Reelle Analysis

Teil I: Gewöhnliche Differentialgleichungen

Skript von Prof. Dr. Helge Glöckner zu einer Vorlesung an der Universität Paderborn im WS 2019/20

Inhaltsverzeichnis

1 Grundbegriffe und erste Beispiele

Beispiel 1.1 (Ein Populationsmodell/Exponentielles Wachstum). Wir betrachten eine Population von Bakterien in einer Petrischale. Zu einer festen Zeit t_0 (zu der wir die Beobachtung beginnen) gebe es $N(t_0) = N_0 > 0$ Bakterien. Wie groß ist die Zahl $N(t)$ der Bakterien zu einer beliebigen Zeit t? Wir modellieren das Problem wie folgt: Da $N(t)$ sehr groß ist und gezeichnet von einer Kurve nicht zu unterscheiden wäre, machen wir die vereinfachende Annahme, dass $N(t)$ eine stetig differenzierbare Funktion von t ist. Wir machen weiter die Annahme, dass die Änderungsrate/Ableitung $N'(t)$ zur Zeit t proportional zu $N(t)$ ist (liegen z.B. doppelt so viele Bakterien vor, erwarten wir die doppelte Zuwachsrate).[1] Mit einer reellen Konstanten $\alpha > 0$ ist also

$$N'(t) = \alpha N(t). \tag{1}$$

Um die Gleichung zu lösen, teilen wir duch $N(t) > 0$ und erhalten

$$\alpha = \frac{N'(t)}{N(t)} = \frac{d}{dt}(\ln(N(t)).$$

Integrieren beider Seiten von t_0 bis t liefert

$$\alpha(t - t_0) = \ln(N(t)) - \ln(N(t_0)) = \ln \frac{N(t)}{N_0}.$$

Wir wenden die Exponentialfunktion auf beide Seten an und lösen nach $N(t)$ auf, mit dem Ergebnis

[1] Dies ist vernünftig, so lange die Polulation nicht zu groß ist, so dass ausreichend Nährstoffe und Platz zur Verfügung stehen.

$$N(t) \;=\; N_0\, e^{\alpha(t-t_0)}. \tag{2}$$

Dann ist $N(t_0) = N_0 e^0 = N_0$ und in der Tat (unter Benutzung der Kettenregel)

$$N'(t) = \alpha N_0 e^{\alpha(t-t_0)} = \alpha N(t).$$

Die durch (2) gegebene Funktion $N\colon \mathbb{R} \to \mathbb{R}$ löst also die Gleichung (1) unter der Anfangsbedingung $N(t_0) = N_0$ und ist die einzige solche (stets positive) Funktion (da wir in jedem Schritt des Rechenwegs nur Schlussfolgerungen gezogen haben). Nach (2) zeigt die Bakterienpopulation exponentielles Wachstum. Da $N(t) \to \infty$ für $t \to \infty$, ist die Lösung für große t übrigens nicht realistisch und unsere Modellierung war zu schlecht (in eine Petrischale passen nicht beliebig viele Bakterien).

Im vorigen Beispiel haben wir eine Differentialgleichung gelöst (genauer: ein Anfangswertproblem). Die Begriffe sind wie folgt:

Definition 1.2 Eine *Differentialgleichung* (erster Ordnung) in $\mathbb{R}^n$ ist eine Gleichung der Form

$$y'(t) = f(t, y(t)) \tag{3}$$

(oder kurz: $y' = f(t,y)$) mit einer gegebenen Funktion $f\colon U \to \mathbb{R}^n$ auf einer Teilmenge $U \subseteq \mathbb{R} \times \mathbb{R}^n$. Eine *Lösung der Differentialgleichung* (3) ist eine C^1-Funktion $\gamma\colon I \to \mathbb{R}^n$ auf einem nicht entarteten Intervall $I \subseteq \mathbb{R}$ derart, dass

$$(t, \phi(t)) \in U \quad \text{für alle } t \in I$$

(also $\operatorname{graph}(\phi) := \{(t,\phi(t))\colon t \in I\} \subseteq U$) und

$$\phi'(t) = f(t, \phi(t)) \quad \text{für alle } t \in I.$$

Definition 1.3 Gegeben f wie zuvor und $(t_0, y_0) \in U \subseteq \mathbb{R} \times \mathbb{R}^n$ nennt man eine Lösung $\phi\colon I \to \mathbb{R}^n$ der Differentialgleichung (3) eine *Lösung des Anfangswertproblems*

$$\begin{cases} y' &= f(t,y) \\ y(t_0) &= y_0, \end{cases}$$

wenn zudem $t_0 \in I$ und $\phi(t_0) = y_0$ gilt.

Bemerkung 1.4 (a) Wir betrachten ausschließlich *explizite* Differentialgleichungen der Form (3), nicht implizite Differentialgleichungen der Form

$$0 = f(t, y(t), y'(t)).$$

(b) In der Literatur spricht man oft lediglich im Falle $n = 1$ von Differentialgleichungen, im Falle $n \geq 2$ von *Systemen* von Differentialgleichungen.

(c) Später betrachten wir auch Differentialgleichungen höherer Ordnung, der Form

$$y^{(m)}(t) = f(t, y(t), y'(t), y''(t), \ldots, y^{(m-1)}(t)).$$

(d) Die von uns betrachteten Differentialgleichungen sind *gewöhnliche* Differentialgleichungen, wir suchen also Lösungen $t \mapsto \phi(t)$, die Funktionen einer reellen Variablen sind. Sucht man hingegen Funktionen ϕ mehrerer Variablen und stellt Gleichungen unter Benutzung partieller Ableitungen auf, so spricht man von einer *partiellen Differentialgleichung*; solche Gleichungen sind nicht Gegenstand dieser Vorlesung. Ein Beispiel ist die Laplace-Gleichung

$$0 = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2},$$

deren Lösungen sogenannte *harmonische Funktionen* $(x, y) \mapsto \phi(x, y)$ sind.

(e) Wir kürzen "Differentialgleichung" mitunter als "DGL" ab und "Anfangswertproblem" als "AWP." Im Englischen kürzt man "ordinary differential equation" als "ODE" ab, "partial differential equation" als "PDE".

Fortsetzung von Beispiel 1.1.
In Beispiel 1.1 haben wir also das Anfangswertproblem

$$\begin{cases} y'(t) &= f(t, y(t)), \\ y(t_0) &= N_0 \end{cases}$$

gelöst mit der durch $f(t, y) = \alpha y$ gegebenen Funktion. Was wir als Definitionsbereich U von f nehmen, haben wir nun zu präzisieren. Natürliche Wahlen sind $U := \mathbb{R} \times \mathbb{R}$ oder (weil die Lösung überall positiv sein soll) $U := \mathbb{R}\times]0, \infty[$. Die Formulierung in Beispiel 1.1 war noch unpräzise und unvollständig, da der Definitionsbereich nicht angegeben wurde.

Beispiel 1.6 (Radioaktiver Zerfall). Wir betrachten eine Probe einer radioaktiv zerfallenden Substanz. Die Anzahl der Atome der Substanz zur Zeit t sei $N(t) > 0$ (wobei wir wieder näherungsweise N als stetig dfferenzierbare Funktion annehmen) und zu einer Anfangszeit t_0 liegen N_0 Atome vor. Modellierung: Wir nehmen an, dass $N'(t)$ proportional zu $N(t)$ ist, mit einer negativen reellen Proportonalitätskonstanten $-k$ (mit $k > 0$). Zu lösen haben wir also das Anfangswertproblem

$$\begin{cases} N'(t) &= -kN(t), \\ N(t_0) &= N_0. \end{cases}$$

Der gleiche Rechenweg wie in Beispiel 1.1 zeigt, dass es auf $\mathbb{R}$ genau eine überall positive Lösung gibt,

$$N(t) = N_0 \, e^{-k(t-t_0)}.$$

Die Zahl der Atome nimmt also exponentiell ab.

