Die Exponentialfunktion auf dem Raum der linearen Abbildungen von $\mathbb{R}^N$ in sich

Teilnehmer:

8 Schülerinnen und Schüler	Heinrich-Hertz-Gymnasium Herder-Gymnasium Immanuel-Kant-Gymnasium

mit tatkräftiger Unterstützung durch:

Charlotte Walter	Humboldt-Universität zu Berlin

Gruppenleiter:

Konrad Gröger	Humboldt-Universität zu Berlin

Die gewöhnliche Exponentialfunktion spielt für die Lösung von Anfangswertaufgaben für reellwertige Funktionen eine wichtige Rolle. Die Gruppe hat gelernt, dass man auch Anfangswertaufgaben für *vektorwertige* Funktionen mit Hilfe einer Art von Exponentialfunktion lösen kann.

Um die neue Exponentialfunktion definieren zu können, waren vorbereitende Schritte nötig.

1. Für eine natürliche Zahl N ist der Raum $\mathbb{R}^N$ als Menge der N-Tupel reeller Zahlen eingeführt worden. Es ist definiert worden, wie man Elemente von $\mathbb{R}^N$ addiert und mit reellen Zahlen multipliziert. Außerdem ist die euklidische Norm auf $\mathbb{R}^N$ eingeführt und mit deren Hilfe die Konvergenz von Folgen aus $\mathbb{R}^N$ erklärt worden.

2. Es ist definiert worden, was man unter einer *linearen Abbildung* von $\mathbb{R}^N$ in sich versteht. Dieser Begriff ist durch konkrete Beispiele illustriert worden. In der Menge aller linearen Abbildungen von $\mathbb{R}^N$ in sich, die man mit $L(\mathbb{R}^N)$ bezeichnet, sind eine Addition, eine Multiplikation und eine Norm definiert worden. Wie bei $\mathbb{R}^N$ ist mit Hilfe der Norm auf $L(\mathbb{R}^N)$ eine Konvergenz in $L(\mathbb{R}^N)$ eingeführt worden.

Nach diesen Vorbereitungen konnte $\exp(A) \in L(\mathbb{R}^N)$ für $A \in L(\mathbb{R}^N)$ mit Hilfe einer konvergenten Reihe definiert werden. Diese Reihe ist der Taylorreihe für die gewöhnliche Exponentialfunktion nachempfunden. So kommt man zu einer Exponentialfunktion $\exp : L(\mathbb{R}^N) \longrightarrow L(\mathbb{R}^N)$.

Als Hauptergebnis hat sich die Gruppe folgenden Satz erarbeitet:

Eine Funktion $u : \mathbb{R} \longrightarrow \mathbb{R}^N$ genügt für gegebene $x \in \mathbb{R}^N$ und $A \in L(\mathbb{R}^N)$ genau dann den Bedingungen $u(0) = x,\ \ u'(t) = A(u(t))$ für $t \in \mathbb{R}$, wenn $u(t) = e^{tA}(x)$ ist für $t \in \mathbb{R}$.

0. Einführung und Motivation

Wir betrachten zuerst eine einfache Anfangswertaufgabe in $\mathbb{R}$, d. h. eine Differenzialgleichung, bei der eine Funktion $u : \mathbb{R} \to \mathbb{R}$ anhand einer Gleichung und einem bestimmten Anfangswert gesucht wird. Speziell fixieren wir ein $a \in \mathbb{R}$ sowie einen Startwert $x \in \mathbb{R}$ und betrachten die folgende Aufgabe:

$$u'(t) = au(t) \quad \text{für alle } t \in \mathbb{R}, \qquad\qquad u(0) = x. \tag{1}$$

Gleichungen wie diese treten an vielen Stellen in der Naturwissenschaft auf, wie beispielsweise bei Prozessen des Zerfalls oder dem Wachstum von Populationen, aber auch in Wirtschaftswissenschaften oder anderer Forschung. Diese Aufgabe lässt sich leicht lösen, wir erhalten

$$u(t) = e^{at}x \qquad \text{für alle } t \in \mathbb{R}$$

als einzige Lösung. Es treten in der Natur allerdings auch oft Anfangswertprobleme auf, an denen mehrere Variablen beteiligt sind. Diese können beispielsweise verwendet werden, um die Entwicklung von mehreren Populationen verschiedener Arten zu beschreiben. Bei diesen könnte die Größe der einen Population auch Auswirkungen auf die Größe der zweiten haben kann. Gesucht wären also beispielsweise zwei Funktionen $u_1, u_2 : \mathbb{R} \to \mathbb{R}$, die dem Gleichungssystem

$$\begin{aligned} u_1'(t) &= a_1u_1(t) + a_2u_2(t) \qquad \text{für alle } t \in \mathbb{R}, \qquad && u_1(0) = x_1,\\ u_2'(t) &= b_1u_1(t) + b_2u_2(t) \qquad \text{für alle } t \in \mathbb{R}, \qquad && u_2(0) = x_2, \end{aligned}$$

für fixierte Werte $a_1, a_2, b_1, b_2, x_1, x_2 \in \mathbb{R}$ genügen. Dieses System kann kürzer dargestellt werden, indem wir anstatt von zwei reellwertigen Funktionen u_1 und u_2 eine einzige Funktion u betrachten, die als Werte Paare von zwei reellen Zahlen hat, also Elemente des Raumes $\mathbb{R}^2$. Die Ableitung u' dieser Funktion $u : \mathbb{R} \to \mathbb{R}^2$ kann dann komponentenweise verstanden werden. Genauso können die zwei Anfangswerte x_1 und x_2 zu einem Paar $x = (x_1, x_2) \in \mathbb{R}^2$ zusammengefasst werden. Unser Gleichungssystem kann daher wie folgt geschrieben werden:

$$u'(t) = A(u(t)) \quad \text{für alle } t \in \mathbb{R}, \qquad\qquad u(0) = x,$$

wobei $A : \mathbb{R}^2 \to \mathbb{R}^2$ eine Funktion ist, die die Abhängigkeit der zwei Funktionen voneinander beschreibt und die Werte a_1, a_2, b_1, b_2 zusammenfasst:

$$A(x_1, x_2) = (a_1x_1 + a_2x_2,\ b_1x_1 + b_2x_2) \text{ für } (x_1, x_2) \in \mathbb{R}^2.$$

In unserem eindimensionalen Fall (Gleichung (1)) hätte A dagegen die Form

$$A(x) = ax \qquad \text{für alle } x \in \mathbb{R}.$$

Wir sehen, dass die Gleichungen in dieser Darstellung bis auf die Dimensionszahl genau die gleiche Form haben.

