Die Exponentialfunktion – eine kurze Einführung

1) $\boldsymbol{\exp x := \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!} =: 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \ (x \geq 0)}$.

Die Folge der Werte $e_n = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$ $(n \in \mathbb{N})$ mit $x > 0$ ist streng monoton wachsend. Sie ist auch beschränkt; denn für $n_0 \geq 2x$ folgt
$e_n = e_{n_0} + \frac{x^{n_0+1}}{(n_0+1)!} + \cdots + \frac{x^n}{n!} < e_{n_0} + \frac{x^{n_0}}{n_0!}\big(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2^{n-n_0}}\big) < e_{n_0} + \frac{x^{n_0}}{n_0!}$ $(n > n_0)$.
Also existiert der Grenzwert $\exp x = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{x^k}{k!}$ $(x \geq 0)$.

2) $\boldsymbol{\Big|1 + z + \frac{z^2}{2!} + \cdots + \frac{z^m}{m!} - \big(1 + \frac{z}{n}\big)^n\Big| \leq \frac{|z|^2}{2n} \exp|z| \ (m \geq n \geq 1, z \in \mathbb{C})}$.

$$\Big|1+z+\frac{z^2}{2!}+\cdots+\frac{z^m}{m!}-\big(1+\frac{z}{n}\big)^n\Big| = \Big|\sum_{k=0}^{m}\frac{z^k}{k!} - \sum_{k=0}^{n}\frac{z^k}{k!}\big(1-\frac{1}{n}\big)\cdots\big(1-\frac{k-1}{n}\big)\Big|$$

$$\leq \sum_{k=n+1}^{m}\frac{|z|^k}{k!} + \sum_{k=2}^{n}\frac{|z|^k}{k!}\underbrace{\Big(1-\big(1-\frac{1}{n}\big)\cdots\big(1-\frac{k-1}{n}\big)\Big)}_{\leq (k-1)k/\,2n}$$

$$\leq \sum_{k=n+1}^{m}\frac{|z|^k}{k!} + \frac{|z|^2}{2n}\sum_{k=2}^{n}\frac{|z|^{k-2}}{(k-2)!} \leq \frac{|z|^2}{n(n+1)}\sum_{k=n-1}^{m-2}\frac{|z|^k}{(k+2)!}\,n(n+1) + \frac{|z|^2}{2n}\sum_{k=0}^{n-2}\frac{|z|^k}{k!}$$

$$\leq \frac{|z|^2}{2n}\sum_{k=0}^{m-2}\frac{|z|^k}{k!} \leq \frac{|z|^2}{2n}\exp|z|.$$

3) $\boldsymbol{|a_m - b_n| \leq \varepsilon_n \ (m \geq n \geq 1), \quad (\varepsilon_n)}$ Nullfolge
$\Rightarrow$ $\boldsymbol{(a_n), (b_n)}$ konvergent, $\boldsymbol{\lim a_n = \lim b_n}$.

Denn $|a_m - a_n| \leq |a_m - b_n| + |b_n - a_n| \leq 2\varepsilon_n$ $(m \geq n \geq 1)$, also (a_n) Cauchyfolge; ebenso $|b_m - b_n| \leq |b_m - a_m| + |a_m - b_n| \leq \varepsilon_m + \varepsilon_n$ $(m \geq n \geq 1)$ und damit auch (b_n) Cauchyfolge.

4) $$\boldsymbol{\exp z := \lim_{n\to\infty}\Big(1+\frac{z}{n}\Big)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{z^k}{k!} =: \sum_{k=0}^{\infty}\frac{z^k}{k!} \quad (z \in \mathbb{C})}.$$

Dies ist eine sinnvolle Definition nach 2) und 3).

5) $\boldsymbol{\Big|\big(1+\frac{z}{n}\big)^n - \big(1+\frac{w}{n}\big)^n\Big| \leq |z-w|\,\exp\big(\max(|z|,|w|)\big) \ (z, w \in \mathbb{C})}$.

Denn mit $a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1})$ folgt
$\Big|\big(1+\frac{z}{n}\big)^n - \big(1+\frac{w}{n}\big)^n\Big| \leq \frac{|z-w|}{n}\Big|\big(1+\frac{z}{n}\big)^{n-1} + \big(1+\frac{z}{n}\big)^{n-2}b + \cdots + \big(1+\frac{w}{n}\big)^{n-1}\Big|$
$\leq \frac{|z-w|}{n}\,n\big(1+\frac{\max(|z|,|w|)}{n}\big)^{n-1} \leq |z-w|\,\exp\big(\max(|z|,|w|)\big)$.
($(1+\frac{c}{n})^n = \sum_{k=0}^{n}\frac{c^k}{k!}(1-\frac{1}{n})\cdots(1-\frac{k-1}{n})$ ist für $c \geq 0$ offenbar monoton wachsend mit n.)

