1.4. Die Baker-Campbell-Hausdorff-Formel

Generalvoraussetzungen:
Sei $\mathrm{G} \subseteq \mathrm{GL}_n(\mathbb{R})$ eine lineare Lie-Gruppe und $\mathfrak{g} = \mathfrak{Lie}(\mathrm{G})$ die dazugehörige Lie-Algebra. Bezeichne $\exp := \exp_{\mathrm{G}} : \mathfrak{g} \to \mathrm{G}$.
Desweiteren sei $U_0 \subseteq \mathfrak{g}$ eine Exponentialkoordinatenumgebung von $0 \in \mathfrak{g}$.
Somit ist $V_{\mathrm{I}_n} = \exp(U_0)$ eine Umgebung der Einheitsmatrix $\mathrm{I}_n \in \mathrm{G}$, sodass $\exp : U_0 \to V_{\mathrm{I}_n}$ einen Homöomorphismus mit Umkehrabbildung $\log : V_{\mathrm{I}_n} \to U_0$ darstellt.
Bezeichne $\mathrm{Exp} := \exp_{\mathrm{GL}_n(\mathbb{R})} : \mathfrak{gl}(\mathfrak{g}) \to \mathrm{GL}(\mathfrak{g})$.
Sei $U \subseteq U_0$ eine offene Umgebung der 0 in $\mathfrak{g}$, sodass

$$\exp(U) \cdot \exp(U) \subseteq V_{\mathrm{I}_n}$$

gilt.

Bemerkung 1.65 (Baues 1.67)

Seien $X, Y \in U$. Dann definiert

$$X * Y := \log(\exp(X) \cdot \exp(Y))$$

ein Produkt $* : U \times U \to U_0$.

Ziel dieses Abschnittes ist es, das Produkt $*$ zu verstehen. Dafür erst einmal folgendes

Beispiel 1.66 (Baues 1.68)

Für $\mathfrak{g}$ gelte $[\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{z}(\mathfrak{g})$, wobei $\mathfrak{z}(\mathfrak{g})$ die Menge der Elemente aus $\mathfrak{g}$ ist, die mit allen Elementen aus $\mathfrak{g}$ kommutiert. Somit ist diese Voraussetzung zu $[\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]]=\{0\}$ äquivalent, d.h. $\mathfrak{g}$ ist nilpotent der Stufe 2. Dann gilt

$$X * Y = X + Y + \tfrac{1}{2}[X, Y]$$
$$\Leftrightarrow \exp(X * Y) = \exp(X) \cdot \exp(Y) = \exp(X + Y + \frac{1}{2}[X, Y])$$

und in diesem Fall kann $U_0 = U = \mathfrak{g}$ gewählt werden.

Beweis. Um diese Aussage zu zeigen, benötigen wir folgendes Hilfslemma: Für $C, D \in \mathfrak{g}$ gilt

$$\exp(C)D\exp(-C) = D + [C, D] + \frac{1}{2!}[C, [C, D]] + \frac{1}{3!}[C, [C, [C, D]]] + ...$$

Definiere dafür

$$F(t) := \exp(tC)D\exp(-tC), \tag{1}$$

$t \in \mathbb{R}$ und $C, D \in \mathfrak{g}$. Taylorentwicklung von F an der Stelle $t = 0$ liefert

$$F(t) = F(0) + F'(0)t + \frac{1}{2}F''(0)t^2 + \frac{1}{6}F'''(0)t^3 + ...$$

Es gilt

$$\frac{\mathrm{d}F(t)}{\mathrm{d}t} = C\exp(tC)D\exp(-tC) - \exp(tC)DC\exp(-tC) \tag{2}$$

$$\begin{aligned}\frac{\mathrm{d}^2F(t)}{\mathrm{d}t^2} &= C^2\exp(tC)D\exp(-tC) - C\exp(tC)DC\exp(-tC)\\ &\quad -C\exp(tC)DC\exp(-tC) + \exp(tC)DC^2\exp(-tC)\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\mathrm{d}^3F(t)}{\mathrm{d}t^3} &= C^3\exp(tC)D\exp(-tC) - C^2\exp(tC)DC\exp(-tC)\\ &\quad -C^2\exp(tC)DC\exp(-tC) + C\exp(tC)DC^2\exp(-tC)\\ &\quad -C^2\exp(tC)DC\exp(-tC) + C\exp(tC)DC^2\exp(-tC)\\ &\quad +C\exp(tC)DC^2\exp(-tC) - \exp(tC)DC^3\exp(-tC)\end{aligned} \tag{4}$$