Daher können wir vermuten, dass sich derartige zwei-(und mehr-)dimensionale Anfangswertprobleme auch auf eine ähnliche Art lösen lassen, dass also wie im eindimensionalen Fall eine Abbildung e^{At} existiert, sodass

$$u(t) = e^{At}(x), \qquad \text{für alle } t \in \mathbb{R}$$

eine Lösung der Aufgabe bildet.

1. Grundbegriffe und Definitionen

Wir beginnen zuerst mit der Formalisierung der Idee, mehrere reelle Funktionen $u_1, u_2 : \mathbb{R} \to \mathbb{R}$ in eine einzige Funktion $u : \mathbb{R} \to \mathbb{R}^2$ zusammenzufassen. Wir führen die Räume $\mathbb{R}^1, \mathbb{R}^2, \dots, \mathbb{R}^N$ ein und definieren elementare Operationen.

Definition 1. Für ein fixiertes $N \in \mathbb{N},\ N \geq 1$, definieren wir $\mathbb{R}^N := \{(x_1, x_2, \dots, x_n), x_1, x_2, \dots, x_n \in \mathbb{R}\}$ als die Menge aller N-Tupel reeller Zahlen, diese nennen wir Vektoren. (In diesem Bericht sei $\mathbb{N}$ die Menge aller natürlichen Zahlen mit Null.) Wir definieren den Nullvektor $0 := (0, 0, \dots, 0)$ als einen Vektor mit nur Nullen als Komponenten. Für zwei Vektoren $x = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ und $y = (y_1, y_2, \dots, y_N) \in \mathbb{R}^N$ können wir dann deren Summe $x + y$ komponentenweise definieren:

$$x + y := (x_1 + y_1, x_2 + y_2, \dots, x_N + y_N).$$

Indem wir die i-te Komponente von $x + y$ mit $(x + y)_i$ bezeichnen, lässt sich diese Gleichung wie folgt darstellen:

$$(x + y)_i := x_i + y_i, \ \ i = 1, \dots, N.$$

Genauso definieren wir für einen Vektor $x = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ und eine reelle Zahl $t \in \mathbb{R}$ das skalare Vielfache tx des Vektors x über die folgende Gleichung:

$$(tx)_i := tx_i, \ \ i = 1, \dots, N.$$

Zum Schluss definieren wir die euklidische Norm $|x|$ eines Vektors $x \in \mathbb{R}^N$ wie gewohnt:

$$|x| := \left(\sum_{i=1}^{N} (x_i)^2 \right)^{\frac{1}{2}}.$$

Der Raum $\mathbb{R}^N$ bildet mit der Vektoraddition und der Multiplikation mit Skalaren einen Vektorraum. Weiterhin gilt für die von uns definierte Norm auch die Dreiecksungleichung, d. h. für beliebige $x, y \in \mathbb{R}^N$ gilt

$$|x + y| \leq |x| + |y|.$$

Wir benutzen die üblichen Einheitsvektoren e_k mit $(e_k)_i := \begin{cases} 1, & \text{wenn } k = i \\ 0, & \text{wenn } k \neq i \end{cases}$ (für $i, k \in \{1, \dots, N\}$), durch welche sich jeder Vektor $x = (x_1, \dots, x_N) \in \mathbb{R}^N$ als $x = \sum_{i=1}^{N} x_i e_i$. darstellen lässt.

In unserer Motivation haben wir unter anderem die *Ableitung* einer Funktion mit Werten in $\mathbb{R}^N$ betrachtet. Um diesen Begriff zu präzisieren, müssen wir zuerst definieren, was man unter Konvergenz und Differenziation im Raum $\mathbb{R}^N$ versteht. Hierfür betrachten wir eine Folge $(x_n)_{n \in \mathbb{N}} = (x_{n,1}, x_{n,2}, \dots, x_{n,N})_{n \in \mathbb{N}}$ von Vektoren in $\mathbb{R}^N$. Wir beginnen damit, für derartige Folgen eine Art der Konvergenz zu definieren. Dies verläuft analog zur Konvergenz in den reellen Zahlen, die bereits aus dem Schulunterricht bekannt ist, allerdings wird hier eine andere Art des Betrages verwendet.

Definition 2. Eine Folge $(x_n)_{n\in\mathbb{N}}$ aus $\mathbb{R}^N$ heißt *konvergent in* $\mathbb{R}^N$ *gegen den Grenzwert* $x^* \in \mathbb{R}^N$, wenn zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ so existiert, dass $|x_n - x^*| < \varepsilon$ ist für $n \geq n_\varepsilon$. Das ist gleichbedeutend mit $\lim\limits_{n\to\infty} |x_n - x^*| = 0$. Wir schreiben für den Grenzwert: $\lim\limits_{n\to\infty} x_n = x^*$. Eine Folge (x_n) aus $\mathbb{R}^N$ heißt *konvergent*, wenn sie gegen irgendeinen Punkt aus $\mathbb{R}^N$ konvergent ist.