6) $\boldsymbol{\exp z \cdot \exp w = \exp(z+w) \quad (z, w \in \mathbb{C})}$.

Denn $\left(1+\frac{z}{n}\right)^n\left(1+\frac{w}{n}\right)^n = \left(1+\frac{z+w+zw/n}{n}\right)^n$, also nach 5)
$\left|\left(1+\frac{z}{n}\right)^n\left(1+\frac{w}{n}\right)^n - \left(1+\frac{z+w}{n}\right)^n\right| \le \frac{|zw|}{n}\exp(|z+w|+|zw|)\ (n \ge 1)$.

7) Aus 6) folgt $\exp(-z)\exp z = 1$ ($z \in \mathbb{C}$), also insbesondere $\exp z \neq 0$ ($z \in \mathbb{C}$), ferner $\exp x > 0$ ($x \in \mathbb{R}$), und $\exp x$ ist auf $\mathbb{R}$ streng monoton wachsend. Außerdem per Induktion mit $\mathbf{e := \exp(1) = 2.71828182845904523536...}$ (*Eulersche Zahl*):

$\exp n = \mathrm{e}^n$ ($n \in \mathbb{Z}$), $\quad \exp(\frac{1}{n}) = \sqrt[n]{\mathrm{e}}$, $\quad \boldsymbol{\exp r = \mathrm{e}^r\ (r \in \mathbb{Q})}$.

Daher: $\quad \boldsymbol{\mathrm{e}^z := \exp z\ (z \in \mathbb{C})}$.

8) $\displaystyle \boldsymbol{\lim_{h\to 0}\frac{\mathrm{e}^h-1}{h} = 1, \quad (\mathrm{e}^z)' = \lim_{h\to 0}\frac{\mathrm{e}^{z+h}-\mathrm{e}^z}{h} = \mathrm{e}^z\ (z \in \mathbb{C})}$.

(Dabei darf h sich *beliebig in* $\mathbb{C}\backslash\{0\}$ an 0 annähern.) Für $h \neq 0$ gilt:

$\displaystyle \left|\frac{\mathrm{e}^h-1}{h}-1\right| = \left|\sum_{k=1}^{\infty}\frac{h^{k-1}}{k!}-1\right| \le \sum_{k=2}^{\infty}\frac{|h|^{k-1}}{k!} = \frac{|h|}{2}\sum_{k=2}^{\infty}\frac{2}{(k-1)k}\frac{|h|^{k-2}}{(k-2)!} < \frac{|h|}{2}\,\mathrm{e}^{|h|}$, und

$\frac{\mathrm{e}^{z+h}-\mathrm{e}^z}{h} = \mathrm{e}^z\,\frac{\mathrm{e}^h-1}{h}$. (Formale gliedweise Ableitung ergibt:
$(a_0+a_1x+a_2x^2+\cdots)' \overset{!}{=} a_0+a_1x+a_2x^2+\cdots \quad \Rightarrow \quad a_{k+1} = \frac{1}{k+1}\,a_k\ (k \ge 0)$.
Das erklärt den Ausgangspunkt 1) der exp-Definition.)

Nachbemerkungen:

I) Die Abschätzung 5) impliziert insbesondere ganz allgemein:
$\boldsymbol{z_n \to z\ (n\to\infty) \quad \Rightarrow \quad \left(1+\frac{z_n}{n}\right)^n \to \mathrm{e}^z\ (n\to\infty)}$.

II) Aus 5) folgt $|\mathrm{e}^z-\mathrm{e}^w| \le |z-w|\,\max(\mathrm{e}^{|z|},\mathrm{e}^{|w|})$. Es gilt aber auch:
$\boldsymbol{\left|\mathrm{e}^z-\mathrm{e}^w\right| \le \min\left(|\mathrm{e}^z|,|\mathrm{e}^w|\right)\left(\mathrm{e}^{|w-z|}-1\right)\ (z,w \in \mathbb{C})}$.
Denn $\mathrm{e}^z-\mathrm{e}^w = \mathrm{e}^z(1-\mathrm{e}^{w-z})$, und $|1-\mathrm{e}^{w-z}| = |-\sum_{k=1}^{\infty}\frac{(w-z)^k}{k!}| \le \sum_{k=1}^{\infty}\frac{|w-z|^k}{k!}$.

