pmatrix} r \\ \phi \end{pmatrix} - \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n+1)} \\ &= \varepsilon^{(n)} - \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n)} + df_{(r, \phi)_0}^{-1} \left( f \left( \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n)} \right) - f \left( \begin{pmatrix} r \\ \phi \end{pmatrix} \right) \right) \\ &= \varepsilon^{(n)} + df_{(r, \phi)_0}^{-1} \left( (df_{(r, \phi)_0} \varepsilon^{(n)}) + \mathcal{O}(\|\varepsilon^{(n)}\|^2) \right) \\ &\Rightarrow \frac{\|\varepsilon^{(n+1)}\|}{\|\varepsilon^{(n)}\|} \leq \left\| df_{(r, \phi)_0}^{-1} df_{(r, \phi)_0} \right\|_{\mathcal{O}_p} + \mathcal{O}(\|\varepsilon^{(n)}\|). \end{aligned}$$

<sup>2</sup>Kriterien für „nahe genug“ werden nach der Fehlerberechnung angegeben.

mit der induzierten Matrixnorm  $\|\cdot\|_{Op}$ .<sup>3</sup> Hierbei haben wir die mehrdimensionale Taylorformel verwendet. Wir sehen, dass das Verfahren linear konvergiert wenn das Taylorrestglied  $R_2$  klein genug ist auf unserer Umgebung  $U$  mit  $(r, \phi), (r, \phi)_0 \in U$ . Der Fehler bzgl. Supremumsnorm ist

$$\begin{aligned} \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(0)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} &= 0, 1045, \quad \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(1)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} &= 0, 0093, \\ \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(2)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} &= 0, 0010, \quad \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(3)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} &< 0, 0001. \end{aligned}$$

Wir ordnen dem Auswerten der trigonometrischen Funktionen die Einheit Tri und jeder der Rechenoperationen Addition, Multiplikation, Subtraktion, Division ein FLOP<sup>4</sup> zu. Damit erhalten wir für  $(f((r, \phi)^{(n)} - (1, 1; 0, 1)))$  4FLOP + 2Tri Operationen, für die abschließende Addition weitere 2FLOP und für die Matrixmultiplikation jeweils 6FLOP. Die Berechnung der Matrix liefert schließlich einmalig 2 weitere Operationen.<sup>5</sup> Zusammen erhalten wir  $12n\text{FLOP} + 2n\text{Tri} + \mathcal{O}(1)$ .

- (b) Nach Bemerkung 2.63 sei  $(r, \phi)^{(n+1)} = (r, \phi)^{(n)} - df_{(r, \phi)^{(n)}}^{-1}(f((r, \phi)^{(n)} - (1, 1; 0, 1)))$ ,  $n \in \mathbb{N}_0$ . Wähle wieder  $\begin{pmatrix} r \\ \phi \end{pmatrix}^{(0)} = \begin{pmatrix} r \\ \phi \end{pmatrix}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^{(1)}$ :

$$\begin{pmatrix} r \\ \phi \end{pmatrix}^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1^{-1} \end{pmatrix} \begin{pmatrix} 1 - 1, 1 \\ 0 - 0, 1 \end{pmatrix} = \begin{pmatrix} 1, 1 \\ 0, 1 \end{pmatrix},$$

$$\begin{pmatrix} r \\ \phi \end{pmatrix}^{(2)} = \begin{pmatrix} 1, 1045 \\ 0, 0906 \end{pmatrix}, \quad \begin{pmatrix} r \\ \phi \end{pmatrix}^{(3)} = \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix}.$$

Dieser gerundete Wert entspricht dem gerundeten Werte der exakten Lösung  $(r, \phi) = (\sqrt{x^2 + y^2}, \arctan(y/x)) \sim (1, 1045; 0, 0907)$ . Allgemein gilt für den Fehler

$$\begin{aligned} \varepsilon^{(n+1)} &:= \begin{pmatrix} r \\ \phi \end{pmatrix} - \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n+1)} \\ &= \varepsilon^{(n)} - \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n)} + df_{(r, \phi)^{(n)}}^{-1} \left( f \left( \begin{pmatrix} r \\ \phi \end{pmatrix}^{(n)} \right) - f \left( \begin{pmatrix} r \\ \phi \end{pmatrix} \right) \right) \\ &= \varepsilon^{(n)} - \varepsilon^{(n)} + \frac{df_{(r, \phi)^{(n)}}^{-1}(\varepsilon^{(n)})^t d_{((r, \phi)^{(n)})^t}^2 f \varepsilon^{(n)}}{2} + \mathcal{O}(\|\varepsilon^{(n)}\|^3) \\ \Rightarrow \frac{\|\varepsilon^{(n+1)}\|}{\|\varepsilon^{(n)}\|^2} &\leq \left\| df_{(r, \phi)_0}^{-1} \right\|_{Op} \left\| \frac{df_{(r, \phi)^{(n)}}^{-1} d_{((r, \phi)^{(n)})^t}^2 f}{2} \right\|_{Op} + \mathcal{O}(\|\varepsilon^{(n)}\|), \end{aligned}$$

wobei wir jeweils um  $((r, \phi)^{(n)})^t$  entwickelt haben. Wir erhalten quadratische Konvergenz für  $((r, \phi)^{(n)})^t$  nahe genug bei  $(\sqrt{x^2 + y^2}, \arctan(y/x))$ .