Bemerkung 1. Wie in den reellen Zahlen sind auch Grenzwerte in $\mathbb{R}^N$ eindeutig, eine Folge (x_n) aus $\mathbb{R}^N$ konvergiert immer gegen höchstens einen Grenzwert. Der Beweis hierfür verläuft analog zu dem Beweis für reelle Zahlenfolgen. Wir sehen weiterhin leicht, dass die Konvergenz einer Folge aus $\mathbb{R}^N$ äquivalent zu der Konvergenz der Folge jeder einzelnen Koordinate ist.

Soll Konvergenz bestimmt werden, ohne dass der Grenzwert bekannt ist, eignet sich der Begriff der Cauchy-Folge. Dieser ist für die reellen Zahlen bereits bekannt und lässt sich ähnlich zu dem Konvergenzbegriff direkt auf Folgen in $\mathbb{R}^N$ übertragen.

Definition 3. Eine Folge (x_n) aus $\mathbb{R}^N$ heißt genau dann *Cauchy-Folge*, wenn $\lim\limits_{m,n\to\infty} |x_n - x_m| = 0$ ist.

Die wichtigste Eigenschaft von Cauchyfolgen in den reellen Zahlen ist die Äquivalenz zwischen Konvergenz und der Cauchyfolgen-Eigenschaft, die bereits aus dem Schulunterricht bekannt ist.

Behauptung. Eine Folge (x_n) aus $\mathbb{R}$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Auch dieser Zusammenhang lässt sich auf Folgen in $\mathbb{R}^N$ erweitern.

Behauptung. Eine Folge (x_n) aus $\mathbb{R}^N$ ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Beweis: Die Hinrichtung folgt analog zu reellen Cauchyfolgen. Für die Rückrichtung verwenden wir unsere Beobachtung, dass eine Folge aus $\mathbb{R}^N$ genau dann konvergiert, wenn jede Komponente in $\mathbb{R}$ konvergiert, und ähnlich dazu genau dann eine Cauchy-Folge ist, wenn jede Komponente eine Cauchy-Folge in $\mathbb{R}$ bildet. Jede dieser Cauchyfolgen in $\mathbb{R}$ ist dann gleichzeitig konvergent, wodurch auch die ursprüngliche Folge von Vektoren konvergent sein muss. $\square$

Um die eindimensionale Anfangswertaufgabe (1) auf mehrere Dimensionen zu verallgemeinern, haben wir eine Abbildung $A : \mathbb{R}^N \to \mathbb{R}^N$ verwendet, die den Zusammenhang der Funktion u mit ihrer Ableitung beschreibt. Die Klasse von Abbildungen A, für die derartige Anfangswertaufgaben tatsächlich auf unserem Weg lösbar sind, werden wir im Folgenden betrachten.

Definition 4. Die Funktion $A : \mathbb{R}^N \to \mathbb{R}^N$ heißt *lineare Abbildung*, falls

$$\begin{aligned} A(x+y) &= A(x) + A(y) && \text{(Additivität)} \\ \text{und} \qquad A(tx) &= tA(x) && \text{(Homogenität)} \end{aligned}$$

für alle $x, y \in \mathbb{R}^N$ und $t \in \mathbb{R}$ gilt. Statt $A(x)$ schreiben wir auch kurz Ax.

Definition 5. Wir bezeichnen mit $L(\mathbb{R}^N)$ die Menge aller linearen Abbildungen von $\mathbb{R}^N$ in sich.

Definition 6 (Elementare lineare Abbildungen)**.** Für beliebiges $N \in \mathbb{N}$ definieren wir $Ix := x$ und $0x := 0 \in \mathbb{R}^N$, jeweils für alle $x \in \mathbb{R}^N$, als die *identische Abbildung* bzw. die *Nullabbildung* von $\mathbb{R}^N$.

Beispiel 1. Für $N = 1$ besteht $L(\mathbb{R}^N)$ aus den linearen Funktionen $A(x) = mx$ $(m \in \mathbb{R})$ durch den Ursprung.

Beispiel 2. Für $N = 2$ und $x = (x_1, x_2)$ definieren wir $Dx := (-x_2, x_1)$ als eine Drehung um $\frac{\pi}{2}$ um den Ursprung und $Sx := (x_2, x_1)$ als Spiegelung an einer Diagonalen. Für $N = 3$ und $x = (x_1, x_2, x_3)$ definieren wir $Qx := (x_2, x_3, 0)$ als eine "Verschiebungs-Funktion" und $Px := (x_1, x_2, 0)$ als eine Projektion auf die Ebene mit $x_3 = 0$.

Ähnlich wie wir in $\mathbb{R}^N$ die Rechenoperationen Addition und Multiplikation mit einem Skalar koordinatenweise definiert haben, können wir derartige Operationen auch für lineare Abbildungen, diesmal über die einzelnen Funktionswerte, definieren.

Definition 7. Für zwei Abbildungen $A, B \in L(\mathbb{R}^N)$ definieren wir die Abbildung $A + B : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ durch die Vorschrift $(A + B)x := Ax + Bx \in \mathbb{R}^N$ für alle $x \in \mathbb{R}^N$.

Definition 8. Für eine Abbildung $A \in L(\mathbb{R}^N)$ und einen Skalar $t \in \mathbb{R}$ definieren wir die Abbildung $tA : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ durch die Vorschrift $(tA)x := t(Ax)$ für alle $x \in \mathbb{R}^N$.

Weiterhin kann eine Art der Multiplikation zweier Abbildungen durch Hintereinanderausführung definiert werden.

Definition 9. Für zwei lineare Abbildungen $A, B \in L(\mathbb{R}^N)$ definieren wir die Abbildung $AB : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ durch die Vorschrift $(AB)x := A(Bx)$ für alle $x \in \mathbb{R}^N$.