III) Die Abschätzung 2) ist relativ scharf; wir illustrieren dies am Fall $z=1$.
$\left(1+\frac{1}{n}\right)^n < \mathrm{e} < \left(1+\frac{1}{n}\right)^{n+1}$ und $\left(1+\frac{1}{n}\right)^{n+1} - \left(1+\frac{1}{n}\right)^n = \frac{1}{n}\left(1+\frac{1}{n}\right)^n \approx \frac{\mathrm{e}}{n}$, und nach 2) $|\mathrm{e}-\left(1+\frac{1}{n}\right)^n| \le \frac{\mathrm{e}}{2n}$. Genauer:
$\mathrm{e}-\left(1+\frac{1}{n}\right)^n = \mathrm{e}-\mathrm{e}^{n\ln(1+1/n)} = \mathrm{e}\left(1-\mathrm{e}^{n(1/n-1/2n^2+1/3n^3\mp\cdots)-1}\right)$
$\displaystyle = -\mathrm{e}\sum_{k=1}^{\infty}\frac{(-\frac{1}{2n}+\frac{1}{3n^2}\mp\cdots)^k}{k!} = \frac{\mathrm{e}}{2n}+\mathcal{O}(\frac{1}{n^2})$. D.h.: $\frac{\mathrm{e}}{2n}$ ist die *bestmögliche* Abschätzung erster Ordnung.

Dies zeigt exemplarisch, wie *langsam* $(1+\frac{z}{n})^n$ gegen e^z konvergiert. Im Gegensatz dazu gilt $\left|\mathrm{e}^z-\sum_{k=0}^{n}\frac{z^k}{k!}\right| \le \frac{1}{\sqrt{2\pi(n+1)}}\left(\frac{\mathrm{e}\,|z|}{n+1}\right)^{n+1}\cdot\frac{1}{1-|z|/(n+2)}$ $(n+2>|z|)$.

Höhere Dimensionen

Die Exponentialfunktion lässt sich leicht auf den Fall *beschränkter linearer Operatoren in Banachräumen* verallgemeinern; wichtiger Spezialfall: *die Matrizen-Exponentialfunktion.* Einiges wird im folgenden genau ausgeführt, manches nur skizziert; mehr mathematische Vorkenntnisse als beim Einführungsteil werden vorausgesetzt.

Sei X ein Banachraum. Für beschränkte lineare Operatoren $A, B : X \to X$ sei das Produkt AB – wie allgemein üblich – als Hintereinanderausführung von A und B verstanden, also $(AB)x := A(Bx) \quad (x \in X)$. Dann sind Potenzen von Operatoren wohldefiniert und damit auch Polynome $a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n$ mit dem Identitätsoperator $I : x \mapsto x \quad (x \in X)$. Wir können die Abschätzung bei 2) praktisch unverändert übernehmen, mit der Operatornorm $\|\cdot\|$ anstelle des Betrages komplexer Zahlen:

$$\boldsymbol{\left\|\sum_{k=0}^{m}\frac{A^k}{k!}-\Big(I+\frac{A}{n}\Big)^n\right\| \le \frac{\|A\|^2}{2n}\exp\|A\| \ \big(m\ge n\ge 1, A\in\mathcal{B}(X)\big)}$$

Dabei bezeichne $\mathcal{B}(X)$ den Banachraum der beschränkten linearen Operatoren auf X. Auch 3) und 4) können wir ohne Änderung übernehmen und setzen:

$$\boldsymbol{\exp A := \lim_{n\to\infty}\Big(I+\frac{A}{n}\Big)^n = \lim_{n\to\infty}\sum_{k=0}^{n}\frac{A^k}{k!} =: \sum_{k=0}^{\infty}\frac{A^k}{k!} \quad (A\in\mathcal{B}(X))}$$

Im Spezialfall $X = \mathbb{C}$ ist $\mathcal{B}(X)$ nichts als die Gesamtheit der Multiplikationsoperatoren $a : z \mapsto a\,z \ \ (z \in \mathbb{C})$ und damit die Operator-Exponentialfunktion mit der „gewöhnlichen“ Exponentialfunktion identifizierbar.

Auch die Abschätzung 5) lässt sich auf den Operator-Fall übertragen; allerdings müssen wir die Begründung ein klein wenig modifizieren. Die Multiplikation der Operatoren ist nämlich *nicht kommutativ*, sogar schon im einfachsten nichttrivialen Fall $X = \mathbb{C}^2$ (oder auch $X = \mathbb{R}^2$).

$a^n - b^n = a^{n-1}(a-b) + a^{n-2}(a-b)b + a^{n-3}(a-b)b^2 + \cdots + a(a-b)b^{n-2} + (a-b)b^{n-1}$
gilt aber auch für nichtkommutative Produkte, und damit folgt

$$\|A^n - B^n\| \le \|A - B\|\left(\|A^{n-1}\| + \|A^{n-2}\|\,\|B\| + \cdots + \|B^{n-1}\|\right).$$

Dies genügt offenbar, um 5) zu erschließen:

$$\boldsymbol{\left\|\Big(1+\frac{A}{n}\Big)^n-\Big(1+\frac{B}{n}\Big)^n\right\| \le \|A-B\|\exp\big(\max(\|A\|,\|B\|)\big)\ (A,B\in\mathcal{B}(X))}$$

Die zu 6) führende Identität $\left(1+\frac{z}{n}\right)^n\left(1+\frac{w}{n}\right)^n = \left(1+\frac{z+w+zw/n}{n}\right)^n$ setzt aber *wesentlich* die Kommutativität voraus; und in der Tat kann man durch Gegenbeispiele zeigen, dass 6) im allgemeinen *nicht* gilt – ein schmerzliches Handicap beim Umgang mit Operator-Exponentialfunktionen. Wir nennen hier ein Gegenbeispiel aus dem Bereich der 2x2-Matrizen und leiten vorab eine allgemeine Formel her für die Exponentialfunktion in diesem einfachsten mehrdimensionalen Fall.