Für den Fehler bzgl. Supremumsnorm gilt

$$\left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(0)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} = 0, 1045, \quad \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(1)} - \begin{pmatrix} 1, 1045 \\ 0, 0907 \end{pmatrix} \right\|_{\infty} = 0, 0093,$$

<sup>3</sup>Vgl. Operatornorm.

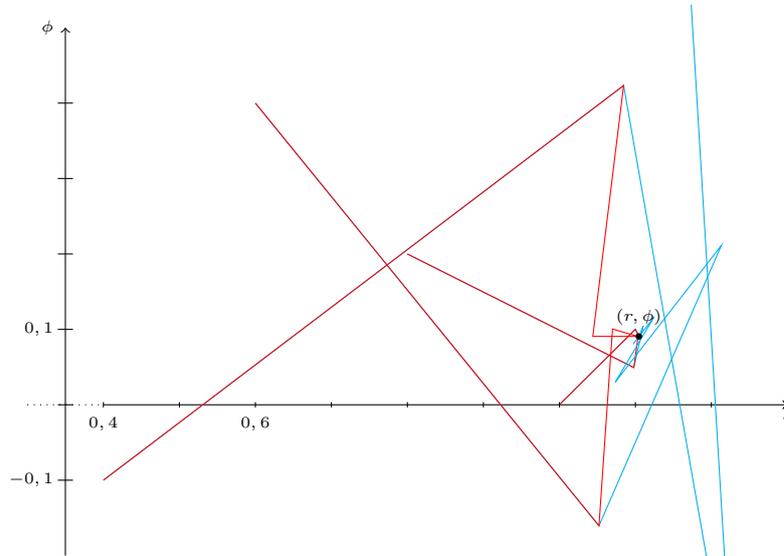
<sup>4</sup>FLOP = floating-point operation. Multiplikationen mit  $\pm 1$  werden nicht gewertet.

<sup>5</sup>Vorausgesetzt wir speichern die Werte der trigonometrischen Funktionen.

$$\left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(2)} - \begin{pmatrix} 1,1045 \\ 0,0907 \end{pmatrix} \right\|_{\infty} = 0,0001, \quad \left\| \begin{pmatrix} r \\ \phi \end{pmatrix}^{(3)} - \begin{pmatrix} 1,1045 \\ 0,0907 \end{pmatrix} \right\|_{\infty} < 0,0001.$$

Wie zuvor berechnet sich der Rechenaufwand jetzt zu  $14n\text{FLOP} + 2n\text{Tri} + \mathcal{O}(1)$ .

Man beachte, dass der Rechenaufwand für unsere spezielle Funktion  $f$  im Übergang zum vereinfachten Newtonverfahren nur leicht sinkt. Für andere Funktionen  $f$ , insbesondere auf Räumen höherer Dimension, ergibt sich aber ein wesentlich größerer Unterschied.<sup>6</sup>



**Fig. 1:** Die roten Polygonzüge zeigen die Annäherung mit Newton-Verfahren an die Lösung, die blauen Polygonzüge den abweichenden Verlauf des vereinfachten Newton-Verfahrens nach der ersten Iteration.

### 12.3 Quadratwurzel von Matrizen und Operatoren. Bearbeiten Sie entweder Teil (a) oder Teil (b) der folgenden Aufgabe:

(a) Zeigen Sie, daß die Abbildung

$$Q : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}, \quad Q(A) = A^2$$

stetig differenzierbar mit der Ableitung  $dQ_A(B) = A \cdot B + B \cdot A$  für  $A, B \in \mathbb{R}^{n \times n}$  ist. Folgern Sie, dass es eine stetig differenzierbare, in einer offenen Umgebung  $V_1$  von  $\text{Id} \in \mathbb{R}^{n \times n}$  definierte Abbildung  $\text{sqr}t : V_1 \rightarrow \mathbb{R}^{n \times n}$ ,  $A \mapsto \sqrt{A}$ , mit  $\sqrt{\text{Id}} = \text{Id}$  gibt, die  $Q$  genügend nahe bei  $\text{Id}$  invertiert. Zeigen Sie  $d\text{sqr}t_{\text{Id}}(A) = \frac{1}{2}A$  für  $A \in \mathbb{R}^{n \times n}$ .