Für zwei lineare Abbildungen $A, B \in L(\mathbb{R}^N)$ und einen Skalar $t \in \mathbb{R}$ sind auch die Summe $A+B$ und das Produkt AB der Abbildungen, sowie das skalare Vielfache tA einer der Abbildungen selbst linear, und daher in $L(\mathbb{R}^N)$.

Auch hier gelten Rechengesetze wie die Assoziativgesetze $(A + B) + C = A + (B + C)$ und $s(tA) = (st)A$ sowie die Distributivgesetze $s(A+B) = sA+sB$, $(s+t)A = sA+sB$, $A(B+C) = AB+AC$, $(A+B)C = AC+AB$ jeweils für beliebige $A, B, C \in L(\mathbb{R}^N)$ und $s, t \in \mathbb{R}$. Die Funktion I ist das neutrale Element der Multiplikation linearer Abbildungen, es gilt also $IA = AI = A$ für alle $A \in L(\mathbb{R}^N)$.

Die Kommutativität der Multiplikation linearer Abbildungen ist allerdings nicht immer gegeben: Im Allgemeinen gilt $AB \neq BA$. Ein Beispiel hierfür ist das Produkt der in Beispiel 2 definierten Drehung $D : \mathbb{R}^2 \to \mathbb{R}^2$ um den Ursprung und einer Projektion P auf die x-Achse, definiert durch $Px = (x_1, 0)$ für alle $x = (x_1, x_2) \in \mathbb{R}^2$. Wird zuerst die Projektion und danach die Drehung angewendet, so erhalten wir das Produkt DP, beschrieben durch die Vorschrift $(DP)x = (0, x_1)$. Wird dagegen zuerst die Drehung und dann die Projektion durchgeführt, erhalten wir die Abbildung PD mit der Vorschrift $(PD)x = (-x_2, 0)$, eine vollkommen andere Funktion.

Anhand der genannten Rechengesetze erkennen wir, dass $L(\mathbb{R}^N)$ mit der soeben definierten Addition und Multiplikation linearer Abbildungen einen (für $N \geq 2$ nichtkommutativen) Ring mit Einselement bildet. Es stellt sich die Frage, für welche dieser linearen Abbildungen auch multiplikativ inverse lineare Abbildungen existieren. Da wir die Multiplikation als Hintereinanderausführung definiert hatten, müssen diese inversen Elemente Umkehrabbildungen der linearen Abbildungen seien.

Wir wissen, dass eine Funktion $A : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ genau dann eine Umkehrfunktion besitzt, wenn sie bijektiv ist, d. h., wenn die Gleichung $A(x) = y$ für jedes $y \in \mathbb{R}^N$ genau ein Lösung $x \in \mathbb{R}^N$ besitzt. Es sei also $A \in L(\mathbb{R}^N)$ bijektiv. Dann können wir eine Umkehrfunktion $A^{-1} : \mathbb{R}^N \longrightarrow \mathbb{R}^N$ definieren:

$$A^{-1}(y) := x, \quad \text{wenn } A(x) = y.$$

Diese Funktion A^{-1} ist, wie leicht zu sehen ist, auch linear und daher in $L(\mathbb{R}^N)$, wodurch also alle bijektiven Abbildungen in $L(\mathbb{R}^N)$ auch multiplikative Inverse besitzen.

Wir können auch Potenzen linearer Abbildungen auf die übliche Art definieren.

Definition 10. Für ein $A \in L(\mathbb{R}^N)$ definieren wir $A^0 := I$ und $A^{k+1} := AA^k$ für $k \in \mathbb{N}$.

Nachdem wir den Begriff der Konvergenz für Vektoren definiert haben, können wir dies auch für lineare Abbildungen tun. Hierfür definieren wir wie bei den Vektoren zuerst eine Art des Betrages, eine *Norm*, für lineare Abbildungen.

Definition 11. Für $A \in L(\mathbb{R}^N)$ setzt man

$$|A| := \sup_{\substack{x \neq 0 \\ x \in \mathbb{R}^N}} \frac{|Ax|}{|x|}.$$

Dieses Supremum ist immer endlich: Mithilfe der Basisvektoren lässt sich ein Vektor $x \in \mathbb{R}^N$ als $\sum\limits_{i=1}^{N} x_i e_i$ darstellen. Dann ist:

$$\begin{aligned} Ax &= A\left(\sum_{i=1}^{N} x_i e_i\right) = \sum_{i=1}^{N} x_i A e_i \\ \Rightarrow \quad |Ax| &\leq \sum_{i=1}^{N} |x_i A e_i| = \sum_{i=1}^{N} |x_i||Ae_i| \leq |x| \sum_{i=1}^{N} |Ae_i|. \end{aligned}$$

Die Menge der Quotienten $\dfrac{|Ax|}{|x|}$, $x \neq 0$, ist also durch die Summe $\sum\limits_{i=1}^{N} |Ae_i|$ von oben beschränkt.

Aus Definition 11 folgt auch direkt, dass $|Ax| \leq |A||x|$ ist für alle $x \in L(\mathbb{R}^N)$.

Nachdem wir eine Norm auf $L(\mathbb{R}^N)$ definiert haben, können wir nun den Begriff der Konvergenz in $L(\mathbb{R}^N)$ definieren.

Definition 12. Eine Folge (A_n) aus $L(\mathbb{R}^N)$ heißt *konvergent in* $L(\mathbb{R}^N)$ *gegen* $A^* \in L(\mathbb{R}^N)$, falls $\lim\limits_{n\to\infty} |A_n - A^*| = 0$ ist. Wir schreiben für den Grenzwert: $\lim\limits_{n\to\infty} A_n = A^*$. Eine Folge (A_n) aus $L(\mathbb{R}^N)$ heißt *konvergent*, wenn sie in $L(\mathbb{R}^N)$ gegen irgendeine Abbildung aus $L(\mathbb{R}^N)$ konvergiert.