Zuerst eine Bemerkung zu einem Trivialfall: *Wenn* A und B kommutieren ($AB = BA$), *dann* gilt die zu 6) führende Identität, und die Funktionalgleichung 6) folgt mit demselben kurzen Beweis. Daher benutzen wir auch weiterhin die generelle Schreibweise $\mathbf{exp}\,\boldsymbol{A} =: \mathbf{e}^{\boldsymbol{A}}$. Also:

$$\boldsymbol{A}\,\boldsymbol{B} = \boldsymbol{B}\,\boldsymbol{A} \quad \Rightarrow \quad \mathbf{e}^{\boldsymbol{A}}\mathbf{e}^{\boldsymbol{B}} = \mathbf{e}^{\boldsymbol{A}+\boldsymbol{B}}$$

Zum 2x2-Fall: Mit $\boldsymbol{A} := \begin{pmatrix} \boldsymbol{a} & \boldsymbol{b} \\ \boldsymbol{c} & \boldsymbol{d} \end{pmatrix}$ und $\boldsymbol{I} := \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}$ gilt $A - \frac{a+d}{2}I = \begin{pmatrix} \frac{a-d}{2} & b \\ c & \frac{d-a}{2} \end{pmatrix}$, $\left(A - \frac{a+d}{2}I\right)^2 = \left(\left(\frac{a-d}{2}\right)^2 + bc\right) I$ (denn Spur 0 ergibt zwei entgegengesetzt gleiche Eigenwerte!) und mit $\boldsymbol{\omega^2} := -\left(\frac{\boldsymbol{a-d}}{\mathbf{2}}\right)^{\mathbf{2}} - \boldsymbol{bc}$ daher

$$\mathrm{e}^{\boldsymbol{A}} = \mathrm{e}^{\frac{a+d}{2}I}\,\mathrm{e}^{A-\frac{a+d}{2}I} = \mathrm{e}^{\frac{a+d}{2}} \sum_{n=0}^{\infty} \frac{1}{n!}\left(A - \tfrac{a+d}{2}I\right)^n$$

$$= \mathrm{e}^{\frac{a+d}{2}}\left(\sum_{k=0}^{\infty} \tfrac{(-1)^k\omega^{2k}}{(2k)!}I + \sum_{k=0}^{\infty} \tfrac{(-1)^k\omega^{2k}}{(2k+1)!}\left(A - \tfrac{a+d}{2}I\right)\right) = \mathbf{e}^{\frac{\boldsymbol{a+d}}{\mathbf{2}}}\left(\mathbf{cos}\,\boldsymbol{\omega}\,\boldsymbol{I} + \mathbf{sinc}\,\boldsymbol{\omega}\,\left(\boldsymbol{A} - \tfrac{\boldsymbol{a+d}}{\mathbf{2}}\boldsymbol{I}\right)\right)$$

mit $\mathbf{sinc}\,\boldsymbol{\omega} := \begin{cases} \frac{\sin\omega}{\omega}, & \omega \neq 0, \\ 1, & \omega = 0 \end{cases}$. Da cos und sinc *gerade* Funktionen sind, d.h. Funktionen von ω^2, kommt es auf das Vorzeichen von ω nicht an.

Mit $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ folgt $AB = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \neq BA = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, ferner $\mathrm{e}^A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $\mathrm{e}^B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, $\mathrm{e}^A\mathrm{e}^B \neq \mathrm{e}^{A+B} = \begin{pmatrix} \cosh 1 & \sinh 1 \\ \sinh 1 & \cosh 1 \end{pmatrix} \neq \mathrm{e}^B\mathrm{e}^A \neq \mathrm{e}^A\mathrm{e}^B$.

Man kann im $n \times n$-Fall zeigen: Haben die Matrizen A, B lauter *algebraische* Elemente, folgt aus $\mathrm{e}^A\mathrm{e}^B = \mathrm{e}^B\mathrm{e}^A$ stets auch $AB = BA$; bei nichtkommutierenden Matrizen mit algebraischen Elementen ist die Funktionalgleichung also *stets* verletzt. (Siehe EMEW, *Two Remarks on Matrix Exponentials*, Linear Algebra and its Applications 117: 127-132 (1989); siehe auch N.J. Higham, *Functions of Matrices*, SIAM 2008, Theorem 10.3.)