(b) Es sei  $(U, \|\cdot\|)$  ein Banachraum. Zeigen Sie, dass die Abbildung

$$Q : (\mathcal{B}(U, U), \|\cdot\|_{U \rightarrow U}) \rightarrow (\mathcal{B}(U, U), \|\cdot\|_{U \rightarrow U}), \quad Q(A) = A \circ A$$

stetig differenzierbar mit der Ableitung  $dQ_A(B) = A \circ B + B \circ A$  für  $A, B \in \mathcal{B}(U, U)$  ist. Folgern Sie, dass es eine stetig differenzierbare, in einer offenen Umgebung  $V_1$  von  $\text{id}_U \in \mathcal{B}(U, U)$  definierte Abbildung  $\text{sqr}t : V_1 \rightarrow \mathcal{B}(U, U)$ ,  $A \mapsto \sqrt{A}$ , mit  $\sqrt{\text{id}_U} = \text{id}_U$  gibt, die  $Q$  genügend nahe bei  $\text{id}_U$  invertiert. Zeigen Sie  $d\text{sqr}t_{\text{id}_U}(A) = \frac{1}{2}A$  für  $A \in \mathcal{B}(U, U)$ . **[10 Punkte]**

<sup>6</sup>Die Berechnung der Inversen einer  $k \times k$ -Matrix mit den uns bekannten Verfahren erfordert i. A.  $\mathcal{O}(k^3)$  Operationen.

## Solution

(a) We can calculate the derivative of  $Q$  either by the definition, i.e.

$$DQ(A)B = \lim_{t \rightarrow 0} \frac{Q(A+tB) - Q(A)}{t} = \lim_{t \rightarrow 0} \frac{(A+tB)^2 - A^2}{t} = AB + BA,$$

or using the Leibniz rule, i.e.  $Q(A) = I(A)I(A)$  where  $I : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  is the identity map. Hence,

$$DQ_A B = DI_A B \times I(A) + I(A) \times DI_A B = BA + AB.$$

Therefore,  $DQ(A)B = AB + BA$  and in particular  $DQ(I) = 2I$  which is invertible. By the theorem of inverse local, there exist an open  $U_1$  around  $I$  and  $V_1$  around  $Q(I) = I$  such that  $Q : U_1 \rightarrow V_1$  is invertible. We denote the inverse of  $Q$  by  $\text{sqrt} : V_1 \rightarrow U_1$ .  $\text{sqrt} \circ Q(A) = A$ . Thus,  $D\text{sqrt}(Q(A))DQ(A)(B) = B$  and in particular  $D\text{sqrt}(Q(I))DQ(I)(B) = D\text{sqrt}(I)(2B) = B$ . Therefore,

$$D\text{sqrt}(I)(B) = \frac{1}{2}B.$$

(b) Similarly as previous part,

$$\begin{aligned} DQ_A B &= \lim_{t \rightarrow 0} \frac{Q(A+tB) - Q(A)}{t} = \lim_{t \rightarrow 0} \frac{(A+tB) \circ (A+tB) - A \circ A}{t} \\ &= \lim_{t \rightarrow 0} \frac{t(A \circ B + B \circ A) + t^2 B \circ B}{t} = A \circ B + B \circ A, \end{aligned}$$

and in particular,  $DQ_I A = 2A$ . Hence, the local inverse theorem implies that there exist open subsets  $U_1$  and  $V_1$  around  $I$  and  $Q(I) = I$  such that  $Q : U_1 \rightarrow V_1$  is invertible and the derivative of  $Q^{-1} = \text{sqrt} : V_1 \rightarrow U_1$  is

$$d\text{sqrt}_A = dQ_{A \circ A}^{-1},$$

and in particular  $d\text{sqrt}_I(A) = dQ_I^{-1}(A) = \frac{1}{2}A$ .

**12.4** Zeigen Sie mit dem Banachschen Fixpunktsatz und dem Satz von den impliziten Funktionen, dass durch das Gleichungssystem

$$\begin{aligned} g_1(x_1, x_2) &= \frac{1}{10} \arctan(g_1(x_1, x_2) + g_2(x_1, x_2) + x_1) \\ g_2(x_1, x_2) &= \frac{1}{20} \arctan(g_1(x_1, x_2) - g_2(x_1, x_2) + x_2) \end{aligned}$$

es eine eindeutige glatte Funktion  $g = (g_1, g_2) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  global implizit definiert wird. Berechnen Sie  $dg_1(0, 0)$  und  $dg_2(0, 0)$ . **[10 Punkte]**

## Solution

We consider the following system of equation:

$$\begin{aligned} y_1 &= \frac{1}{10} \arctan(y_1 + y_2 + x_1) \\ y_2 &= \frac{1}{20} \arctan(y_1 - y_2 + x_2). \end{aligned} \tag{*}$$