Setze $t = 0$ in (1),(2),(3) und (4) und erhalte

$$\begin{aligned}F(t) &= D + (CD - DC)t + \frac{1}{2}(C^2D - CDC - CDC + DC^2)t^2\\ &\quad +\frac{1}{6}(C^3D - C^2DC - C^2DC + CDC^2 - C^2DC + CDC^2\\ &\quad +CDC^2 - DC^3)t^3 + ...\end{aligned} \tag{5}$$

Setze in (5) $t = 1$ und es folgt

$$\exp(C)D\exp(-C) = F(1) = D + [C, D] + \frac{1}{2}[C, [C, D]] + \frac{1}{6}[C, [C, [C, D]]] + ...$$

Somit haben wir das Hilfslemma bewiesen.
Nun zu unserem Beweis:
Definiere

$$G(t) := \exp(Xt) \cdot \exp(Yt),$$

$t \in \mathbb{R}$ und $X, Y \in \mathfrak{g}$. Dann gilt

$$\begin{aligned}\frac{\mathrm{d}G(t)}{\mathrm{d}t} &= X \cdot \exp(Xt) \cdot \exp(Yt) + \exp(Xt) \cdot Y \cdot \exp(Yt)\\ &= X \cdot G(t) + \exp(Xt) \cdot Y \cdot \exp(-Xt) \cdot G(t)\\ &= (X + \exp(Xt) \cdot Y \cdot \exp(-Xt)) \cdot G(t)\\ &\overset{\text{Hilfslemma}}{=} (X + Y + [Xt, Y] + \frac{1}{2}[Xt, [Xt, Y]]\\ &\quad +\frac{1}{6}[Xt, [Xt, [Xt, Y]]] + ...) \cdot G(t)\\ &= (X + Y + t[X, Y] + \frac{1}{2}t^2[X, [X, Y]]\\ &\quad +\frac{1}{6}t^3[X, [X, [X, Y]]] + ...) \cdot G(t)\\ &\overset{\text{Voraussetzung}}{=} (X + Y + t[X, Y]) \cdot G(t)\end{aligned}$$

Somit haben wir eine Differentialgleichung für G erhalten. Zudem setzen wir die Anfangsbedingung $G(0) = \mathrm{I}_n$. Wir versuchen diese Differentialgleichung mit dem Ansatz

$$\begin{aligned} H(t) &:= \exp((X+Y)t + [X,Y]\frac{t^2}{2}) \\ &\underset{\text{kommutieren}}{\overset{\text{(X+Y),[X,Y]}}{=}} \exp((X+Y)t) \cdot \exp([X,Y]\frac{t^2}{2}) \end{aligned}$$

Wir müssen zeigen, dass H die Differentialgleichung erfüllt. Es gilt

$$\begin{aligned} \frac{\mathrm{d}H(t)}{\mathrm{d}t} &\overset{\text{Produktregel}}{=} (X+Y) \cdot \exp((X+Y)t) \cdot \exp([X,Y]\frac{t^2}{2}) \\ &\quad + \exp((X+Y)t) \cdot [X,Y] \cdot t \cdot \exp([X,Y]\frac{t^2}{2}) \\ &= X \cdot \exp((X+Y)t) \cdot \exp([X,Y]\frac{t^2}{2}) \\ &\quad + Y \cdot \exp((X+Y)t) \cdot \exp([X,Y]\frac{t^2}{2}) \\ &\quad + t \cdot [X,Y] \cdot \exp((X+Y)t) \cdot \exp([X,Y]\frac{t^2}{2}) \\ &= (X + Y + t \cdot [X,Y]) \cdot H(t) \end{aligned}$$

Zudem gilt $H(0) = \mathrm{I}_n$. Somit folgt

$$G(t) = H(t) \Leftrightarrow \exp(Xt) \cdot \exp(Yt) = \exp((X+Y)t + [X,Y]\frac{t^2}{2})$$

Setze $t = 1$ und wir erhalten die Behauptung.

$$\exp(X) \cdot \exp(Y) = \exp(X + Y + \frac{1}{2}[X,Y])$$

□

Für diesen Spezialfall verstehen wir das Produkt $*$. Um es allgemein zu verstehen, benötigen wir Informationen über das Differential der Exponentialabbildung. Dafür folgende

Erinnerung

Sei $f : \mathbb{R}^n \to \mathbb{R}^m$ eine Abbildung und fixiere $x \in \mathbb{R}^n$. $D_x : \mathbb{R}^n \to \mathbb{R}^m$ heißt *Differential von f an der Stelle x*, falls für

$$\varphi_x(y) := f(x+y) - f(x) - D_x(y)$$

mit $y \in \mathbb{R}^n$

$$\lim_{\|y\| \to 0} \frac{\|\varphi_x(y)\|}{\|y\|}$$

gilt.
Notation: $(df)_x$ *Differential von f an der Stelle x*

Lemma 1.67 (Baues 1.65)

Seien $X, Y \in \mathfrak{g}$.