Allgemeiner gilt: Sind A, B nichtkommutierende beschränkte lineare Operatoren in einem Banachraum X, und gilt $\sigma(A) \cap (\sigma(A) + 2k\pi\mathrm{i}) = \emptyset$ und $\sigma(B) \cap (\sigma(B) + 2k\pi\mathrm{i}) = \emptyset$ für deren Spektren $\sigma(A), \sigma(B)$ bei beliebigem $k \in \mathbb{Z}\backslash\{0\}$, so folgt auch $\mathrm{e}^A\mathrm{e}^B \neq \mathrm{e}^B\mathrm{e}^A$. (Siehe EMEW, *A Remark on Commuting Operator Exponentials*, Proceedings of the AMS 125: 1685-1688 (1997).)

Das Gegenbeispiel $A = \begin{pmatrix} 0 & \pi \\ -\pi & 0 \end{pmatrix}$, $\mathrm{e}^A = -I$ und $B = t\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ mit $t \in \mathbb{C}\backslash\{0\}$ und damit $AB \neq BA$ zeigt, dass die Voraussetzung bezüglich der Spektren nicht wegfallen kann.

Da die Abschätzung 5) auch ohne Kommutativitäts-Voraussetzung gültig bleibt, gilt dies auch für die wichtige Folgerung aus Nachbemerkung I):

$$\boldsymbol{A_n} \to \boldsymbol{A}\ (\boldsymbol{n} \to \infty) \quad \Rightarrow \quad \left(\boldsymbol{I} + \frac{\boldsymbol{A_n}}{\boldsymbol{n}}\right)^{\boldsymbol{n}} \to \mathbf{e}^{\boldsymbol{A}}\ (\boldsymbol{n} \to \infty)$$

Dabei sind die A_n eine Folge beschränkter linearer Operatoren in einem Banachraum X, und A ist ein ebensolcher Operator; d.h. es gilt $\|A_n - A\| \to 0$.

Wir wenden dies an auf einen interessanten *drei*dimensionalen Fall, die Drehungen um eine Achse $\mathbf{a} \in \mathbb{R}^3$, $|\mathbf{a}| = 1$, mit Winkel φ.

Ist ε ein sehr kleiner Drehwinkel, wird jeder Punkt $\mathbf{x}$ in erster Ordnung *senkrecht zur Drehachse* bewegt; die Verschiebung: $\varepsilon\,\mathbf{a} \times \mathbf{x} + \mathcal{O}(\varepsilon^2)$, $|\mathcal{O}(\varepsilon^2)| \le \frac{\varepsilon^2 |\mathbf{a}\times\mathbf{x}|}{2-\varepsilon^2}$. Also

$$D_{\mathbf{a},\varepsilon}\,\mathbf{x} = \mathbf{x} + \varepsilon\,\mathbf{a} \times \mathbf{x} + \mathcal{O}(\varepsilon^2) = \big(I + \varepsilon\, A_\times + \mathcal{O}(\varepsilon^2)\big)\mathbf{x}$$

mit $\boldsymbol{A}_\times := \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$. Mit $\varepsilon = \varphi/k$, $k \to \infty$ folgt

$$D_{\mathbf{a},\varphi}\,\mathbf{x} = D^k_{\mathbf{a},\varphi/k}\mathbf{x} = \Big(I + \frac{\varphi A_\times}{k} + \mathcal{O}(k^{-2})\Big)^k \mathbf{x} \to \mathrm{e}^{\varphi A_\times}\,\mathbf{x}\ (k \to \infty).$$

Da $A_\times = (\mathbf{a} \times \mathbf{e}_1, \mathbf{a} \times \mathbf{e}_2, \mathbf{a} \times \mathbf{e}_3)$ und $\mathbf{a} \times (\mathbf{a} \times \mathbf{e}_i) = a_i\,\mathbf{a} - \mathbf{e}_i$, gilt $A_\times^2 = \mathbf{a}\,\mathbf{a}^T - I$, $A_\times^3 = -A_\times = A_\times^T$. (Man beachte die Analogie zu $\mathrm{i}^3 = -\mathrm{i}$ und $\mathrm{e}^{\mathrm{i}\varphi}$ als Drehfaktor in $\mathbb{C}$.) Also ist die spezielle Exponentialreihe $\sum_{k=0}^\infty \frac{1}{k!}\big(\varphi\, A_\times\big)^k$ leicht auszurechnen:

$$\boldsymbol{D_{\mathbf{a},\varphi} = \mathrm{e}^{\varphi A_\times} = \sum_{k=0}^{\infty} \frac{\varphi^k}{k!} A_\times^k = I + \sin\varphi\, A_\times + (1-\cos\varphi) A_\times^2}\ .$$