2 Lokale Eindeutigkeit von Lösungen

Beispiel 2.1. Das Anfangswertproblem

$$\begin{cases} y'(t) &= 3 \sqrt[3]{y(t)^2} \\ y(0) &= 0 \end{cases}$$

hat die Lösungen $\phi\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto 0$ und $\psi\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto t^3$ (Nachrechnen!). Es ist $\phi \neq \psi$ und sogar $\phi|_W \neq \psi|_W$ für jede 0-Umgebung $W \subseteq \mathbb{R}$.

Die DGL $y' = 3\sqrt[3]{y^2}$ erfüllt also nicht lokale Eindeutigkeit der Lösungen im Sinne der folgenden Definition.

Definition 2.2. Sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f\colon U \to \mathbb{R}^n$ eine Funktion. Wir sagen, die Differentialgleichung

$$y' = f(t, y) \tag{4}$$

erfüllt *lokale Eindeutigkeit von Lösungen*, wenn gilt: Sind $\phi_1\colon I_1 \to \mathbb{R}^n$ und $\phi_2\colon I_2 \to \mathbb{R}^n$ Lösungen von (4) und ist $\phi_1(t_0) = \phi_2(t_0)$ für ein $t_0 \in I_1 \cap I_2$, so gilt $\phi_1|_W = \phi_2|_W$ für eine t_0-Umgebung W in $I_1 \cap I_2$.

Definition 2.3. Es sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f\colon U \to \mathbb{R}^n$ eine Funktion. Wir versehen $\mathbb{R}^n$ mit einer Norm $\|\cdot\|$.

(a) Gibt es ein $L \in [0, \infty[$ derart, dass

$$\|f(t, y_2) - f(t, y_2)\| \leq L\|y_2 - y_1\| \quad \text{für alle } (t, y_1), (t, y_2) \in U,$$

so sagt man, f erfüllt eine (globale) *Lipschitzbedingung*. Man nennt dann L eine Lipschitzkonstante.

(b) Het jedes $(t, y) \in U$ eine Umgebung $V \subseteq U$ derart, dass $f|_V$ eine Lipschitzbedingung erfüllt, so sagt man, f erfüllt eine *lokale Lipschitzbedingung.*

Bemerkung. (a) Aus der globalen Lipschitzbedingung folgt die lokale Lipschitzbedingung (man nehme $V := U$).

(b) Es ist egal, welche Norm auf $\mathbb{R}^n$ benutzt wird, da alle Normen äquivalent sind (der Wert der Lipschitzkonstanten kann sich allerdings ändern beim Wechsel der Norm, im Falle einer globalen Lipschitzbedingung).

(c) Ist f wie oben Lipschitzstetig, so erfüllt f eine Lipschitzbedingung; die umgekehrte Implikation gilt nicht. (Beachten Sie, dass wir im Falle der Lipschitzbedingung $\|f(t_2, y_2) - f(t_1, y_1)\|$ nur abschätzen, wenn $t_1 = t_2$).

Satz 2.4 (Eindeutigkeitssatz). *Es sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ eine Teilmenge und $f : U \to \mathbb{R}^n$ eine Funktion. Erfüllt f eine lokale Lipschitzbedingung, so erfüllt die Differentialgleichung*

$$y' = f(t, y)$$

lokale Eindeutigkeit von Lösungen.

Beweis. Seien $\phi_1 : I_1 \to \mathbb{R}^n$ und $\phi_2 : I_2 \to \mathbb{R}^n$ Lösungen der Differentialgleichung und $t_0 \in I_1 \cap I_2$ derart, dass $\phi_1(t_0) = \phi_2(t_0)$. Wir haben eine t_0-Umgebung $W \subseteq I_1 \cap I_2$ zu finden mit $\gamma_1|_W = \gamma_2|_W$.

1. Fall. Ist $I_1 \cap I_2 = \{t_0\}$, so nehmen wir $W := \{t_0\}$.

2. Fall: Sei $I_1 \cap I_2$ ein nicht entartetes Intervall. Wir wählen eine Norm $\|\cdot\|$ auf $\mathbb{R}^n$. Da f eine lokale Lipschitzbedingung erfüllt, gibt es eine Umgebung V von $(t_0, \phi(t_0))$ in U derart, dass $f|_V$ eine Lipschitzbedingung erfüllt. Sei L eine Lipschitzkonstante. Es sei W ein in $I_1 \cap I_2$ enthaltenes, in $\mathbb{R}$ abgeschlossenes und beschränktes Intervall, das eine t_0-Umgebung in $I_1 \cap I_2$ ist. Da $W \to U$, $t \mapsto (t, \phi_j(t))$ für $j \in \{1, 2\}$ stetig ist, können wir nach Verkleinern von W annehmen, dass

$$(t, \phi_j(t)) \in V \quad \text{für } j \in \{1, 2\} \text{ und alle } t \in W.$$

Sei ℓ die Länge des Intervalls W; nach Verkleinern von W dürfen wir annehmen, dass

$$\ell L < 1.$$

Wir betrachten nun die C^1-Funktion

$$\psi : W \to \mathbb{R}^n, \quad t \mapsto \phi_2(t) - \phi_1(t).$$

Da

$$\psi(t_0) = \phi_2(t_0) - \phi_1(t_0) = 0,$$

ist $\psi(t) = \psi(t) - \psi(t_0)$ für alle $t \in W$ und somit nach dem Hauptsatz der Differential- und Integralrechnung

$$\begin{aligned}
\|\psi(t)\| &= \left\| \int_{t_0}^{t} \psi'(s)\, ds \right\| = \left\| \int_{t_0}^{t} \phi_2'(s) - \phi_1'(s)\, ds \right\| \\
&= \left\| \int_{t_0}^{t} f(s, \phi_2(s)) - f(s, \phi_1(s))\, ds \right\| \\
&\leq |t - t_0| \sup_{s \in W} \underbrace{\|f(s, \phi_2(s)) - f(s, \phi_1(s))\|}_{\leq L\|\phi_2(s) - \phi_1(s)\|} \\
&\leq \ell L \sup\{\|\phi_2(s) - \phi_1(s)\| \colon s \in W\} = \ell L \|\psi\|_\infty
\end{aligned}$$

unter Benutzung der Supremumsnorm

$$\|\psi\|_\infty := \sup\{\|\psi(s)\| \colon s \in W\}$$

von ψ. Bildung des Supremums über alle $t \in W$ liefert

$$\|\psi\|_\infty \leq \ell L \|\psi\|_\infty.$$

Da $\ell L < 1$, folgt $\|\psi\|_\infty = 0$. Also ist $\psi = 0$ und somit $\phi_1|_W = \phi_2|_W$. $\square$

Satz 2.5. *Es sei $J \subseteq \mathbb{R}$ ein Intervall, $Y \subseteq \mathbb{R}^n$ eine offene Teilmenge und*

$$f \colon J \times Y \to \mathbb{R}^n, \quad (t, y) \mapsto f(t, y)$$

eine Funktion derart, dass die partiellen Ableitungen

$$\frac{\partial f}{\partial y_j}(t, y)$$

für alle $t \in J$ und $y = (y_1, \ldots, y_n) \in Y$ existieren und stetige $\mathbb{R}^n$-wertige Funktionen von $(t, y) \in J \times Y$ sind. Dann erfüllt f eine lokale Lipschitzbedingung.