Zu der neu definierten Norm sowie unserem neuen Begriff der Konvergenz werden wir nun einige Zusammenhänge beweisen.

Lemma 1. *Für* $A, B \in L(\mathbb{R}^N)$ *gilt* $|A + B| \leq |A| + |B|$ *und* $|AB| \leq |A||B|$.

Beweis: Es gilt

$$\sup_{x \neq 0} \frac{|(A+B)x|}{|x|} \leq \sup_{x \neq 0} \frac{|Ax| + |Bx|}{|x|} \leq |A| + |B|$$

sowie

$$\begin{aligned} & |(AB)x| = |A(Bx)| \leq |A||Bx| \leq |A||B||x| \text{ für alle } x \in \mathbb{R}^N \\ \Rightarrow \quad & |AB| \leq |A||B|. \end{aligned}$$

□

Lemma 2. *Es konvergiere* (A_n) *in* $L(\mathbb{R}^N)$ *gegen* A^* *und* (B_n) *in* $L(\mathbb{R}^*)$ *gegen* B^*. *Dann gilt*

$$\lim_{n\to\infty} (A_n + B_n) = A^* + B^*, \quad \lim_{n\to\infty} A_n B_n = A^* B^*.$$

Beweis: Es gilt (wegen der vorausgesetzten Existenz der Einzelgrenzwerte)

$$\begin{aligned} \lim_{n\to\infty} |(A_n + B_n) - (A^* + B^*)| &\leq \lim_{n\to\infty} (|A_n - A^*| + |B_n - B^*|) \\ &= \lim_{n\to\infty} |A_n - A^*| + \lim_{n\to\infty} |B_n - B^*| = 0 \end{aligned}$$

und

$$\begin{aligned}\lim_{n\to\infty}|A_nB_n - A^*B^*| &= \lim_{n\to\infty}|(A_n - A^*)B_n + A^*(B_n - B^*)|\\ &\le \left(\lim_{n\to\infty}|A_n - A^*||B_n|\right) + \left(\lim_{n\to\infty}|A^*|\right)\left(\lim_{n\to\infty}|B_n - B^*|\right)\\ &\le \lim_{n\to\infty}|A_n - A^*||B_n - B^* + B^*| + 0\\ &\le \lim_{n\to\infty}|A_n - A^*|\left(|B_n - B^*| + |B^*|\right) = 0.\end{aligned}$$

□

Auch den Begriff der Cauchyfolge, sowie dessen Bedeutung, können wir auf Folgen in $L(\mathbb{R}^N)$ verallgemeinern.

Definition 13. Eine Folge (A_n) aus $L(\mathbb{R}^N)$ heißt *Cauchyfolge* in $L(\mathbb{R}^N)$, falls $\lim\limits_{m,n\to\infty}|A_n - A_m| = 0$ ist.

Lemma 3. *Eine Folge (A_n) aus $L(\mathbb{R}^N)$ konvergiert in $L(\mathbb{R}^N)$ genau dann, wenn sie eine Cauchy-Folge ist.*

Beweis: 1. Es sei (A_n) eine Cauchyfolge in $L(\mathbb{R}^N)$. Wir fixieren ein $x \in \mathbb{R}^N$ und betrachten die Folge $(A_nx)_{n\in\mathbb{N}}$. Es gilt für $n, m \in \mathbb{N}$:

$$|A_nx - A_mx| \le |A_n - A_m||x| \to 0, \quad \text{für } m, n \to \infty.$$

Die Folge (A_nx) ist also eine Cauchyfolge in $\mathbb{R}^N$. Wir können also $A^*x := \lim\limits_{n\to\infty} A_nx$ definieren. Es folgt für beliebige $x, y \in \mathbb{R}^N$ und $t \in \mathbb{R}$:

$$\begin{aligned} A^*(x+y) &= \lim_{n\to\infty} A_n(x+y) &&= \lim_{n\to\infty} A_nx + \lim_{n\to\infty} A_ny = A^*x + A^*y\\ \text{sowie}\quad A^*(tx) &= \lim_{n\to\infty} A_n(tx) &&= \lim_{n\to\infty} tA_nx = tA^*x.\end{aligned}$$

Die Abbildung A^* ist also linear. Nun gilt

$$\begin{aligned}|(A_n - A^*)x| &= |(A_n - A_m + A_m - A^*)x|\\ &\le |A_n - A_m||x| + |A_mx - A^*x|\end{aligned}$$

Sei $\varepsilon > 0$ gegeben und n_ε so gewählt, dass $|A_n - A_m| < \varepsilon$ gilt, falls $n \ge n_\varepsilon, m \ge n_\varepsilon$ ist. Der Grenzübergang $m \to \infty$ liefert

$$|(A_n - A^*)x| \le \varepsilon|x| + 0$$

und daher für $x \neq 0$:

$$\frac{|(A_n - A^*)x|}{|x|} \le \varepsilon.$$

Da dies für alle $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$ gilt, folgt dass $\lim\limits_{n\to\infty}|A_n - A^*| = 0$ ist. Da die Abbildung A^* wie eben gezeigt auch in $L(\mathbb{R}^N)$ liegt, ist sie Grenzwert der Cauchyfolge (A_n).
2. Die Hinrichtung verläuft analog zu konvergenten Folgen in $\mathbb{R}$. □

Wir wollen nun unendliche Reihen linearer Abbildungen untersuchen. Um einen Ausdruck wie $\sum\limits_{n=0}^{\infty} A_n$ für lineare Abbildungen $A_n \in L(\mathbb{R}^N)$, $n \in \mathbb{N}$, zu verstehen, betrachten wir die Partialsummen

$$S_m := \sum_{n=0}^{m} A_n$$

und definieren die unendliche Reihe als Grenzwert der Folge dieser Partialsummen.