Define  $F : \mathbb{R}^4 \rightarrow \mathbb{R}^2$  as

$$F(x_1, x_2, y_1, y_2) = \left( \frac{1}{10} \arctan(y_1 + y_2 + x_1) - y_1, \frac{1}{20} \arctan(y_1 - y_2 + x_2) - y_2 \right).$$

Take  $(x_1, x_2, y_1, y_2) \in \mathbb{R}^4$  such that  $F(x_1, x_2, y_1, y_2) = 0$ . Then

$$D_2F_{(x_1, x_2, y_1, y_2)} = \begin{pmatrix} \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} - 1 & \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} \\ \frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} & -\frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} - 1 \end{pmatrix}$$

is invertible since

$$\det D_2F_{(x_1, x_2, y_1, y_2)} = 1 - \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} + \frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} - \frac{1}{100} \frac{1}{(1+(y_1-y_2+x_1)^2)(1+(y_1-y_2+x_1)^2)} \neq 0.$$

The implicit function theorem implies that for every  $(x_1, x_2) \in \mathbb{R}^2$  there exists an open subset  $U$  around  $(x_1, x_2)$  and a map

$$G^U : U \rightarrow \mathbb{R}^2 \\ (x_1, x_2) \mapsto (g_1(x_1, x_2), g_2(x_1, x_2))$$

such that  $F(x_1, x_2, G^U(x_1, x_2)) = 0$  and

$$DG_{(x_1, x_2)}^U = -D_2F_{(x, G^U(x))}^{-1} \cdot D_1F_{(x, G^U(x))}. \quad (**)$$

We show that  $G^U = G^V$  on  $U \cap V$  and we conclude that  $G(x_1, x_2) = G^U(x_1, x_2)$  for  $(x_1, x_2) \in U$  is well defined globally.

For every given  $(x_1, x_2) \in \mathbb{R}^2$  define

$$H(y_1, y_2) := \left( \frac{1}{10} \arctan(y_1 + y_2 + x_1), \frac{1}{20} \arctan(y_1 - y_2 + x_2) \right).$$

We show that  $H$  is a contraction mapping on  $\mathbb{R}^2$  and therefore, the Banach fixed-point theorem implies that for every  $(x_1, x_2) \in \mathbb{R}^2$  the system of equation (\*) has a unique solution.

$$DH_{(y_1, y_2)} = \begin{pmatrix} \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} & \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} \\ \frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} & -\frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} \end{pmatrix}.$$

Hence for every  $y = (y_1, y_2)$ , we have (See Blatt 5, 5.4)

$$\|DH_y\|_{2 \rightarrow 2}^2 \leq 2 \left( \frac{1}{10} \frac{1}{1+(y_1+y_2+x_1)^2} \right)^2 + 2 \left( \frac{1}{20} \frac{1}{1+(y_1-y_2+x_1)^2} \right)^2 \leq \frac{1}{40}.$$

Thus, for every  $y_0 = (y_1^0, y_2^0)$ ,  $\|DH_{y_0}\| \leq \frac{1}{6}$  and there exists  $\varepsilon > 0$  small enough such that for every  $y \in U_\varepsilon(y_0)$

$$\begin{aligned} \|H(y) - H(y_0)\| &\leq \|DH_{y_0}(y - y_0)\| + O(\|y - y_0\|^2) \\ &\leq \|DH_{y_0}\|_{2 \rightarrow 2} \|y - y_0\| + O(\|y - y_0\|^2) \\ &\leq \frac{1}{6} \|y - y_0\| + O(\|y - y_0\|^2) \leq \frac{1}{2} \|y - y_0\|. \end{aligned}$$

Moreover, for a given  $(x_1, x_2) \in \mathbb{R}^2$ ,  $H$  is a contraction mapping on  $\mathbb{R}^2$ . Indeed, let  $y_0, y \in \mathbb{R}^2$ . Take a partition  $y_0, y_1, \dots, y_{n-1}, y_n = y$  on the segment joining  $y$  to  $y_0$  such that for every  $i = 0, \dots, n-1$

$$\|H(y_i) - H(y_{i-1})\| \leq \frac{1}{2} \|y_i - y_{i-1}\|.$$

Note that this partition exists according to what we have shown above. Thus,

$$\begin{aligned} \|H(y_0) - H(y)\| &\leq \|H(y_0) - H(y_1)\| + \dots + \|H(y_{n-1}) - H(y)\| \\ &\leq \frac{1}{2} (\|y_0 - y_1\| + \dots + \|y_{n-1} - y\|) = \frac{1}{2} \|y - y_0\|. \end{aligned}$$