Beweis. Sei $t_0 \in J$, $y_0 \in Y$. Wir versehen $\mathbb{R}^n$ und $\mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$ mit der Maximumnorm. Für $t \in J$ betrachten wir die Funktion

$$f_t := f(t, \cdot) \colon Y \to \mathbb{R}^n, \quad y \mapsto f(t, y),$$

die durch Festhalten der ersten Variablen entsteht. Dann existiert für jedes $j \in \{1, \ldots, n\}$ die partiellen Ableitung

$$\frac{\partial f_t}{\partial y_j}(y) = \frac{\partial f}{\partial y_j}(t, y)$$

und ist eine stetige $\mathbb{R}^n$-wertige Funktion von $y \in U$. Also ist f_t stetig differenzierbar. Bezeichnet $J_{f_t}(y) \in \mathbb{R}^{n \times n}$ die Jacobimatrix von f_t an der Stelle y, so ist die Abbildung

$$J \times Y \to \mathbb{R}^{n \times n}, \quad (t, y) \mapsto J_{f_t}(y) = \Big(\frac{\partial f_i}{\partial y_j}(t, y\Big)_{i,j=1,\ldots n}$$

stetig, wobei $f = (f_1, \ldots f_n)$ in Komponenten. Also ist auch die Abbildung

$$h \colon J \times Y \to [0, \infty[, \quad (t, y) \mapsto \|J_{f_t}(y)\|_{\mathrm{op}}$$

stetig (wobei wir die Operatornorm bilden). Folglich existiert ein $\delta > 0$ derart, dass

$$h(t, y) \leq h(t_0, y_0) + 1$$

für alle (t, y) in der offenen Teilmenge

$$V := \{(t, y) \in J \times Y \colon |t - t_0| < \delta \text{ und } \|y - y_0\|_\infty < \delta\}$$

von $J \times Y$. Unter Benutzung der Kugeln $I := \{t \in J \colon |t - t_0| < \delta\}$ und $B_\delta(y_0) \subseteq \mathbb{R}^n$ ist

$$V = I \times B_\delta(y_0).$$

Da die Kugel $B_\delta(y_0)$ konvex ist, können wir den Mittelwertsatz in Integralform anwenden und erhalten für alle (t, y), $(t, z) \in V$

$$\begin{aligned}
\|f(t, z) - f(t, y)\|_\infty &= \|f_t(z) - f_t(y)\|_\infty \\
&= \left\| \int_0^1 J_{f_t}(y + t(z - y))(z - y)\, dt \right\|_\infty \\
&\leq \sup\{\|J_{f_t}(y + t(z - y))\|_{\mathrm{op}} \|z - y\|_\infty \colon t \in [0, 1]\} \\
&\leq L\, \|z - y\|_\infty
\end{aligned}$$

mit $L := h(t_0, y_0) + 1$. $\square$

Folgerung 2.6. *Ist $U \subseteq \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge und $f \colon U \to \mathbb{R}^n$*

stetig differenzierbar, so erfüllt f eine lokale Lipschitzbedingung.

Beweis. Sei $(t_0, y_0) \in U$ mit $y_0 = (y_1, \dots, y_n)$. Da U offen ist, existiert ein $\delta > 0$ derart, dass

$$J \times Y \subseteq U$$

mit $J :=]t_0 - \delta, t_0 + \delta[$ und

$$Y :=]y_1 - \delta, y_1 + \delta[\times \cdots \times]y_n - \delta, y_n + \delta[.$$

Nach Satz 2.5 existiert eine offene (t_0, y_0)-Umgebung $V \subseteq J \times Y$ derart, dass $(f|_{J\times Y})|_V = f|_V$ eine Lipschitzbedingung erfüllt. Da $J \times Y$ in $\mathbb{R} \times \mathbb{R}^n$ offen ist, ist auch V in $\mathbb{R} \times \mathbb{R}^n$ offen und somit eine Umgebung von (t_0, y_0) in $\mathbb{R} \times \mathbb{R}^n$. $\square$

Beispiele 2.7. (a) Da das Beispiel 2.1 lokale Eindeutigkeit verletzt, kann die Funktion

$$f \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (t, y) \mapsto \sqrt[3]{y^2}$$

keine lokale Lipschitzbedingung erfüllen. Alternativ folgt wir dies rechnerisch aus[2]

$$\frac{|f(0, y) - f(0, 0)|}{|y - 0|} = \frac{y^{2/3}}{y} = y^{-1/3} \to \infty$$

für $0 < y \to 0$.

(b) Für $\alpha > 0$ ist die Funktion $f \colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(t, y) \mapsto \alpha y$ stetig differenzierbar. Die DGL $y' = \alpha y$ aus Beispiel 1.1 erfüllt somit lokale Eindeutigkeit von Lösungen, nach Folgerung 2.6 und Satz 2.4.

Definition 2.8. Eine *lineare Differentialgleichung* erster Ordnung ist eine Differentialgleichung der Form

$$y'(t) = A(t)y(t) + b(t),$$

wobei $J \subseteq \mathbb{R}$ ein nicht entartetes Intervall ist, $A \colon J \to \mathbb{R}^{n \times n} \cong \mathbb{R}^{n^2}$ eine stetige matrixwertige Funktion und $b \colon J \to \mathbb{R}^n$ eine stetige vektorwertige Funktion.[3] Ist $b(t) = 0$ für alle $t \in J$, die DGL also von der Form

$$y'(t) = A(t)y(t),$$

[2] Gäbe es ein $L \in [0, \infty[$ mit $|f(t, z) - f(t, y)| \leq L\,|z - y|$ für all (t, z), (t, y) in einer Nullumgebung $V \subseteq \mathbb{R} \times \mathbb{R}$, so wäre $|f(0, y) - f(0, 0)|/|y - 0| \leq L$ für $0 < y$ so klein, dass $(0, y) \in V$, Widerspruch.

[3] Es ist also $A(t) = (a_{ij}(t))_{i,j=1,\dots,n}$ mit stetigen Funktionen $a_{ij} \colon J \to \mathbb{R}$ für $i, j \in \{1, \dots, n\}$ und $b = (b_1, \dots, b_n)$ mit stetigen Funktionen $b_1, \dots, b_n \colon J \to \mathbb{R}$.

so spricht man von einer *homogenen* linearen Differentialgleichung erster Ordnung; andernfalls heißt die lineare Differentialgleichung *inhomogen*.

Wir betrachten also die Differentialgleichung

$$y'(t) = f(t, y(t))$$

in $\mathbb{R}^n$ mit $f\colon J \times \mathbb{R}^n \to \mathbb{R}^n$, $(t, y) \mapsto A(t)y + b(t)$.

Satz 2.9. *Jede lineare Differentialgleichung $y(t) = A(t)y(t) + b(t)$ erfüllt lokale Eindeutigkeit von Lösungen.*

Beweis. Mit Notationen wir in Definition 2.8 betrachten wir

$$f\colon J \times \mathbb{R}^n \to \mathbb{R}^n, \quad (t, y) \mapsto A(t)y + b(t).$$

Für festes $t \in J$ ist $f_t\colon \mathbb{R}^n \to \mathbb{R}^n$, $y \mapsto A(t)y + b(t)$ affin-linear. Wir können also partielle Ableitungen nach $y_1, \ldots, y_n$ (mit $y = (y_1, \ldots, y_n)$) bilden und erhalten für $j \in \{1, \ldots, n\}$

$$\frac{\partial f}{\partial y_j}(t, y) = A(t)e_j = (a_{1j}(t), \ldots, a_{nj}(t))^T,$$

was eine stetige $\mathbb{R}^n$-wertige Funktion von $(t, y) \in J \times \mathbb{R}^n$ ist. Nach Satz 2.5 erfüllt f folglich eine lokale Lipschitzbedingung. Nach dem Eindeutigkeitssatz erfüllt die zugehörige Differentialgleichung somit lokale Eindeutigkeit von Lösungen. □

Definition 2.10. Ein topologischer (oder metrischer) Raum X heißt *unzusammenhängend*, wenn es offene nicht leere Teilmengen A und B von X gibt mit $A \cap B = \emptyset$ und $X = A \cup B$. Ist X nicht unzusammenhängend, so wird X *zusammenhängend* genannt.