Da die Verallgemeinerung von 5) stets gilt, haben wir auch die allgemeine Abschätzung

$$\boldsymbol{\big\|\mathrm{e}^{A} - \mathrm{e}^{B}\big\| \le \|A - B\|\ \max\big(\mathrm{e}^{\|A\|}, \mathrm{e}^{\|B\|}\big) \quad \big(A, B \in \mathcal{B}(X)\big)}\,.$$

Aber die Herleitung der meist weit schärferen Abschätzung aus Nachbemerkung II) benutzt die Funktionalgleichung, gilt also *nicht* allgemein. Aber ein wichtiger Spezialfall trifft zu: Ist A ein beschränkter *normaler* Operator in einem Hilbertraum X und E ein beliebiger weiterer beschränkter linearer Operator in X, eine „Störung", so gilt die *Störungs-Abschätzung*

$$\boldsymbol{\big\|\mathrm{e}^{A+E} - \mathrm{e}^{A}\big\| \le \big\|\mathrm{e}^{A}\big\|\,\big(\mathrm{e}^{\|E\|} - 1\big)}\,.$$

(Siehe EMEW, *Two Remarks on Matrix Exponentials*, Linear Algebra and its Applications 117: 127-132 (1989); siehe auch R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, CambridgeUP 1991 (korr. Nachdr. 1994), Theorem 6.5.29.)

Die wohl bekannteste Anwendung der *Matrizen*-Exponentialfunktion betrifft *lineare Differenzialgleichungssysteme erster Ordnung mit konstanten Koeffizienten*:

$$\mathbf{y}'(t) = A\,\mathbf{y}(t) + \mathbf{f}(t), \quad \mathbf{y}(t_0) = \mathbf{y}_0,$$

mit $A \in \mathbb{R}^{n\times n}$, $\mathbf{y}_0 \in \mathbb{R}^n$, $\mathbf{f} : I \to \mathbb{R}^n$ stetig, $t_0 \in I$, $I \subseteq \mathbb{R}$ ein Intervall, hat die Lösung (*komponentenweises* Integral)

$$\mathbf{y}(t) = \mathbf{y}_0\,\mathrm{e}^{A(t-t_0)} + \int_{t_0}^{t} \mathrm{e}^{A(t-\tau)}\mathbf{f}(\tau)\mathrm{d}\tau \quad (t \in I)\,.$$

Eine Verallgemeinerung auf Systeme mit nichtkonstanten Koeffizienten ist der *Dysonsche Propagator* (nach Freeman Dyson). (Eine erste Einführung dazu findet man im *Teubner-Taschenbuch der Mathematik*, Herausgeber und Hauptautor Eberhard Zeidler.)

Beim Beweis der Störungs-Abschätzung kann man – als Ersatz für die nicht verfügbare Funktionalgleichung – die *Liesche Produkt-Formel* (nach Sophus Lie) benutzen; sie lautet

$$\lim_{n\to\infty}\left(\mathrm{e}^{A/n}\mathrm{e}^{B/n}\right)^n = \mathrm{e}^{A+B} \quad (A, B \in \mathcal{B}(X))\,.$$

Zum Beweis nutzen wir wieder die ohne Kommutativität gültige Identität $a^n - b^n = a^{n-1}(a-b) + a^{n-2}(a-b)b + a^{n-3}(a-b)b^2 + \cdots + a(a-b)b^{n-2} + (a-b)b^{n-1}$.

Damit $\quad \left\|\left(\mathrm{e}^{\frac{A}{n}}\mathrm{e}^{\frac{B}{n}}\right)^n - \left(\mathrm{e}^{\frac{A+B}{n}}\right)^n\right\| \le \left\|\mathrm{e}^{\frac{A}{n}}\mathrm{e}^{\frac{B}{n}} - \mathrm{e}^{\frac{A+B}{n}}\right\| \cdot$
$\cdot\left(\mathrm{e}^{(n-1)c} + \mathrm{e}^{(n-2)c}\mathrm{e}^{d} + \cdots + \mathrm{e}^{c}\mathrm{e}^{(n-2)d} + \mathrm{e}^{(n-1)d}\right) \le \left\|\mathrm{e}^{\frac{A}{n}}\mathrm{e}^{\frac{B}{n}} - \mathrm{e}^{\frac{A+B}{n}}\right\| n\mathrm{e}^{(n-1)c}$
mit $c := \frac{\|A\|+\|B\|}{n}$, $d := \frac{\|A+B\|}{n}$, also $c \ge d$.
Es gilt $\mathrm{e}^{A/n} = I + \frac{A}{n} + R_1$, $\mathrm{e}^{B/n} = I + \frac{B}{n} + R_2$, $\mathrm{e}^{(A+B)/n} = I + \frac{A+B}{n} + R_3$ mit