In conclusion, let  $(x_1, x_2) \in U \cap V$  and,  $y = G^U(x_1, x_2)$  and  $y_0 = G^U(x_1, x_2)$  be the solution of the system of equation, i.e.  $H(y) = y$  and  $H(y_0) = y_0$ . We have

$$\|H(y_0) - H(y)\| = \|y - y_0\| \leq \frac{1}{2} \|y - y_0\|.$$

Therefore,  $y = y_0$  and in conclusion  $G$  is well defined. We now calculate  $dg_1(0, 0)$  and  $dg_2(0, 0)$  using equation (\*\*). Take  $(y_1, y_2) = (g_1(0, 0), g_2(0, 0))$ . The point  $(0, 0)$  is the fix point of  $H(y_1, y_2) = (\frac{1}{10} \arctan(y_1 + y_2 + x_1), \frac{1}{20} \arctan(y_1 - y_2 + x_2))$  and it is the only solution by the Banach fixed-point theorem. Therefore,

$$\begin{aligned} DG_{(0,0)} &= - \begin{pmatrix} \frac{1}{10} \frac{1}{1+(y_1+y_2)^2} - 1 & \frac{1}{10} \frac{1}{1+(y_1+y_2)^2} \\ \frac{1}{20} \frac{1}{1+(y_1-y_2)^2} & -\frac{1}{20} \frac{1}{1+(y_1-y_2)^2} - 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{10} \frac{1}{1+(y_1+y_2)^2} & \frac{1}{10} \frac{1}{1+(y_1+y_2)^2} \\ \frac{1}{20} \frac{1}{1+(y_1-y_2)^2} & -\frac{1}{20} \frac{1}{1+(y_1-y_2)^2} \end{pmatrix} \\ &= - \begin{pmatrix} \frac{9}{10} & \frac{1}{10} \\ \frac{1}{20} & -\frac{21}{20} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{20} & \frac{1}{20} \\ \frac{1}{20} & \frac{1}{20} \end{pmatrix} = -\frac{1}{190} \begin{pmatrix} -22 & -22 \\ 8 & 8 \end{pmatrix}. \end{aligned}$$

Thus,  $dg_1(0, 0) = \frac{22}{190}(dx_1 + dx_2)$  and  $dg_2(0, 0) = -\frac{8}{190}(dx_1 + dx_2)$ .

## 12.5 Ableitung der Matrix-Exponentialfunktion und Matrix-Logarithmus. Es sei $n \in \mathbb{N}$ .

- (a) Zeigen Sie für gegebenes  $j \in \mathbb{N}$ , dass die Matrix-Potenzfunktion  $f : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ ,  $f(A) = A^j$  die Ableitung  $df_A(B) = \sum_{k=1}^j A^{k-1} B A^{j-k}$  für  $A, B \in \mathbb{R}^{n \times n}$  besitzt.
- (b) Es sei  $U \subseteq \mathbb{R}^{n \times n}$  offen und beschränkt. Zeigen Sie, dass die Matrix-Exponentialreihe

$$U \ni A \mapsto \exp(A) = \sum_{j=1}^{\infty} \frac{1}{j} A^j$$

in  $(C_b^1(U, \mathbb{R}^{n \times n}), \|\cdot\|_{C^1})$  konvergiert.

- (c) Folgern Sie für  $A, B \in \mathbb{R}^{n \times n}$ :

$$d \exp_A(B) = \int_0^1 e^{tA} B e^{(1-t)A} dt.$$

Insbesondere gilt  $d \exp_{\text{Id}} = \text{id}_{\mathbb{R}^{n \times n}}$ .

- (d) Zeigen Sie, dass es eine offene Umgebung  $U$  von  $0 \in \mathbb{R}^{n \times n}$  gibt, so dass die Matrix-Exponentialfunktion  $\exp|_U : U \rightarrow V := \exp[U] \subseteq \mathbb{R}^{n \times n}$  ein Homöomorphismus mit einer stetig differenzierbaren Umkehrung  $\log : V \rightarrow U$  ist. Berechnen Sie  $d \log_{\text{Id}}$ . **[10 Punkte]**

12.5.a. Diese wunderschöne Aufgabe ist  
 ganz einfach und doch nicht ganz  
 einfach. Ich jedenfalls habe viel  
 gelernt (j).