1. Es gilt $(d\exp)_{tX}X = X \cdot \exp(tX)$.

2. Definiere $L_A(B) = A \cdot B$. Dann gilt für $(d\exp)_X : \mathfrak{g} \to T_{\exp(X)}\mathrm{G}$

$$\begin{aligned}(d\exp)_X &= L_{\exp(X)} \circ \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!}\mathrm{ad}_X^k\right)\\ &= L_{\exp(X)} \circ \left(\frac{\mathrm{id}_{\mathfrak{g}} - \mathrm{Exp}(-\mathrm{ad}_X)}{\mathrm{ad}_X}\right),\end{aligned}$$

   wobei der Bruch als formale Schreibweise zu verstehen ist.
   Anmerkung: Nach Lemma 1.54 gilt

$$T_{\mathrm{I}_n}\mathrm{G} = \mathfrak{Lie}(\mathrm{G}) = \mathfrak{g}$$

   Analog gilt

$$T_A\mathrm{G} = A \cdot T_{\mathrm{I}_n}\mathrm{G} = A \cdot \mathfrak{g}$$

   Wir wissen $\exp : \mathfrak{g} \to \mathrm{G} \subseteq \mathrm{GL}_n(\mathbb{R}) \subseteq \mathrm{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. Also gilt $(d\exp)_X : \mathfrak{g} \to \mathbb{R}^{n^2}$. Wir können das Bild von $(d\exp)_X$ noch konkretisieren durch

$$\mathrm{Bild}(d\exp_X) = \exp(X) \cdot \mathfrak{g} = T_{\exp(X)}\mathrm{G},$$

   da

$$\left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!}\mathrm{ad}_X^k\right) \in \mathfrak{gl}(\mathfrak{g}).$$

Lemma 1.68 (Baues 1.66)

Seien

$$\begin{aligned}\varphi(z) &= \frac{1-\mathrm{e}^{-z}}{z} = \sum_{k=0}^{\infty} \frac{(-z)^k}{(k+1)!},\\ \psi(z) &= \frac{z\log(z)}{z-1} = z\sum_{k=1}^{\infty}(-1)^{k-1}\frac{(z-1)^{k-1}}{k}.\end{aligned}$$

Dann gilt

$$\psi(\mathrm{e}^z)\varphi(z) = 1.$$

Beweis. Es gilt

$$\psi(\mathrm{e}^z)\varphi(z) = \frac{\mathrm{e}^z\log(\mathrm{e}^z)}{\mathrm{e}^z-1} \cdot \frac{1-\mathrm{e}^{-z}}{z} = \frac{\mathrm{e}^z z}{\mathrm{e}^z-1} \cdot \frac{1-\mathrm{e}^{-z}}{z} = \frac{\mathrm{e}^z z - z}{\mathrm{e}^z z - z} = 1.$$

□

Anmerkung: Für den folgenden Satz benutzen wir dieses Resultat:
Für $f : \mathbb{R} \to \mathbb{R}^n$ differenzierbar gilt

$$f'(t) = (df)_t(1). \tag{6}$$

Satz 1.69 (BCH-Formel, Integraldarstellung)

Seien $X, Y \in U$. Definiere für $\xi \in \mathfrak{gl}(\mathfrak{g})$

$$\Psi(\xi) = \xi \circ \left(\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (\xi - \mathrm{id}_{\mathfrak{g}})^{k-1} \right) = \xi \circ \frac{\log(\xi)}{\xi - \mathrm{id}_{\mathfrak{g}}}.$$

Dann gilt die folgende Fassung der BCH-Formel

$$X * Y = X + \int_0^1 \Psi(\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y))(Y)\ dt$$

Beweis. Definiere für $\xi \in \mathfrak{gl}(\mathfrak{g})$

$$\Phi(\xi) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} \xi^k$$

Dann gilt nach Lemma 1.67 2.

$$(d\exp)_X = L_{\exp(X)} \circ \Phi(\mathrm{ad}_X) = \exp(X) \cdot \Phi(\mathrm{ad}_X). \tag{7}$$