Definition 14. Man sagt, eine Reihe der Form $\sum\limits_{n=0}^{\infty} A_n$ konvergiere in $L(\mathbb{R}^N)$, wenn die Folge (S_m) der Partialsummen $S_m := \sum\limits_{n=0}^{m} A_n$ in $L(\mathbb{R}^N)$ konvergiert. Ist $S^* = \lim\limits_{n\to\infty} S_m$, so nennt man S^* den *Wert* der Reihe $\sum\limits_{n=0}^{\infty} A_n$. Eine Gleichung $\sum\limits_{n=0}^{\infty} A_n = S^*$ besagt dann, dass $\lim\limits_{m\to\infty} S_m$ existiert und gleich S^* ist.

Als ein hilfreiches Kriterium für die Konvergenz von unendlichen Reihen führen wir den Begriff der absoluten Konvergenz ein, mithilfe dessen das Problem auf eine unendliche Reihe in den reellen Zahlen reduziert werden kann.

Definition 15. Man sagt, eine Reihe $\sum\limits_{n=0}^{\infty} A_n, A_n \in L(\mathbb{R}^n)$, konvergiere *absolut*, wenn die Reihe $\sum\limits_{n=0}^{\infty} |A_n|$ in $\mathbb{R}$ konvergiert.

Lemma 4. *Wenn eine Reihe* $\sum\limits_{n=0}^{\infty} A_n$ *absolut konvergiert, so konvergiert sie auch, d. h.* $\lim_{m\to\infty} S_m$ *existiert in* $L(\mathbb{R}^N)$.

Beweis: Wir betrachten zwei Partialsummen S_m, S_n (für $n, m \in \mathbb{N}$), o.B.d.A. mit $m \leq n$. Es folgt:

$$|S_n - S_m| = \left| \sum_{k=m+1}^{n} A_k \right| \leq \sum_{k=m+1}^{n} |A_k|$$

Wir definieren $s_m := \sum\limits_{n=0}^{m} |A_k|$. Nach Voraussetzung konvergiert die Folge (s_m) und ist daher eine Cauchy-Folge. Also folgt:

$$|S_n - S_m| \leq s_n - s_m = |s_n - s_m| \to 0 \quad \text{für } m, n \to \infty$$

Die Folge (S_m) ist also eine Cauchyfolge in $L(\mathbb{R}^N)$ und daher nach Lemma 3 konvergent. □

2. Die Exponentialfunktion und einige ihrer Eigenschaften

Nach den Vorbereitungen des vorigen Abschnitts können wir nun unsere Exponentialfunktion auf $L(\mathbb{R}^N)$ definieren:

Lemma 5. *Für jedes* $A \in L(\mathbb{R}^N)$ *konvergiert die Reihe* $\sum\limits_{n=0}^{\infty} \frac{A^n}{n!}$ *absolut.*

Beweis: Wir betrachten die Reihe

$$\sum_{n=0}^{\infty} \left| \frac{A^n}{n!} \right| = \sum_{n=0}^{\infty} \frac{|A^n|}{n!} \leq \sum_{n=0}^{\infty} \frac{|A|^n}{n!}.$$

Die letzte Umformung ergibt sich mithilfe von vollständiger Induktion aus Lemma 1. Es seien beliebige $n \geq m \geq 2|A|$ gegeben. Dann folgt

$$\frac{|A|^n}{n!} \leq \frac{|A|^m}{m!} \cdot \frac{|A|^{n-m}}{m^{n-m}} \leq \frac{(2|A|)^m}{m!} \left(\frac{1}{2}\right)^n.$$

Der linke Faktor ist von n unabhängig und kann aus der Summe ausgeklammert werden, es verbleibt eine geometrische Reihe mit der Basis $\frac{1}{2}$. Da diese konvergiert, ist auch die Reihe $\sum\limits_{n=0}^{\infty} \frac{A^n}{n!}$ selbst absolut konvergent. □

Definition 16. Man definiert die *Exponentialfunktion* $\exp : L(\mathbb{R}^N) \to L(\mathbb{R}^N)$ durch

$$\exp(A) := \sum_{n=0}^{\infty} \frac{A^n}{n!} \qquad \text{für } A \in L(\mathbb{R}^N).$$

Statt $\exp(A)$ schreiben wir auch e^A.

Beispiel 3. Für beliebige $N \in \mathbb{N},\ N > 0$, und für $A = tI$ mit $t \in \mathbb{R}$ gilt:

$$\exp(A) = \sum_{n=0}^{\infty} \frac{(tI)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} I = \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} \right) I = e^t I.$$

Wir werden nun einige Eigenschaften der Exponentialfunktion auf $L(\mathbb{R}^N)$ beweisen und auf diese Art einige Gemeinsamkeiten mit der Exponentialfunktion auf $\mathbb{R}$ aufzeigen.

Lemma 6. *Es seien $A, B \in L(\mathbb{R}^N)$ vertauschbar, d. h., es gelte $AB = BA$. Dann ist*

$$(A + B)^n = \sum_{k=0}^{n} \binom{n}{k} A^k B^{n-k}.$$

Beweis: Der Beweis verläuft analog zum binomischen Lehrsatz für reelle Zahlen. □

Satz 1. *Es seien $A, B \in L(\mathbb{R}^N)$ vertauschbar. Dann gilt*

$$\exp(A + B) = \exp(A) \exp(B).$$

Beweis: Nach Lemma 6 ist

$$\sum_{k=0}^{2n} \frac{(A + B)^k}{k!} = \sum_{k=0}^{2n} \frac{1}{k!} \sum_{i=0}^{k} \frac{k!}{i!(k - i!)} A^i B^{k-i}.$$