Bemerkung 2.11. (a) Da für $A, B \subseteq X$ die Bedingung $A \cap B = \emptyset$ und $A \cup B = X$ zu $B = X \setminus A$ äquivalent ist, ist X genau dann unzusammenhängend, wenn gilt:

Es existiert eine offene und abgeschlossene, nicht leere Teilmenge $A \subseteq X$ derart, dass $A \neq X$.

(b) Per Definition ist ein topologischer Raum X also zusammenhängend, wenn für alle disjunkten offenen Teilmengen A und B in X mit $A \cup B = X$

folgt, dass $A = \emptyset$ oder $B = \emptyset$.

(c) Verneinen der Bedingung aus (a) zeigt: Ein topologischer Raum X ist genau dann zusammenhängend, wenn für jede nicht leere Teilmenge $A \subseteq X$, die in X offen und abgeschlossen ist, $A = X$ folgt.

Satz 2.12. *Jedes Intervall $I \subseteq \mathbb{R}$ ist zusammenhängend.*

Beweis. Sei A eine nicht leere Teilmenge von I, welche in I offen und abgeschlossen ist. Es existiert ein $a \in A$. Wir zeigen, dass

$$A \cap [a, \infty[= I \cap [a, \infty[\,. \tag{5}$$

Analog ist $A \cap]-\infty, a] = I \cap]-\infty, a]$ und somit $A = I$, folglich I zusammenhängend. Die Inklusion "$\subseteq$" in (5) ist trivial. Um die Inklusion "$\supseteq$" nachzuweisen, sei $b \in I \cap [a, \infty[$; wir zeigen $b \in A \cap [a, \infty[$. Dies ist trivial, wenn $b = a$. Sei nun also $b > a$. Da A in I offen ist, existiert ein $x \in]a, b]$ mit $[a, x] \subseteq A$. Setzen wir

$$J := \{x \in]a, b] \colon [a, x] \subseteq A\} \quad \text{und} \quad s := \sup J,$$

so ist $s \in [a, b] \subseteq I$. Seien $x_n \in J$ mit $x_n \to s$ für $n \to \infty$. Da $x_n \in A$ und A in I abgeschlossen ist, folgt

$$s = \lim_{n \to \infty} x_n \in A.$$

Also ist

$$[a, s] = \{s\} \cup \bigcup_{x \in]a, s[} [a, x] \subseteq A.$$

Wäre $s < b$, so gäbe es wegen der Offenheit von A in I ein $y \in]s, b]$ mit $[s, y] \subseteq A$. Dann wäre $[a, y] = [a, s] \cup [s, y] \subseteq A$, also $y \in J$. Da $y > s$, ergäbe sich ein Widerspruch zu $s = \sup J$. Also muss doch $s = b$ sein und somit ist $b = s \in A$. $\square$

Wir erwähnen für die Allgemeinbildung, dass auch die Umkehrung gilt:

Eine Teilmenge $I \subseteq \mathbb{R}$ ist genau dann zusammenhängend, wenn I ein Intervall ist.[4]

[4] Ist I kein Intervall, so gibt es Zahlen $x < z$ in I und ein $y \in \mathbb{R}$ mit $x < y < z$ und $y \notin I$. Dann sind $A :=]-\infty, y[\cap I$ und $B :=]y, \infty[\cap I$ offene Teilmengen von I und beide nicht leer (da $x \in A$, $z \in B$). Weiter ist $A \cap B = \emptyset$ und $I = A \cup B$, somit I nicht zusammenhängend.

Satz 2.13. *Es sei $f: U \to \mathbb{R}^n$ eine Funktion auf einer Teilmenge $U \subseteq \mathbb{R}\times\mathbb{R}^n$ derart, dass die Differentialgleichung*

$$y' = f(t, y) \tag{6}$$

lokale Eindeutigkeit von Lösungen erfüllt. Sind $\phi_1: I_1 \to \mathbb{R}^n$ und $\phi_2: I_2 \to \mathbb{R}^n$ Lösungen der Differentialgleichung (6) *derart, dass $\phi_1(t_0) = \phi_2(t_0)$ für ein $t_0 \in I_1 \cap I_2$, so gilt $\phi_1|_{I_1\cap I_2} = \phi_2|_{I_1\cap I_2}$.*

Beweis. Die Menge $A := \{t \in I_1 \cap I_2 : \phi_1(t) = \phi_2(t)\}$ ist per Voraussetzung nicht leer. Ist $t \in A$, so gibt es nach dem Eindeutigkeitssatz eine t-Umgebung W in $I_1 \cap I_2$ mit $W \subseteq A$. Also ist A offen in $I_1 \cap I_2$. Weiter ist A abgeschlossen in $I_1 \cap I_2$, denn ist $t \in I_1 \cap I_2$ und $(t_m)_{m\in\mathbb{N}}$ eine Folge in A mit $t_m \to t$ für $m \to \infty$, so gilt wegen der Stetigkeit von ϕ_1 und ϕ_2

$$\phi_1(t) = \phi_1\left(\lim_{m\to\infty} t_m\right) = \lim_{m\to\infty} \underbrace{\phi_1(t_m)}_{=\phi_2(t_m)} = \phi_2\left(\lim_{m\to\infty} t_m\right) = \phi_2(t)$$

und somit $t \in A$. Da $I_1 \cap I_2$ ein Intervall und somit zusammenhängend ist, folgt mit Bemerkung 2.4(c), dass $A = I_1 \cap I_2$, also $\phi_1|_{I_1\cap I_2} = \phi_2|_{I_1\cap I_2}$. □

Satz 2.14. *Es sei $f: U \to \mathbb{R}^n$ eine Funktion auf einer Teilmenge $U \subseteq \mathbb{R}\times\mathbb{R}^n$ derart, dass die Differentialgleichung $y' = f(t, y)$ lokale Eindeutigkeit der Lösungen erfüllt. Sei $(t_0, y_0) \in U$. Hat das Anfangswertproblem*

$$\begin{cases} y' &= f(t,y) \\ y(t_0) &= y_0 \end{cases} \tag{7}$$

eine Lösung, so hat es eine Lösung $\gamma_{\max}: I_{\max} \to \mathbb{R}^n$, welche im folgenden Sinne **maximal** *ist: Für jede Lösung $\gamma: I \to \mathbb{R}^n$ von* (7) *gilt $I \subseteq I_{\max}$ und $\gamma = \gamma_{\max}|_I$.*

Beweis. Sei Γ die Menge aller Lösungen $\gamma: I_\gamma \to \mathbb{R}^n$ des Anfangswertproblems (7). Dann ist

$$I_{\max} := \bigcup_{\gamma\in\Gamma} I_\gamma$$

ein Intervall. Da $\gamma|_{I_\gamma\cap I_\eta} = \eta|_{I_\gamma\cap I_\eta}$ für alle $\gamma, \eta \in \Gamma$ (nach Satz 2.13), ist

$$\gamma_{\max}: I_{\max} \to \mathbb{R}^n, \quad t \mapsto \gamma(t) \quad \text{wenn } \gamma \in \Gamma \text{ und } t \in I_\gamma$$

wohldefiniert.

Wir zeigen nun, dass an jeder Stelle t des Definitionsbereichs $I_{\max}$, an welcher man von rechts- bzw. linksseitigen Ableitungen sprechen kann, diese für $\gamma_{\max}$ existieren und durch $f(t, \gamma_{\max}(t))$ gegeben sind. Folglich ist die Funktion γ an jeder Stelle $t \in I_{\max}$ differenzierbar (also insbesondere dort stetig) und $\gamma'_{\max}(t) = f(t, \gamma_{\max}(t))$, was eine stetige Funktion von t ist. Somit ist $\gamma_{\max}$ eine C^1-Funktion und eine Lösung der Differentialgleichung $y' = f(t, y)$. Per Konstruktion gilt $I_\gamma \subseteq I_{\max}$ und $\gamma_{\max}|_{I_\gamma} = \gamma$ für jede Lösung $\gamma \colon I_\gamma \to \mathbb{R}^n$ des Anfangswertproblems (7). Insbesondere ist $\gamma_{\max}(t_0) = \gamma(t_0) = y_0$ und somit $\gamma_{\max}$ eine Lösung von (7).