Die Funktion Ψ bzw. Φ entspricht in ihrer Reihendarstellung formal der Potenzreihe ψ bzw. φ aus Lemma 1.68. Somit gilt analog

$$\Psi(\mathrm{Exp}(\mathrm{ad}_X)) \circ \Phi(\mathrm{ad}_X) = \mathrm{id}_{\mathfrak{g}}. \tag{8}$$

Definiere für $t \in \mathbb{R}$

$$\begin{aligned} F(t) &:= \log(\exp(X) \cdot \exp(tY)), \\ G(t) &:= \exp(F(t)). \end{aligned}$$

Dann gilt

$$\begin{aligned} F(0) &= \log(\exp(X) \cdot \exp(0)) = \log(\exp(X)) = X, \\ F(1) &= \log(\exp(X) \cdot \exp(Y)) = X * Y, \\ F(1) &= F(0) + F(1) - F(0) = X + \int_0^1 F'(t)dt, \end{aligned}$$

mit gesuchtem $F'(t)$.
Um $X * Y$ zu verstehen, müssen wir die Ableitung von F bestimmen. Dafür

betrachten wir zwei Darstellungen der Ableitung von G.
Einerseits gilt

$$\begin{aligned}
G'(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\exp(F(t)) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\exp(X)\cdot\exp(tY) \\
&= \exp(X)\cdot\frac{\mathrm{d}}{\mathrm{d}t}\exp(tY) \\
&= \exp(X)\cdot Y\cdot\exp(tY) \\
&\underset{\text{kommutieren}}{\overset{\text{Y,exp(tY)}}{=}} \exp(X)\cdot\exp(tY)\cdot Y \\
&= \exp(\log(\exp(X)\cdot\exp(tY)))\cdot Y \\
&= \exp(F(t))\cdot Y.
\end{aligned}$$

Andererseits gilt

$$\begin{aligned}
G'(t) &\overset{(6)}{=} (dG)_t(1) \\
&= (d\exp(F))_t(1) \\
&\overset{\text{Kettenregel}}{=} ((d\exp)_{F(t)}\circ(dF)_t)(1) \\
&= (d\exp)_{F(t)}(dF)_t(1) \\
&\overset{(6)}{=} (d\exp)_{F(t)}F'(t) \\
&\overset{(7)}{=} \exp(F(t))\cdot\Phi(\mathrm{ad}_{F(t)})(F'(t))
\end{aligned}$$

Somit folgt

$$\begin{aligned}
\exp(F(t))\cdot Y &= \exp(F(t))\cdot\Phi(\mathrm{ad}_{F(t)})(F'(t)) \\
\Leftrightarrow Y &= \Phi(\mathrm{ad}_{F(t)})(F'(t))
\end{aligned}$$

Die Verknüpfung mit $\Psi(\mathrm{Exp}(\mathrm{ad}_{F(t)}))$ liefert nach (8)

$$\begin{aligned}
F'(t) &= \Psi(\mathrm{Exp}(\mathrm{ad}_{F(t)}))(Y) \\
&= \Psi(\mathrm{Ad}_{\exp(F(t))})(Y) \\
&\overset{\text{Def. von F}}{=} \Psi(\mathrm{Ad}_{\exp(X)\cdot\exp(tY)})(Y) \\
&= \Psi(\mathrm{Ad}_{\exp(X)}\circ\mathrm{Ad}_{\exp(tY)})(Y) \\
&= \Psi(\mathrm{Exp}(\mathrm{ad}_X)\circ\mathrm{Exp}(\mathrm{ad}_{tY}))(Y) \\
&= \Psi(\mathrm{Exp}(\mathrm{ad}_X)\circ\mathrm{Exp}(t\cdot\mathrm{ad}_Y))(Y)
\end{aligned}$$

Damit erhalten wir

$$\begin{aligned}
X * Y &= X + \int_0^1 F'(t)dt \\
&= X + \int_0^1 \Psi(\mathrm{Exp}(\mathrm{ad}_X)\circ\mathrm{Exp}(t\cdot\mathrm{ad}_Y))(Y)dt
\end{aligned}$$

□

Korollar 1.70 (BCH-Formel, Reihendarstellung)

Es gilt

$$X * Y = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] - [Y,[X,Y]]) + ...,$$

wobei "..." für Terme höherer Ordnung in ad_X und ad_Y steht.