Es sei $K_n := \{(i, j), i \in \mathbb{N}, j \in \mathbb{N}, i + j \leq 2n\}$ sowie $L_n := \{(i, j), i \in \mathbb{N}, j \in \mathbb{N}, i \leq n, j \leq n\}$. Es folgt direkt, dass $L_n \subseteq K_n$ ist. Zum Schluss sei $M_n := K_n \setminus L_n$. Dann folgt:

$$\sum_{k=0}^{2n} \frac{(A + B)^k}{k!} = \sum_{i,j \in K_n} \frac{A^i}{i!} \frac{B^j}{j!} = \left(\sum_{i,j \in L_n} \frac{A^i}{i!} \frac{B^j}{j!} \right) + \left(\sum_{i,j \in M_n} \frac{A^i}{i!} \frac{B^j}{j!} \right) = \left(\sum_{i=0}^{n} \frac{A^i}{i!} \sum_{j=0}^{n} \frac{B^j}{j!} \right) + \left(\sum_{i,j \in M_n} \frac{A^i}{i!} \frac{B^j}{j!} \right).$$

Wir definieren

$$S_n := \sum_{k=0}^{2n} \frac{(A + B)^k}{k!} - \sum_{i=0}^{n} \frac{A^i}{i!} \sum_{j=0}^{n} \frac{B^j}{j!}.$$

Dann gilt

$$|S_n| = \Big| \sum_{(i,j) \in M_n} \frac{A^i}{i!} \frac{B^j}{j!} \Big| \leq (n + 1)n \frac{\max\{|A|, |B|, 1\}^{2n}}{(n + 1)!} = \frac{c^2 \left(c^2\right)^{n-1}}{(n - 1)!} \longrightarrow 0, \quad \text{für } n \to \infty.$$

mit $c := \max\{|A|, |B|, 1\}$. Folglich gilt $\lim\limits_{n\to\infty} S_n = 0$. Zugleich gilt nach Lemma 2

$$\lim_{n\to\infty} S_n = \exp(A + B) - \exp(A) \exp(B).$$

Daher ist $\exp(A + B) = \exp(A) \exp(B)$. □

Korollar 1. Für jedes $A \in L(\mathbb{R}^N)$ ist e^A invertierbar: $I = e^0 = e^{A-A} = e^A e^{-A}$

Die letzte Gleichung folgt, da A und $-A$ vertauschbar sind. Somit ist e^{-A} ein Inverses von e^A.

Zum Schluss können wir nun bestätigen, dass unsere neu definierte Exponentialfunktion tatsächlich zu der Lösung der von uns beschriebenen Anfangswertaufgaben beiträgt.

Lemma 7. *Für jedes $A \in L(\mathbb{R}^N)$ ist*

$$\lim_{t\to 0} \frac{\exp(tA) - I}{t} = A.$$

Beweis: Es gilt

$$\frac{1}{t}\left(\sum_{n=0}^{\infty} \frac{(tA)^n}{n!} - I\right) = \sum_{n=1}^{\infty} \frac{t^{n-1}A^n}{n!} = A + t\sum_{n=2}^{\infty} \frac{t^{n-2}A^n}{n!}.$$

Für $|t| \leq 1$ ist

$$\sum_{n=2}^{\infty} |t|^{n-2} \frac{|A|^n}{n!} \leq \sum_{n=2}^{\infty} \frac{|A|^n}{n!}.$$

Diese Reihe konvergiert, wodurch für $t \to 0$ der rechte Summand im ursprünglichen Ausdruck gegen 0 konvergiert. Es folgt:

$$\lim_{t\to 0} \frac{1}{t}\left(\sum_{n=0}^{\infty} \frac{(tA)^n}{n!} - I\right) = A.$$

□

Definition 17. Eine Funktion $u : \mathbb{R} \to \mathbb{R}^N$ heißt differenzierbar in $t \in \mathbb{R}$, falls

$$\lim_{h\to 0} \frac{u(t+h) - u(t)}{h}$$

existiert. Für den Grenzwert schreibt man $u'(t)$. Ist u in allen $t \in \mathbb{R}$ differenzierbar, so heißt u differenzierbar.

Wir können nun unsere zentrale Aussage formulieren.

Satz 2. *Die durch die Vorschrift*

$$u(t) := \exp(tA)x \text{ für } t \in \mathbb{R}, \qquad x \in \mathbb{R}^N,\ A \in L(\mathbb{R}^N) \text{ gegeben,}$$

definierte Funktion $u : \mathbb{R} \to \mathbb{R}^N$ ist die einzige Lösung der (Anfangswert-) Aufgabe

$$u'(t) = Au(t) \text{ für } t \in \mathbb{R}, \quad u(0) = x.$$

Beweis: 1. Es sei $u(t) := \exp(tA)x$, $t \in \mathbb{R}$. Wir berechnen u':

$$\begin{aligned}\frac{u(t+h) - u(t)}{h} &= \frac{1}{h}\left(\exp((h+t)A) - \exp(tA)\right)x = \frac{1}{h}\left(\exp(hA)\exp(tA) - \exp(tA)\right)x \\ &= \frac{1}{h}\left(\exp(hA) - I\right)\exp(tA)x \longrightarrow A\exp(tA)x \qquad \text{für } h \to 0 \text{ nach Lemma 7.}\end{aligned}$$

Es gilt weiterhin $u(0) = \exp(0A)x = \exp(0)x = Ix = x$. Folglich ist u eine Lösung der Aufgabe.

2. Sei nun v eine Lösung von $v'(t) = Av(t)$ für $t \in \mathbb{R}$ mit $v(0) = x$. Wir setzten $w(t) := \exp(-tA)v(t)$ und bilden die Ableitung von w. Unter Benutzung der Differenzialgleichung $v'(t) = Av(t)$ ergibt sich, dass $w' = 0$ sein muss. Folglich ist w konstant. Wegen $w(0) = \exp(0)v(0) = x$ ist $w(t) = x$ und daher $v(t) = \exp(tA)w(t) = \exp(tA)x = u(t)$ für alle $t \in \mathbb{R}$. Damit ist gezeigt, dass u die einzige Lösung der betrachteten Anfangswertaufgabe ist. □

Wir gehen nun zurück zu unseren ursprünglichen Beispielen für lineare Abbildungen und berechnen für diese die Exponentialfunktion.