Diskussion der einseitigen Ableitungen: Ist $t \in I_{\max} \cap [t_0, \infty[$ und existiert ein $s \in I_{\max}$ mit $s > t$, so ist $s \in I_\gamma$ für ein $\gamma \in \Gamma$, folglich $[t_0, s] \subseteq I_\gamma$ und $\gamma_{\max}|_{[t_0,s]} = \gamma|_{[t_0,s]}$. Die rechtsseitige Ableitung von $\gamma_{\max}$ an der Stelle t existiert also und stimmt mit

$$\gamma'(t) = f(t, \gamma(t)) = f(t, \gamma_{\max}(t)) \tag{8}$$

überein. Analog existiert die linksseitige Ableitung von $\gamma_{\max}$ an einer Stelle $t \in I_{\max} \cap]-\infty, t_0]$ und stimmt mit $f(t, \gamma_{\max}(t))$ überein, wenn ein $s \in I_{\max}$ mit $s < t$ existiert. Sei nun $t \in I_{\max} \cap [t_0, \infty[$ und existiere ein $s \in I_{\max}$ mit $s < t$. Ist $t = t_0$, so wurde die linksseitige Ableitung an der Stelle t gerade diskutiert. Ist $t > t_0$, so können wir $s \geq t_0$ wählen. Es gibt ein $\gamma \in \Gamma$ mit $t \in I_\gamma$. Dann ist $[t_0, t] \subseteq I_\gamma$ und $\gamma_{\max}|_{[t_0,t]} = \gamma|_{[t_0,t]}$; die linksseitige Ableitung von $\gamma_{\max}$ an der Stelle t existiert also und ist durch (8) gegeben. Analog existiert die rechtsseitige Ableitung von $\gamma_{\max}$ und ist gleich $f(t, \gamma_{\max}(t))$ an jeder Stelle $t \in I_{\max} \cap]-\infty, t_0]$, für die ein $s \in I_{\max}$ mit $s > t$ existiert. $\square$.

3 Elementare Lösungsverfahren

Wir stellen drei einfache Typen von Differentialgleichungen in $\mathbb{R}$ vor, für welche Lösungsverfahren bekannt sind.

Wir betrachten zunächst Differentialgleichungen $y' = f(t, y)$ derart, dass $f(t, y) = g(t)h(y)$ mit geeigneten Funktionen g und h.

Satz 3.1 (Differentialgleichungen in getrennten Veränderlichen). *Es seien $J \subseteq \mathbb{R}$ und $V \subseteq \mathbb{R}$ nicht entartete Intervalle, $g \colon J \to \mathbb{R}$ eine stetige Funktion* *und* $h\colon V \to \mathbb{R}$ *eine stetige Funktion derart, dass* $h(y) \neq 0$ *für alle* $y \in V$. *Gegeben* $t_0 \in I$ *und* $y_0 \in V$ *definieren wir* C^1*-Funktionen*

$$G\colon J \to \mathbb{R}, \quad G(t) := \int_{t_0}^{t} g(s)\, ds$$

und

$$H\colon V \to \mathbb{R}, \quad H(y) := \int_{y_0}^{y} \frac{1}{h(x)}\, dx.$$

Dann ist H *eine streng monotone Funktion,* $H(V)$ *ein nicht entartetes Intervall und* $H\colon V \to H(V)$ *ein* C^1*-Diffeomorphismus. Weiter gilt:*

(a) *Auf einem nicht entarteten Intervall* $I \subseteq J$ *mit* $t_0 \in I$ *gibt es genau dann eine Lösung* $\gamma\colon I \to \mathbb{R}$ *des Anfangswertproblems*

$$\begin{cases} y' &= g(t)h(y) \\ y(t_0) &= y_0, \end{cases}$$

wenn $G(I) \subseteq H(V)$. *Die Lösung ist dann eindeutig und gegeben durch* $\gamma(t) = H^{-1}(G(t))$ *für alle* $t \in I$.

(b) *Ist* $V \subseteq \mathbb{R}$ *ein offenes Intervall, so existiert ein in* J *offenes Intervall* $I \subseteq J$ *mit* $t_0 \in I$ *und eine Lösung* $\gamma\colon I \to \mathbb{R}$ *des Anfangswertproblems aus* (a).

Beweis. Nach dem Zwischenwertsatz ist $h(V)$ ein Intervall. Da dieses 0 nicht enthält, ist $h(y) > 0$ für alle $y \in V$ oder $h(y) < 0$ für alle $y \in V$. Im ersten Fall ist wegen $H' = 1/h$ die C^1-Funktion H streng monoton wachsend, im zweiten Fall streng fallend. Zudem wissen wir aus der Analysis 1, dass H^{-1} stetig differenzierbar ist mit

$$(H^{-1})'(H(y)) = \frac{1}{H'(y)} \quad \text{für alle } y \in V. \tag{9}$$

(a) Ist $\gamma\colon I \to \mathbb{R}$ eine Lösung des Anfangswertproblems, so ist $I \subseteq J$ und $\gamma(t) \in V$ für alle $t \in I$. Für alle $t \in I$ gilt weiter

$$\gamma'(t) = g(t)h(\gamma(t)),$$

also

$$\frac{\gamma'(t)}{h(\gamma(t))} = g(t). \tag{10}$$

Da $(H \circ \gamma)'(t) = H'(\gamma(t))\gamma'(t) = \frac{\gamma'(t)}{h(\gamma(t))}$, ist $H \circ \gamma$ eine Stammfunktion der linken Seite von (10). Integration von t_0 bis t ergibt

$$[H \circ \gamma]_{t_0}^t = [G]_{t_0}^t,$$

also

$$H(\gamma(t)) = G(t)$$

(wobei $H(y_0) = G(t_0) = 0$ benutzt wurde). Insbesondere ist $G(t) = H(\gamma(t)) \in H(V)$ für alle $t \in I$, also $G(I) \subseteq H(V)$ und die Lösung γ erfüllt $\gamma(t) = H^{-1}(G(t))$ für alle $t \in I$.

Ist umgekehrt $I \subseteq J$ ein nicht entartetes Intervall mit $t_0 \in I$ und $G(I) \subseteq H(V)$, so ist

$$\gamma\colon I \to \mathbb{R}, \quad t \mapsto H^{-1}(G(t))$$

eine C^1-Funktion mit $\gamma(t_0) = H^{-1}(G(t_0)) = H^{-1}(0) = y_0$. Wegen (9) ist

$$\gamma'(t) = (H^{-1})'(G(t))G'(t) = (H^{-1})'(H(\gamma(t)))G'(t) = \frac{1}{H'(\gamma(t))}g(t) = h(\gamma(t))g(t)$$

und somit γ eine Lösung des Anfangswertproblems aus (a).

(b) Ist V offen in $\mathbb{R}$, so auch $H(V)$. Also ist $H(V)$ eine offene 0-Umgebung in $\mathbb{R}$. Da $G\colon J \to \mathbb{R}$ stetig ist und $G(t_0) = 0$, ist $G^{-1}(H(V))$ eine Umgebung von t_0 in J und enthält als solche ein in J offenes Intervall I mit $t_0 \in I$. $\square$

Insbesondere erfüllen Differentialgleichungen in getrennten Veränderlichen (im obigen Sinn, mit $0 \notin h(V)$) also stets lokale Eindeutigkeit von Lösungen. Lässt man die Bedingung $0 \notin h(V)$ fallen, so kann lokale Eindeutigkeit der Lösungen verletzt sein (siehe Beispiel 2.1).