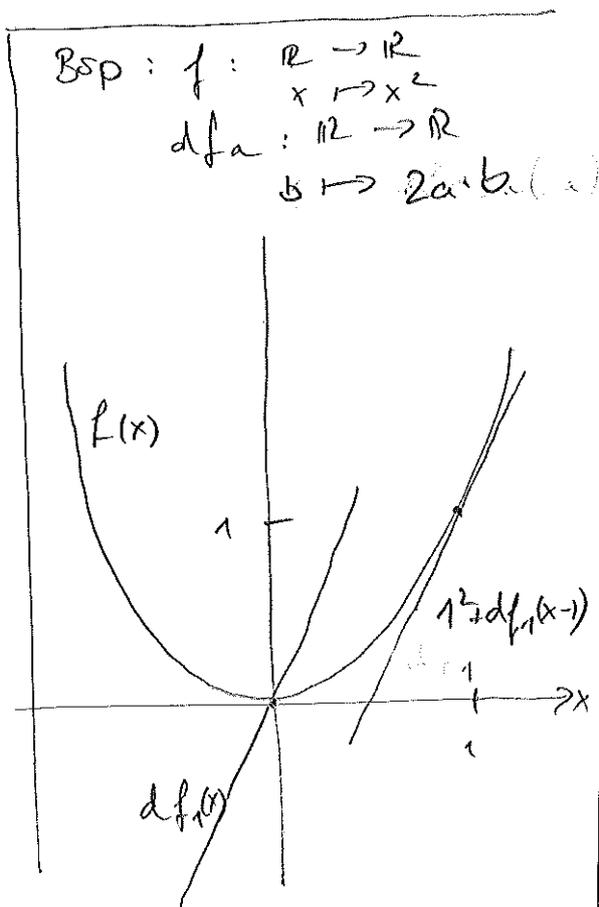
Zunächst müssen wir uns klar  
 machen, was  $f$  für eine Abb. ist:  
 $f$  wirkt Matrizen  $e \in \mathbb{R}^{n \times 2}$  auf Matr.  $e \in \mathbb{R}^{n \times 2}$   
 Ihre Jacobimatrix hat also Format  
 $n^2 \times n^2$  und soll hier weder  
 betrachtet noch bestimmt werden.

Wir denken von vornherein nicht  
 an  $Df$  sondern nur an  $df$ .

$$df_A : B \mapsto df_A(B)$$

ist eine Abb., die  
 die Linearisierung  
 von  $f$  an der Stelle  
 $A$  darstellt. Also

$$df_A : \mathbb{R}^{n \times 2} \rightarrow \mathbb{R}^{n \times 2}, \text{ wie } f \text{ selbst.}$$



Folgende  $d$  werden wir nachher verwenden:

Die Identität

$$(1) \quad g: \mathbb{R}^{n \times 2} \rightarrow \mathbb{R}^{n \times 2}, \quad A \mapsto A, \quad \text{hat}$$

$$dg_A(B) = B, \quad \text{!}$$

weil die Id eine lin. Abb. ist und die Linearisierung eine lin. Abb. ist das gleiche. (Nur dies bedenken, nicht an die Jacobi-Matrix denken.)

Genauso für festes  $C \in \mathbb{R}^{n \times 2}$ :

$$(2) \quad h: \mathbb{R}^{n \times 2} \rightarrow \mathbb{R}^{n \times 2}, \quad A \mapsto A \cdot C$$

$$dh_A(B) = B \cdot C, \quad \text{Matrixmult.}$$

weil auch  $h$  lin. ist. (D.h.

Matrixmult. mit einem festen  $C$  ist linear in den einzelnen Variablen

der Matrix. Hier steckt schon die halbe

Aufgabe drin, macht Bock z.B. klar:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} A & B \\ C & D \end{pmatrix} \text{ (wird) } \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} \mapsto \begin{pmatrix} aA + bC \\ aB + bD \\ cA + dC \\ cB + dD \end{pmatrix}.$$

Zur Illustration des Beweises mache ich ihn erst für  $n=1$ . D.h. ich beweise  $(a^j)' = j a^{j-1}$  mit der mehrdim. Kettenregel in vollst. Induktion

nach  $j$ :

Sei  $\tilde{g} = \text{id}_{\mathbb{R}}$ ,  $\tilde{h}: \mathbb{R} \rightarrow \mathbb{R}$ ,  $a \mapsto a^{j-1}$   
 $\tilde{\alpha}: \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $a \mapsto \begin{pmatrix} \tilde{g}(a) \\ \tilde{h}(a) \end{pmatrix} = \begin{pmatrix} a \\ a^{j-1} \end{pmatrix}$

$\tilde{\beta}: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $(x,y) \mapsto x \cdot y$

Dann sind

$d\tilde{\alpha}_a(b) = \begin{pmatrix} b \\ (j-1)a^{j-2} \cdot b \end{pmatrix} \leftarrow \text{Ind. Voraus.}$   
 bzw.  $d\tilde{\alpha}_a = \begin{pmatrix} da \\ (j-1)a^{j-2} da \end{pmatrix}$

$d\tilde{\beta}_{(x,y)}(u,v) = yu + xv$   
 bzw.  $d\tilde{\beta}_{(x,y)} = ydx + xdy$   
 Nach der Kettenregel gilt

$d(\tilde{\beta} \circ \tilde{\alpha})_a(b) = d\tilde{\beta}_{\tilde{\alpha}(a)} \circ d\tilde{\alpha}_a(b)$   
 $= \underset{y}{a^{j-1}} \cdot \underset{u}{b} + \underset{x}{a} \cdot \underset{v}{(j-1)a^{j-2} \cdot b}$   
 $= j \cdot a^{j-1} \cdot b$

(also  $(a^i)' = j a^{i-1}$ ).