$\|R_1\| \le \sum_{k=2}^{\infty} \frac{\|A\|^k}{k!\,n^k} < \frac{\|A\|^2}{2n^2}\,\mathrm{e}^{\|A\|/n}$, $\|R_2\| < \frac{\|B\|^2}{2n^2}\,\mathrm{e}^{\|B\|/n}$, $\|R_3\| < \frac{\|A+B\|^2}{2n^2}\,\mathrm{e}^{\|A+B\|/n}$, also ($I$, $\frac{A}{n}$, $\frac{B}{n}$, $\frac{A+B}{n}$ fallen beim Ausmultiplizieren weg) $\left\|\mathrm{e}^{\frac{A}{n}}\mathrm{e}^{\frac{B}{n}} - \mathrm{e}^{\frac{A+B}{n}}\right\| \le \frac{c}{n^2}$ mit nur von A und B abhängigem $c > 0$. Es folgt die Liesche Produkt-Formel.

Eine Verallgemeinerung der Lieschen Formel ist die *Trottersche Produkt-Formel*, bei der es um *unbeschränkte Operatoren* sowie Gruppen und Halbgruppen von Operatoren geht; siehe etwa J. Weidmann, *Lineare Operatoren in Hilberträumen, Teil I*, Teubner 2000, oder M. Reed, B. Simon, *Methods of Modern Mathematical Physics, I*, Academic Press 1972. Hier skizzieren wir nur kurz ein paar Fakten der *Halbgruppen-Theorie*.

Diese ist eine *analytische* Ausdehnung der formalalgebraischen Lagrangeschen Taylor-Formel $T_h = \mathrm{e}^{h\,D}$ ($T_h f(x) := f(x+h)$, *Translationsoperator*) auf allgemeinere Klassen linearer Operatoren und wurde unabhängig voneinander Ende der 1940er Jahre von Einar Hille (1894-1980) und Kôsaku Yosida (1909-1990) entwickelt.

Eine *Halbgruppe* $(T_h)_{h\ge 0}$ ist eine Schar beschränkter linearer Operatoren auf einem Banachraum X mit $T_{h_1}T_{h_2} = T_{h_1+h_2}$ $(h_1, h_2 \ge 0)$, und ihren Stetigkeitseigenschaften entsprechen Eigenschaften des sogenannten *infinitesimalen Generators* A, der nur für diejenigen $x \in X$, bei denen der folgende Grenzwert *existiert*, definiert wird durch $A\,x := \lim_{h\searrow 0}(T_h\,x - x)/h$. A muss *nicht* beschränkt sein.

Es gilt $\lim_{h\searrow 0}\|T_h - I\| = 0$ genau dann, wenn $T_h = \mathrm{e}^{hA}$ mit einem *beschränkten* linearen Operator A, und dieser *ist* der (dann für *alle* $x \in X$ definierte) infinitesimale Generator. In diesem Fall hat man sogar eine Operatoren-*Gruppe*, da auch negative h Sinn machen.

Gilt hingegen nur $\lim_{h\searrow 0} T_h\,x = x$ für alle $x \in X$ (*starke* Stetigkeit, im Falle der Halbgruppe gleichwertig zur *schwachen* Stetigkeit, wie man mit dem *Satz von Pettis* zeigen kann), folgt zumindest $\|T_h\| \le M\,\mathrm{e}^{ah}$ mit Konstanten $M \ge 1$ sowie $a \ge 0$, und der infinitesimale Generator A ist ein in X dicht definierter abgeschlossener linearer Operator.

Der Fall $M = 1$, $a = 0$ („Kontraktions-Halbgruppe") ist gleichbedeutend damit, dass $\|(\lambda I - A)^{-1}\| \le 1/\lambda$ für alle $\lambda > 0$ (insbes. $(0,\infty)$ Teil der Resolventenmenge von A); dies ist das *Hille/Yosida-Theorem*. Im Allgemeinfall gilt analog $\|(\lambda I - A)^{-n}\| \le M/(\lambda - a)^n$ für alle $\lambda > a$, $n \in \mathbb{N}$. Für alle stark stetigen Halbgruppen gelten die *Darstellungsformeln*

$$T_h\,x = \mathrm{e}^{h\,A}x := \lim_{n\to\infty}\left(I - \frac{h}{n}A\right)^{-n} x = \lim_{\varepsilon\searrow 0} \mathrm{e}^{h\,(T_\varepsilon - I)/\varepsilon}x \quad (x \in X).$$

Die Konvergenz ist *gleichmäßig* auf endlichen h-Intervallen. Im Falle der *Translation* ($T_h x(t) = x(t+h)$ mit einer stetigen Funktion $x(t)$) ergibt der zweite Limes sozusagen eine *verallgemeinerte Taylor-Formel* für *stetige* Funktionen mit iterierten *Differenzen* statt Ableitungen; vgl. Pazy, p. 32f. (Siehe A. Pazy, *Semigroups of Linear Operators* *and Applications to Partial Differential Equations*, Springer 1983; K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer 2000. Die Klassiker: K. Yosida, *Functional Analysis, Sixth Ed.*, Springer 1980; E. Hille, R.S. Phillips, *Functional Analysis and Semi-groups*, AMS 1957. *Kurze* gute Einführung: W. Rudin, *Functional Analysis, 2nd Ed.*, McGraw-Hill 1991.)