Beispiel 3.2. Wir betrachten das Anfangswertproblem

$$\begin{cases} y'(t) &= & t^2\, y(t) \\ y(1) &= & 2 \end{cases}$$

und fassen dieses als eine Differentialgleichung in getrennten Veränderlichen auf mit $g\colon \mathbb{R} \to \mathbb{R}$, $t \mapsto t^2$ und $h\colon\ ]0, \infty[\to \mathbb{R}$, $y \mapsto y$. Wir erhalten

$$G\colon \mathbb{R} \to \mathbb{R}, \; t \mapsto \int_1^t s^2\, ds = t^3/3 - 1/3 = (t^3 - 1)/3$$

und

$$H\colon\ ]0,\infty[\to\mathbb{R},\quad y\mapsto\int_2^y\frac{1}{x}dx=\ln(y)-\ln(2)=\ln(y/2).$$

Da $H(]0,\infty[)=\mathbb{R}$, ist $G(\mathbb{R})\subseteq H(]0,\infty[)$ erfüllt. Satz 3.1 liefert somit die Lösung

$$\gamma\colon\mathbb{R}\to\mathbb{R},\quad t\mapsto H^{-1}(G(t))=2\,e^{(t^3-1)/3}$$

des Anfangswertproblems.

Satz 3.3. (Homogene lineare Differentialgleichung 1. Ordnung). *Es sei $J\subseteq\mathbb{R}$ ein nicht entartetes Intervall und $a\colon J\to\mathbb{R}$ eine stetige Funktion. Dann erfüllt die homogene lineare Differentialgleichung*

$$y'(t)=a(t)y(t)$$

lokale Eindeutigkeit von Lösungen. Für $t_0\in J$ und $y_0\in\mathbb{R}$ ist die maximale Lösung des Anfangswertproblems

$$\begin{cases} y'(t) &= a(t)y(t)\\ y(t_0) &= y_0\end{cases}$$

die auf ganz J definierte Funktion $\gamma\colon J\to\mathbb{R}$,

$$t\mapsto y_0\,e^{\int_{t_0}^t a(s)\,ds}. \tag{11}$$

Beweis. Nach Kapitel 2 erfüllen lineare Differentialgleichungen lokale Eindeutigkeit der Lösungen. Erster Beweis: Es ist $\gamma(t_0)=y_0e^0=y_0$ und Ableiten mit der Kettenregel zeigt, dass

$$\gamma'(t)=y_0\,e^{\int_{t_0}^t a(s)\,ds}\frac{d}{dt}\int_{t_0}^t a(s)\,ds=y_0\,e^{\int_{t_0}^t a(s)\,ds}a(t)=a(t)\gamma(t).$$

Also löst γ die Differentialgleichung und auch das Anfangswertproblem.

Zweiter Beweis: Ist $y_0=0$, so definiert $\gamma(t):=0$ offenbar eine Lösung des Anfangswertproblems (und stimmt mit (11) überein). Ist $y_0>0$, so erhalten wir die Lösung γ mit Satz 3.1, angewandt mit $g(t):=a(t)$ und $h\colon\ ]0,\infty[\to\mathbb{R}$, $y\mapsto y$ (wobei $H=\ln\colon\ ]0,\infty[\to\mathbb{R}$ und $H(]0,\infty[)=\mathbb{R}$). Analog verfahren wir, wenn $y_0<0$, mit $h\colon\ ]-\infty,0[\to\mathbb{R}$, $y\mapsto y$. $\square$

Beispiele: Siehe Beispiel 1.1, Beispiel 1.6 und Beispiel 3.2.

Satz 3.4 (Inhomogene lineare Differentialgleichungen 1. Ordnung). *Es seien $J \subseteq \mathbb{R}$ ein nicht entartetes Intervall und $a\colon J \to \mathbb{R}$ sowie $b\colon J \to \mathbb{R}$ stetige Funktionen. Weiter seien $t_0 \in J$ und $y_0 \in \mathbb{R}$. Dann erfüllt die lineare Differentialgleichung $y'(t) = a(t)y(t) + b(t)$ lokale Eindeutigkeit von Lösungen. Die maximale Lösung des Anfangswertproblems*

$$\left\{ \begin{array}{rcl} y'(t) & = & a(t)y(t) + b(t) \\ y(t_0) & = & y_0 \end{array} \right. \tag{12}$$

ist die auf ganz J definierte Funktion $\gamma\colon J \to \mathbb{R}$,

$$\gamma(t) = \left(y_0 + \int_{t_0}^{t} b(\tau) e^{-\int_{t_0}^{\tau} a(s)\,ds}\, d\tau \right) e^{\int_{t_0}^{t} a(s)\,ds}.$$

Beweis. Nach Kapitel 2 erfüllen lineare Differentialgleichungen lokale Eindeutigkeit der Lösungen. Wir suchen eine Lösung der Form

$$\gamma(t) = c(t)\eta(t),$$

wobei $\eta(t) = e^{\int_{t_0}^{t} a(s)\,ds}$ eine Lösung der zugehörigen homogenen linearen Differentialgleichung $y'(t) = a(t)y(t)$ ist (diesen Ansatz nennt man auch "Variation der Konstanten.") Dann ist $c(t_0) = y_0$ und

$$\gamma'(t) = c'(t)\eta(t) + c(t)\eta'(t) = c'(t)\eta(t) + c(t)a(t)\eta(t). \tag{13}$$

Wir setzen $\gamma'(t)$ aus (13) und $\gamma(t) = c(t)\eta(t)$ in (12) ein und erhalten die Bedingung

$$c'(t)\eta(t) + c(t)a(t)\eta(t) = a(t)c(t)\eta(t) + b(t),$$

also

$$c'(t)\eta(t) = b(t),$$

also

$$c'(t) = b(t)\eta(t)^{-1} = b(t)e^{-\int_{t_0}^{t} a(s)\,ds} \quad \text{für alle } t \in J.$$

Äquivalent hierzu ist

$$c(t) = y_0 + \int_{t_0}^{t} b(\tau)e^{-\int_{t_0}^{\tau} a(s)\,ds} \quad \text{für alle } t \in J,$$

was den Beweis beendet. □

Bemerkung 3.5. (a) In der Situation von Satz 3.4 sind die auf J definierten Lösungen der zugehörigen homogenen linearen Differentialgleichung $y'(t) = a(t)y(t)$ von der Form

$$c\, e^{\int_{t_0}^t a(s)\, ds}$$

mit einer Konstanten $c \in \mathbb{R}$ (dem Anfangswert y_0 zur Zeit t_0). Beim Ansatz der "Variation der Konstanten" wird die Konstante c durch eine Funktion $c(t)$ ersetzt.

(b) Es dürfte vielen leichter fallen, sich den Ansatz der Variation der Konstanten zu merken und anzuwenden als die fertige Endformel aus Satz 3.4 (die sich aus dem Ansatz jederzeit wieder herleiten lässt).