Das gleiche mache ich jetzt mit  
Matrizen:

Seien  $g = \text{id}_{\mathbb{R}^{n^2}}$ ,  $h: \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{n^2}$ ,  $A \mapsto A^{i-1}$   
 $\alpha: \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{2n^2}$ ,  $A \mapsto \begin{pmatrix} g(A) \\ h(A) \end{pmatrix}$   
 $\beta: \mathbb{R}^{2n^2} \rightarrow \mathbb{R}^{n^2}$ ,  $(X, Y) \mapsto X \cdot Y$   
 $\begin{matrix} \downarrow & \downarrow \\ \in \mathbb{R}^{n^2} & \uparrow \\ & \text{Matrixmult.} \end{matrix}$

Dann sind

$$d\alpha_A(B) = \begin{pmatrix} B \\ \sum_{k=1}^{i-1} A^{k-1} B A^{i-k-1} \end{pmatrix}$$

[Denn: Dass  $d g_A B = B$  ist, haben wir  
in (1) gesehen und unten steht die  
led. Vorausss.]

$$(3) \quad d\beta_{(X,Y)}(u,v) = u \cdot Y + X \cdot v$$

[(3) ist mit Vorsicht zu genießen, das  
ist nicht einfach Gradient  $\cdot \begin{pmatrix} u \\ v \end{pmatrix}$ , wie  
im 1-Dimensionalen. Insbes. kommen hier  
diese Matrixprodukte nicht. Hier

kommt die Herleitung: Dazu  
 müssen wir uns klar machen, dass  
 $X$  und  $Y$  je  $n^2$  reelle Variablen  
 enthalten. Zunächst: Die rechte  
 Seite von (3) ist eine lin.  
 Fkt. in allen  $2n^2$  Variablen.  
 Um zu sehen, dass die die  
 Linearisierung von  $\beta$  ist, machen wir  
 uns folgendes klar: Wenn man  
 bei einer mehrdim. Fkt. einzelne  
 Variablen festhält (d.h.  $\beta \circ f$  betrachtet  
 mit  $f: \mathbb{R}^{n^2} \rightarrow \mathbb{R}^{2n^2}$ ,  $(x_1, \dots, x_{n^2}) \mapsto$   
 $(x_1, \dots, x_{n^2}, y_1, \dots, y_{n^2})$ , ( $y_i$  fest).) erhält  
 man die Linearisierung, indem man  
 dort die gleichen Variablen festhält  
 (d.h.  $d(\beta \circ f) = d\beta \circ \underset{f}{df} = (d\beta) \circ f$ ).  
 Betrachten wir also  $\beta$   
 nur als Fkt. von  $X$  (d.h. der restl.  
 $n^2$  Variablen). Dann ist nach Bsp (2)  
 die Linearisierung die Rechtsult.

mit dem feste  $Y$ . Für festes  $V$  ist das bis auf Konstante die rechte Seite von (3).

Genauso: Betrachte  $\beta$  als Fkt von  $Y$  (der 2. u. 2. Variablen). Dann ist die Linearisierung die Linksmit. mit  $X$ . Das stimmt wieder für festes  $u$  bis auf Konst. mit (3) überein. Damit ist unsere Linearisierung von  $\beta$  eindeutig festgelegt und muss die rechte Seite von (3) sein.]

Der Test ist einfach:

$$\begin{aligned}
 d(\beta \circ \alpha)_A(B) &= d\beta_{\alpha(A)} \circ d\alpha_A(B) \\
 &= \underbrace{\beta \cdot A^{i-1}}_u \cdot \underbrace{A^i}_Y + \underbrace{A}_{X} \underbrace{\sum_{k=1}^{i-1} A^{k-1} \beta A^{i-k-1}}_v \\
 &= \beta \cdot A^{i-1} + A (BA^{i-2} + ABA^{i-3} + \dots + A^{i-3} \beta A + A^{i-2} \beta) \\
 &= \sum_{k=1}^i A^{k-1} \beta A^{i-k}
 \end{aligned}$$

Zu tun!

b. Dies ist Lemma 2.58.

Ganz kurz, wozu das gut ist: In der Aus I macht man manchmal den Satz "Eine plts. konvergente Fkt. Folge, für die die Folge der Ableitungen gleichmäßig konvergiert, hat einen diffbaren Grenzwert", um Potenzreihe gliedweise abzuleiten.