Beispiel 4 (Projektion)**.** Wir betrachten zuerst die durch $Px := (x_1, x_2, 0)$ für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ gegebene Funktion $P \in L(\mathbb{R}^3)$. Wir sehen sofort, dass $P^2 = P$ und daher dass $P^n = P$ für alle $n \in \mathbb{N}$, $n \geq 1$, gilt. Es folgt für $t \in \mathbb{R}$ und $x = (x_1, x_2, x_3) \in \mathbb{R}^3$:

$$e^{tP}x = \sum_{n=0}^{\infty} \frac{(tP)^n x}{n!} = \frac{0^0 P^0 x}{0!} + \sum_{n=1}^{\infty} \frac{t^n P^n x}{n!} = Ix + \left(\sum_{n=1}^{\infty} \frac{t^n}{n!} \right) Px = x + (e^t - 1)Px = \big(e^t x_1, e^t x_2, x_3\big)\,.$$

Wir können auch bestätigen, dass diese Funktion tatsächlich unsere Anfangswertaufgabe (für diese Funktion P) erfüllt: Mit $u(t) := e^{tP}x$ für $t \in \mathbb{R}$ folgt

$$u(0) = (x_1, x_2, x_3) = x \quad \text{und} \quad u'(t) = \big(e^t x_1, e^t x_2, 0\big) = Pu(t) \text{ für } t \in \mathbb{R}.$$

Die Funktion u löst also entsprechend Satz 2 die Anfangswertaufgabe.

Beispiel 5. Die Potenzen der durch die Vorschrift $Dx := (-x_2, x_1)$ für $x = (x_1, x_2) \in \mathbb{R}^2$ definierten linearen Abbildung D sind ähnlich regelmäßig: Es gilt $D^4 = I$ und daher $D^{4n} = I$, $D^{4n+1} = D$, $D^{4n+2} = -I$ und $D^{4n+3} = -D$ für alle $n \in \mathbb{N}$. Für $x = (x_1, x_2) \in \mathbb{R}^2$ und $t \in \mathbb{R}$ folgt:

$$\begin{aligned} e^{tD}x &= \left(\sum_{n=0}^{\infty} \frac{(tD)^{2n}x}{(2n)!} \right) + \left(\sum_{n=0}^{\infty} \frac{(tD)^{2n+1}x}{(2n+1)!} \right) = \left(\sum_{n=0}^{\infty} (-1)^n \frac{t^{2n}}{(2n)!} \right) x + \left(\sum_{n=0}^{\infty} (-1)^n \frac{t^{2n+1}}{(2n+1)!} \right) Dx \\ &= \cos(t)x + \sin(t)Dx = (\cos(t)x_1 - \sin(t)x_2, \sin(t)x_1 + \cos(t)x_2). \end{aligned}$$

(Die Summanden können umgeordnet werden, da wie zuvor gezeigt die Reihe absolut konvergent ist.) Mit $u(t) := e^{tD}x$ für $t \in \mathbb{R}$ folgt daher

$$u(0) = (x_1, x_2) = x \quad \text{und} \quad u'(t) = (-\sin(t)x_1 - \cos(t)x_2,\ \cos(t)x_1 - \sin(t)x_2) = Du(t) \text{ für } t \in \mathbb{R}.$$

Beispiel 6. Für die durch $Sx := (x_2, x_1)$ für $x = (x_1, x_2) \in \mathbb{R}^2$ definierte Abbildung $S \in L(\mathbb{R}^2)$ verläuft die Berechnung ähnlich: Hier gilt $S^2 = I$ und daher $S^{2n} = I$ und $S^{2n+1} = S$ für alle $n \in \mathbb{N}$. Für $x = (x_1, x_2) \in \mathbb{R}^2$ und $t \in \mathbb{R}$ folgt also:

$$\begin{aligned} e^{tS}x &= \left(\sum_{n=0}^{\infty} \frac{(tS)^{2n}x}{(2n)!} \right) + \left(\sum_{n=0}^{\infty} \frac{(tS)^{2n+1}x}{(2n+1)!} \right) = \left(\sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \right) x + \left(\sum_{n=0}^{\infty} \frac{t^{2n+1}}{(2n+1)!} \right) Sx \\ &= \cosh(t)x + \sinh(t)Sx = (\cosh(t)x_1 + \sinh(t)x_2,\ \sinh(t)x_1 + \cosh(t)x_2). \end{aligned}$$

Beispiel 7. Es sei $Qx := (x_2, x_3, 0)$ für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. Dann ist $Q^2x = (x_3, 0, 0)$, $Q^3 = 0$ und daher auch $Q^n = 0$ für alle $n \in \mathbb{N}$, $n \geq 3$. Für $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ und $t \in \mathbb{R}$ gilt also:

$$e^{tQ}x = \sum_{n=0}^{\infty} \frac{(tQ)^n}{n!} = Ix + tQx + \frac{t^2}{2}Q^2x = \left(x_1 + tx_2 + \frac{t^2}{2}x_3,\ x_2 + tx_3,\ x_3 \right).$$

In diesen Beispielen erkennen wir in den Funktionswerten der Exponentialfunktion auf $L(\mathbb{R}^N)$ nicht nur exponentielles Verhalten, sondern auch Funktionen wie Winkelfunktionen oder Polynome. Dies zeigt bereits, dass die Exponentialfunktion auf $L(\mathbb{R}^N)$ weitaus vielfältiger ist als die bekannte Exponentialfunktion auf den reellen Zahlen. Unser Ergebnis in Satz 2 zeigt außerdem, wie weitreichend Anwendungen dieser Funktion sein können.