Beispiel 3.6 (Lineare Differentialgleichungen erster Ordnung mit konstanten Koeffizienten). Wir betrachten den Spezialfall einer linearen Differentialgleichung der Form

$$y'(t) = a\, y(t) + b(t),$$

wobei $b\colon J \to \mathbb{R}$ eine stetige Funktion auf einem nicht entarteten Intervall $J \subseteq \mathbb{R}$ ist und $a \in \mathbb{R}$ eine feste reelle Zahl. Gegeben $t_0 \in J$ und $y_0 \in \mathbb{R}$ ist die maximale Lösung des Anfangswertproblems

$$\begin{cases} y'(t) &= ay(t) + b(t) \\ y(t_0) &= y_0 \end{cases}$$

dann die Funktion $\gamma\colon J \to \mathbb{R}$,

$$\gamma(t) = \left(y_0 + \int_{t_0}^t b(\tau) e^{-(\tau - t_0)a}\, d\tau \right) e^{(t-t_0)a}.$$

Dies können wir (wenn gewünscht) noch umformen zu

$$\gamma(t) = e^{(t-t_0)a} y_0 + \int_{t_0}^t e^{(t-\tau)a} b(\tau)\, d\tau. \tag{14}$$

4 Existenz von Lösungen

Wir untersuchen nun die Existenz von Lösungen zu Anfangswertproblemen und werden sehen, dass diese unter schwachen Voraussetzungen gewährleistet ist. Wieder spielen Lipschitzbedingungen eine Rolle. Wir werden die Lösungen als Fixpunkte gewisser Abbildungen erhalten. Wir kennen aus der Analysis 2 bereits den Banachschen Fixpunktsatz über Fixpunkte von Kontraktionen. Dieser könnte eingesetzt werden, würde aber schlechtere Ergebnisse (kleinere Defintionsbereiche der konstruierten Lösung) liefern als der folgende Fixpunktsatz, den wir statt dessen benutzen werden.

Satz 4.1 (Fixpunktsatz von Banach-Weissinger). *Sei (X,d) ein vollständiger metrischer Raum mit $X \neq \emptyset$ und $f\colon X \to X$ eine Lipschitz-stetige Selbstabbildung derart, dass es für alle $j \in \mathbb{N}_0$ Lipschitzkonstanten L_j für die Iterierten f^j gibt derart, dass*

$$\sum_{j=0}^{\infty} L_j < \infty;$$

hierbei ist $f^0 := \mathrm{id}_X$. Dann gilt:

(a) *f hat genau einen Fixpunkt x_∞.*

(b) *Ist $x_0 \in X$ beliebig, so gilt $f^n(x_0) \to x_\infty$ für $n \to \infty$.*

(c) *In der Situation von* (b) *haben wir für jedes $n \in \mathbb{N}_0$ die a priori Abschätzung*

$$d(x_\infty, f^n(x_0)) \leq \left(\sum_{j=n}^{\infty} L_j\right) d(f(x_0), x_0).$$

Beweis. Da $\sum_{j=1}^{\infty} L_j < \infty$, gibt es ein $N \in \mathbb{N}$ derart, dass $L_j < 1$ für alle $j \geq N$. Also ist $L_N < 1$ und somit $f^N\colon X \to X$ eine Kontraktion. Jeder Fixpunkt von f ist auch ein Fixpunkt von f^N und somit eindeutig nach dem Banachschen Fixpunktsatz. Die Funktion f hat somit höchstens einen Fixpunkt.

Sei nun $x_0 \in X$ und $x_n := f^n(x_0)$ für $n \in \mathbb{N}$. Dann gilt

$$d(x_{j+1}, x_j) \leq L_j\, d(x_1, x_0) \tag{15}$$

für alle $j \in \mathbb{N}$, denn es ist

$$d(x_{j+1}, x_j) = d(f^j(x_1), f^j(x_0)) \leq L_j\, d(x_1, x_0).$$

Mit der Dreiecksungleichung folgt für alle $n \in \mathbb{N}_0$ und $m \in \mathbb{N}$

$$d(x_{n+m}, x_n) \leq \sum_{j=n}^{n+m-1} d(x_{j+1}, x_j) \leq \left(\sum_{j=n}^{n+m-1} L_j\right) d(x_1, x_0). \tag{16}$$

Gegeben $\varepsilon > 0$ existiert ein $N \in \mathbb{N}_0$ derart, dass

$$\sum_{j=N}^{\infty} L_j \leq \varepsilon.$$

Für alle $n \geq N$ und $m \in \mathbb{N}$ gilt nach (16) also

$$d(x_{n+m}, x_n) \;\leq\; \varepsilon\, d(x_1, x_0),$$

was beliebig klein gemacht werden kann. Also ist $(x_n)_{n \in \mathbb{N}_0}$ eine Cauchyfolge in (X, d) und somit konvergent gegen ein Element $x_\infty \in X$ (da (X, d) vollständig angenommen ist). Wegen der Stetigkeit von f ist dann

$$f(x_\infty) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x_\infty,$$

also x_∞ ein Fixpunkt von f. Lassen wir $m \to \infty$ in (16), so folgt die a priori Abschätzung. □

Satz 4.2 (Quantitativer Existenzsatz). *Es seien $a < b$ und $R > 0$ reelle Zahlen, $t_0 \in [a, b]$, $y_0 \in \mathbb{R}^n$ und $\overline{B}_R(y_0) \subseteq \mathbb{R}^n$ die abgeschlossene Kugel um y_0 bezüglich einer Norm $\|\cdot\|$ auf $\mathbb{R}^n$. Weiter sei*

$$f\colon [a, b] \times \overline{B}_R(y_0) \to \mathbb{R}^n$$

eine stetige Funktion, die eine Lipschitzbedingung erfüllt und $M > 0$ eine reelle Zahl derart, dass[5]

$$\|f(t, y)\| \leq M \quad \text{für alle } (t, y) \in [a, b] \times \overline{B}_R(y_0).$$

Dann hat das Anfangswertproblem

$$\begin{cases} y'(t) &= f(t, y) \\ y(t_0) &= y_0 \end{cases}$$

eine Lösung $\gamma\colon I \to \mathbb{R}^n$ auf dem Intervall

$$I := [t_0 - \varepsilon_1, t_0 + \varepsilon_2]$$

mit $\varepsilon_1 := \min\{t_0 - a, \frac{R}{M}\}$ und $\varepsilon_2 := \min\{b - t_0, \frac{R}{M}\}$. Ist $[a, b] = [t_0 - r, t_0 + r]$ mit einem $r > 0$, so ist also

$$I = [t_0 - \varepsilon, t_0 + \varepsilon] \quad \text{mit} \quad \varepsilon := \min\{r, \tfrac{R}{M}\}.$$

[5] Ist f nicht die Nullfunktion, so können wir einfach $M := \|f\|_\infty$ nehmen.

Der Beweis zeigt auch: Schreiben wir $\phi_0 := \widetilde{y}_0$ für die konstante Funktion $I \to \mathbb{R}^n$, $t \mapsto y_0$, so definiert

$$\phi_k(t) := y_0 + \int_{t_0}^t f(s, \phi_{k-1}(s))\, ds$$

für $k \in \mathbb{N}$ eine stetige Funktion $\phi_k \colon I \to \mathbb{R}^n$ mit Werten in $\overline{B}_R(y_0)$ und ϕ_k konvergiert gleichmäßig gegen γ für $k \to \infty$.