Beweis. Man kann die Potenzreihe aus Lemma 1.68 schreiben als

$$\psi(z) = 1 + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)}(z-1)^k$$

(Taylorentwicklung von ψ in $z = 1$). Analog gilt für Ψ aus Satz 1.69

$$\Psi(\xi) = \mathrm{id}_{\mathfrak{g}} + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)}(\xi - \mathrm{id}_{\mathfrak{g}})^k, \tag{9}$$

$\xi \in \mathfrak{gl}(\mathfrak{g})$. Um $X * Y$ zu ermitteln, müssen wir

$$\Psi(\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y)) \overset{(9)}{=} \mathrm{id}_{\mathfrak{g}} + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)}(\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y) - \mathrm{id}_{\mathfrak{g}})^k$$

untersuchen. Hierbei gilt

$$\begin{aligned}
\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y) - \mathrm{id}_{\mathfrak{g}} &= (\mathrm{id}_{\mathfrak{g}} + \mathrm{ad}_X + \frac{1}{2}\mathrm{ad}_X^2 + ...) \circ (\mathrm{id}_{\mathfrak{g}} + t\mathrm{ad}_Y \\
&\quad + \frac{1}{2}t^2\mathrm{ad}_Y^2 + ...) - \mathrm{id}_{\mathfrak{g}} \\
&= (\mathrm{id}_{\mathfrak{g}} + t\mathrm{ad}_Y + \frac{1}{2}t^2\mathrm{ad}_Y^2 + \mathrm{ad}_X + t\mathrm{ad}_X \circ \mathrm{ad}_Y \\
&\quad + \frac{1}{2}t^2\mathrm{ad}_X \circ \mathrm{ad}_Y^2 + \frac{1}{2}\mathrm{ad}_X^2 + \frac{1}{2}t\mathrm{ad}_X^2 \circ \mathrm{ad}_Y \\
&\quad + \frac{1}{4}t^2\mathrm{ad}_X^2 \circ \mathrm{ad}_Y + ...) - \mathrm{id}_{\mathfrak{g}} \\
&= t\mathrm{ad}_Y + \frac{1}{2}t^2\mathrm{ad}_Y^2 + \mathrm{ad}_X + t\mathrm{ad}_X \circ \mathrm{ad}_Y \\
&\quad + \frac{1}{2}t^2\mathrm{ad}_X \circ \mathrm{ad}_Y^2 + \frac{1}{2}\mathrm{ad}_X^2 + \frac{1}{2}t\mathrm{ad}_X^2 \circ \mathrm{ad}_Y \\
&\quad + \frac{1}{4}t^2\mathrm{ad}_X^2 \circ \mathrm{ad}_Y + ...
\end{aligned}$$

Zudem gilt

$$\mathrm{ad}_Y(Y) = [Y,Y] = 0.$$

Somit folgt mit Satz 1.69

$$
\begin{aligned}
X * Y &= X + \int\limits_0^1 \Psi(\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y))(Y)\ dt \\
&= X + \int\limits_0^1 (\mathrm{id}_{\mathfrak{g}} + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k(k+1)} (\mathrm{Exp}(\mathrm{ad}_X) \circ \mathrm{Exp}(t \cdot \mathrm{ad}_Y) - \mathrm{id}_{\mathfrak{g}})^k)(Y)\ dt \\
&= X + \int\limits_0^1 (\mathrm{id}_{\mathfrak{g}} + \frac{1}{2}(\mathrm{ad}_X + \frac{1}{2}\mathrm{ad}_X^2 + ...) - \frac{1}{6}(t\mathrm{ad}_Y \circ \mathrm{ad}_X + \mathrm{ad}_X^2 + ...)(Y)\ dt \\
&= X + \int\limits_0^1 (\mathrm{id}_{\mathfrak{g}} + \frac{1}{2}\mathrm{ad}_X + \frac{1}{4}\mathrm{ad}_X^2 - \frac{1}{6}t\mathrm{ad}_Y \circ \mathrm{ad}_X - \frac{1}{6}\mathrm{ad}_X^2 + ...)(Y)\ dt \\
&= X + \int\limits_0^1 (\mathrm{id}_{\mathfrak{g}} + \frac{1}{2}\mathrm{ad}_X + \frac{1}{12}\mathrm{ad}_X^2 - \frac{1}{6}t\mathrm{ad}_Y \circ \mathrm{ad}_X + ...)(Y)\ dt \\
&= X + \int\limits_0^1 Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{6}t[Y,[X,Y]] + ...\ dt
\end{aligned}
$$

Gliedweises Integrieren liefert

$$
\begin{aligned}
X * Y &= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}[X,[X,Y]] - \frac{1}{12}[Y,[X,Y]] + ... \\
&= X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] - [Y,[X,Y]]) + ...
\end{aligned}
$$

$\square$