Zur Motivation der Operator-Gruppen und -Halbgruppen noch ein paar Bemerkungen.
Die zeitliche Evolution eines physikalischen Systems werde beschrieben durch einen *Zustandsvektor*; $\mathbf{x}(0) = \mathbf{x}_0$ der Anfangszustand zum Zeitpunkt $t = 0$, $\mathbf{x}(t)$ der Zustand zum Zeitpunkt t. Die Evolution werde durch einen linearen Operator T_t erfasst, so dass $\mathbf{x}(t_1+t_2) = T_{t_1+t_2}\mathbf{x}_0 = T_{t_1}\,T_{t_2}\mathbf{x}_0$ für jeden Anfangszustand $\mathbf{x}_0$ und damit $T_{t_1+t_2} = T_{t_1}\,T_{t_2}$.
Beispiel: eindimensionale Diffusion. Sei $u(x,t)$ die Konzentration einer diffundierenden Substanz an der Stelle x zum Zeitpunkt $t > 0$ (Wärmestoff „Phlogiston“ im Falle der *Wärmeleitung*); dann gilt $u(x,t) = T_t u(x,0) = \frac{1}{\sqrt{4\pi t}}\int_{-\infty}^{\infty} u(\xi,0)\mathrm{e}^{(x-\xi)^2/4t}\mathrm{d}\xi$ (Faltung mit dem Hitze-Kern), und die Halbgruppeneigenschaft $T_{t_1+t_2} = T_{t_1}\,T_{t_2}$ entspricht der Tatsache, dass die Summe zweier unabhängiger normalverteilter Zufallsgrößen wieder exakt normalverteilt ist. Infinitesimaler Generator ist die zweite Ableitung, $A = \frac{\partial^2}{\partial x^2}$; denn es gilt die *Diffusionsgleichung* $\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$. (Diffusion reduziert die *Krümmung* von u.)
Zweites Beispiel: ein quantenmechanisches System. Es ist $\mathbf{x}(t)$ normiertes Element im Hilbertraum, $T_t = \mathrm{e}^{\mathrm{i}Ht}$ ist ein unitärer Operator, H der *Hamilton-Operator* (Energie-Operator) des Systems, ein unbeschränkter selbstadjungierter Operator, und $\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathrm{i}H\mathbf{x}(t)$ ist die *Schrödinger-Gleichung*; Formulierung in Physikbüchern meist $H\,\psi = \hbar\mathrm{i}\frac{\partial}{\partial t}\psi$.
Für diesen Fall ist der fundamentale *Satz von M.H. Stone (1932)* maßgebend, der Halbgruppen-Theorie vorausgehend: *Zu jeder einparametrigen stark stetigen Gruppe* $(U_t)_{t\in\mathbb{R}}$ *unitärer Operatoren im komplexen Hilbertraum* H *gibt es genau einen selbstadjungierten Operator* S *in* H *mit* $U_t = \mathrm{e}^{\mathrm{i}tS}$ $(t \in \mathbb{R})$. S ist ein im allgemeinen *unbeschränkter* selbstadjungierter Operator und deshalb $\mathrm{e}^{\mathrm{i}tS}$ definiert über den an den Spektralsatz anknüpfenden *Funktionalkalkül.*

Dies waren ein paar Stichworte zur weit über den eindimensionalen Fall hinausreichenden Bedeutung der Exponentialfunktion. Aber ihre Rolle im *ursprünglichen* – reellen und komplexen – Fall könnte vielfältiger und spektakulärer kaum sein. Wir weisen nur hin auf eines der faszinierendsten und wichtigsten Gebiete der klassischen Analysis: die Darstellung fast beliebiger Funktionen durch *Fourierreihen und Fourierintegrale.*

Ganz zum Schluss eine Anmerkung zur *numerischen Berechnung der Matrizen-Exponentialfunktion.* Heute gibt es sehr leistungsfähige Softwaresysteme, die insbesondere die Matrizen- und Vektorrechnung nahezu perfekt unterstützen: MATLAB, SCILAB und GNU OCTAVE. Die numerische Berechnung von $\exp A$ mit $A \in \mathbb{C}^{n\times n}$ beherrschen sie natürlich auch. Aber dies ist eine gar nicht so ganz leichte Aufgabe. Es gibt einen mittlerweile klassischen Zeitschriftenartikel zu diesem Thema: *Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later*, von Cleve Moler und Charles Van Loan; SIAM Review 45, pp. 3-49 (2003). Es ist die um ein Update erweiterte Fassung eines ursprünglich 1978 erschienenen Papers.