2.58 ist das gleiche mehrdimensional:

"Eine Folge  $f_n \in C_b^1$  ist C-F bzgl.  $\|\cdot\|_{C^1}$  heißt:  $(f_n)$  ist glm. C-F.,  $(df_n)$  ist auch glm. C-F. und  $(C_b^1, \|\cdot\|_{C^1})$  ist vollst."

heißt:  $(f_n)$  hat eine diffbaren Grenzwert. Um das plts hier:

Die Ableitung der Matrix-e-Fkt.

Um zu zeigen, dass  $A \mapsto e^A$  bzgl.  $\|\cdot\|_{C^1}$  konvergiert, müssen wir zeigen,

dass  $A \mapsto e^A$  glm. C-F ist und dass

$A \mapsto \sum_{j=1}^n d(A_j)$  glm. C-F ist.

Beides folgt aus der Submult.  
der Operatornorm

$$\|A^i\| \leq \|A\|^i$$

$$\text{und } \|A^{k-1} B A^{i-k}\| \leq \|A\|^{i-1} \|B\|,$$

weil  $\|A\| < r$  und  $\|B\| < r$

wegen der Beschränktheit von  $u$ .

Damit kann man beide Fälle

mit der geom. Reihe abschätzen,

wie in 1-dim.

Mit 2.58 folgt also die Konvergenz,

und das heißt: Die Matrix-e-Fkt.

auf  $U$  hat eine (Stg., Schr.) Ableitung:

c. Ich rechne wild drauflos:

$$\int_0^1 e^{tA} B e^{(1-t)A} dt =$$

$$= \int_0^1 \left( \sum_{i=0}^{\infty} \frac{t^i A^i}{i!} \right) B \left( \sum_{k=0}^{\infty} \frac{(1-t)^k A^k}{k!} \right) dt$$

$$= \int_0^1 \sum_{i,k \geq 0} \frac{t^i (1-t)^k}{i! k!} A^i B A^k dt$$

$$= \sum_{i,k \geq 0} \underbrace{\int_0^1 \frac{t^i (1-t)^k}{i! k!} dt}_{\frac{1}{(i+k+1)!}} A^i B A^k$$

$$= \sum_{j,k \geq 0} \frac{1}{(j+k+1)!} A^j B A^k$$

$$= \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{m=1}^l A^{l-1} B A^{l-m}$$

$$= \sum_{l=0}^{\infty} \frac{1}{l!} \operatorname{def}_A(B) \quad \text{mit } \operatorname{def}(A) = A^e,$$

wobei alle Umsortierungen erlaubt sind,  
weil alle vorkommenden Reiter kommutativ sind.

(wir arbeiten in  $\mathbb{R}^{n \times 2}$ ) abs. kont.

d. Das ist der lokale Umkehratz.

Wir betrachte

$$d \exp_0 (B) = B,$$

die Identität  $\text{abo}$ , und die Id.

ist bijektiv,  $\text{abo}$   $\exp$  in einer

Umgebung von  $0$  invertierbar

mit diffbar. Umkehrfkt.

$$d \log_{\text{id}} = \left( d \exp_0 \right)^{-1} = \text{id}_{\mathbb{R}^{n \times 2}}.$$

$$\left[ \text{Beispiel: } d \ln \Big|_{\substack{1 \\ e^0}} = (e^0)^{-1} = 1 \right]$$

## 12.6 Funktionale Version der Ableitungsregel $d\frac{1}{1-x} = \frac{dx}{(1-x)^2}$ .

Es sei  $(U, \|\cdot\|)$  ein Banachraum. Wir versehen  $\mathcal{B}(U, U)$  wie immer mit der Operatornorm  $\|\cdot\|_{U \rightarrow U}$ .

- (a) Zeigen Sie für gegebenes  $j \in \mathbb{N}$ , dass die Operator-Potenzfunktion  $f : \mathcal{B}(U, U) \rightarrow \mathcal{B}(U, U)$ ,

$$f(A) = A^j = \underbrace{A \circ \dots \circ A}_{j\text{-mal}}$$

die Ableitung  $df_A(B) = \sum_{k=1}^j A^{k-1} \circ B \circ A^{j-k}$  für  $A, B \in \mathcal{B}(U, U)$  besitzt.

- (b) Es sei  $0 < k < 1$  und  $U_k := \{L \in \mathcal{B}(U, U) \mid \|L\|_{U \rightarrow U} < k\}$ . Zeigen Sie, dass die geometrische Reihe

$$f : U_k \ni A \mapsto (\text{id}_U - A)^{-1} = \sum_{j=1}^{\infty} A^j$$

in  $(C_b^1(U_k, U), \|\cdot\|_{C^1})$  konvergiert.

- (c) Folgern Sie für  $A \in U_k$  und  $B \in \mathcal{B}(U, U)$ :

$$df_A(B) = f(A) \circ B \circ f(A).$$