Beweis. Es ist $F := C(I, \mathbb{R}^n)$ ein Banachraum bezüglich der Supremumsnorm $\|\cdot\|_\infty$. Die abgeschlossene Kugel um $\widetilde{y}_0$ in F vom Radius R ist

$$\overline{B}^F_R(\widetilde{y}_0) = \{\gamma \in C(I, \mathbb{R}^n) \colon (\forall t \in I)\ \underbrace{\|\gamma(t) - y_0\| \leq R}_{\Leftrightarrow \gamma(t) \in \overline{B}_R(y_0)}\}.$$

Als abgeschlossene Teilmenge eines vollständigen metrischen Raumes ist $\overline{B}^F_R(\widetilde{y}_0)$ vollständig (bzgl. der Metrik $(\gamma, \eta) \mapsto \|\gamma - \eta\|_\infty$). Gegeben $\gamma \in B^F_R(\widetilde{y}_0)$ definieren wir eine Funktion $T(\gamma) \colon I \to \mathbb{R}^n$ via

$$T(\gamma)(t) := y_0 + \int_{t_0}^t f(s, \gamma(s))\, ds.$$

Da die Funktion $s \mapsto f(s, \gamma(s))$ stetig ist, ist $T(\gamma)$ nach dem Hauptsatz der Integral- und Differentialrechung differenzierbar mit Ableitung

$$T(\gamma)(t) = \frac{d}{dt} \int_{t_0}^t f(s, \gamma(s))\, ds = f(t, \gamma(t)), \tag{17}$$

welche eine stetige $\mathbb{R}^n$-wertige Funktion von t ist. Also ist $T(\gamma) \in C^1(I, \mathbb{R}^n)$ und insbesondere $T(\gamma) \in C(I, \mathbb{R}^n) = F$. Sei $\varepsilon := \max\{\varepsilon_1, \varepsilon_2\}$. Da

$$\begin{aligned} \|T(\gamma)(t) - y_0\| &= \left\| \int_{t_0}^t f(s, \gamma(s))\, ds \right\| \\ &\leq |t - t_0| \sup\{\|f(s, \gamma(s))\| \colon s \in I\} \leq \varepsilon M \leq R \end{aligned}$$

für alle $t \in I$, ist $T(\gamma) \in \overline{B}^F_R(\widetilde{y}_0)$. Also ist

$$T \colon \overline{B}^f_R(\widetilde{y}_0) \to \overline{B}^F_R(\widetilde{y}_0), \quad \gamma \mapsto T(\gamma)$$

eine Selbstabbildung von $\overline{B}_R^F(\widetilde{y}_0)$. Wir weisen nun nach, dass T die Voraussetzungen des Fixpunktsatzes von Banach-Weissinger erfüllt. Wenn das stimmt, hat T einen Fixpunkt γ (den wir wie im Anschluss an Satz 4.2 beschrieben als gleichmäßigen Grenzwert $\gamma = \lim_{k\to\infty} T^k(\phi_0)$ berechnen können). Dann ist $\gamma = T(\gamma) \in C^1(I, \mathbb{R}^n)$ (siehe Beweisanfang),

$$\gamma(t_0) = T(\gamma)(t_0) = y_0 + \int_{t_0}^{t_0} f(s, \gamma(s))\, ds = y_0$$

und wegen (17) ist

$$\gamma'(t) = T(\gamma)'(t) = f(t, \gamma(t)),$$

somit γ eine Lösung des Anfangswertproblems $y'(t) = f(t, y(t))$, $y(t_0) = y_0$.

Per Voraussetzung gibt es eine Lipschitzkonstante L für f. Wir zeigen nun, dass für alle $j \in \mathbb{N}_0$, $\phi, \psi \in \overline{B}_R^F(\widetilde{y}_0)$ und $t \in I$

$$\|(T^j(\phi) - T^j(\psi))(t)\| \leq \frac{|t - t_0|^j L^j}{j!} \|\phi - \psi\|_\infty \leq \frac{(\varepsilon L)^j}{j!} \|\phi - \psi\|_\infty. \tag{18}$$

Ist dies wahr, so liefert Bildung des Supremumus über $t \in I$

$$\|T^j(\phi) - T^j(\psi)\|_\infty \leq \frac{(\varepsilon L)^j}{j!} \|\psi - \phi\|_\infty,$$

so dass T^j Lipschitz-stetig ist mit mit Lipschitzkonstante

$$L_j := \frac{(\varepsilon L)^j}{j!}.$$

Da $\sum_{j=0}^\infty L_j = \sum_{j=0}^\infty \frac{(\varepsilon L)^j}{j!} = e^{\varepsilon L} < \infty$, sind die Voraussetzungen des Fixpunktsatzes dann nachgewiesen. Der Beweis ist per Induktion nach $j \in \mathbb{N}_0$. Im Falle $j = 0$ gilt

$$\begin{aligned} \|T^0(\phi)(t) - T^0(\psi)(t)\| &= \|\phi(t) - \psi(t)\| \leq \|\phi - \psi\|_\infty \\ &= \frac{(t - t_0)^0 L^0}{0!} \|\phi - \psi\|_\infty = \frac{(\varepsilon L)^0}{0!} \|\phi - \psi\|_\infty. \end{aligned}$$

Ist $j \in \mathbb{N}$ und gilt die Aussage für $j - 1$ an Stelle von j, so folgt für $t \in I$ mit

$t \geq t_0$

$$
\begin{aligned}
\|T^j(\phi)(t) - T^j(\psi)(t)\| &= \|T(T^{j-1}(\phi))(t) - T(T^{j-1}(\psi))(t)\| \\
&= \left\| \int_{t_0}^{t} f(s, T^{j-1}(\phi)(s)) - f(s, T^{j-1}(\psi)(s))\, ds \right\| \\
&\leq \int_{t_0}^{t} \|f(s, T^{j-1}(\phi)(s)) - f(s, T^{j-1}(\psi)(s))\|\, ds \\
&\leq \int_{t_0}^{t} L\|T^{j-1}(\phi)(s) - T^{j-1}(\psi)(s)\|\, ds \\
&\leq L \int_{t_0}^{t} \frac{(s-t_0)^{j-1}L^{j-1}}{(j-1)!} \|\phi - \psi\|_\infty\, ds \\
&= \frac{(t-t_0)^j L^j}{j!} \|\phi - \psi\|_\infty = \frac{|t-t_0|^j L^j}{j!} \|\phi - \psi\|_\infty.
\end{aligned}
$$

Ist $t \in I$ mit $t \leq t_0$, zeigt eine analoge Rechnung

$$
\begin{aligned}
\|T^j(\phi)(t) - T^j(\psi)(t)\| &= \left\| \int_{t_0}^{t} f(s, T^{j-1}(\phi)(s)) - f(s, T^{j-1}(\psi)(s))\, ds \right\| \\
&= \left\| \int_{t}^{t_0} f(s, T^{j-1}(\phi)(s)) - f(s, T^{j-1}(\psi)(s))\, ds \right\| \\
&\leq L \int_{t}^{t_0} \frac{|s-t_0|^{j-1}L^{j-1}}{(j-1)!} \|\phi - \psi\|_\infty\, ds \\
&= L \int_{t}^{t_0} \frac{(t_0-s)^{j-1}L^{j-1}}{(j-1)!} \|\phi - \psi\|_\infty\, ds \\
&= \frac{(t_0-t)^j L^j}{j!} \|\phi - \psi\|_\infty = \frac{|t-t_0|^j L^j}{j!} \|\phi - \psi\|_\infty.
\end{aligned}
$$

Da in beiden Fällen zudem $|t - t_0| \leq \varepsilon$, ist (18) für j nachgewiesen, was den Beweis beendet. $\square$

Satz 4.3 (Qualitativer Existenzsatz von Picard-Lindelöf). *Es sei $U \subseteq \mathbb{R} \times \mathbb{R}^n$ eine offene Teilmenge, $f\colon U \to \mathbb{R}^n$ eine stetige Funktion, die eine lokale Lipschitzbedingung erfüllt und $(t_0, y_0) \in U$. Dann hat das Anfangswertproblem*

$$
\begin{cases} y'(t) &= f(t, y(t)) \\ y(t_0) &= y_0 \end{cases}
$$

eine Lösung $\gamma\colon I \to \mathbb{R}^n$ auf einem offenen Intervall $I \subseteq \mathbb{R}$ mit $t_0 \in I$.

Beweis. Wir versehen $\mathbb{R}^n$ mit der Maximumnorm $\|\cdot\|_\infty$. Da f eine lokale Lipschitzbedingung erfüllt, gibt es eine offene Umgebung V von (t_0, y_0) in U derart, dass $f|_V$ eine Lipschitzbedingung erfüllt. Sei $y_0 = (y_1, \dots, y_n)$ in Komponenten. Da V in $\mathbb{R} \times \mathbb{R}^n$ offen ist, existiert ein $R > 0$ derart, dass

$$[t_0 - R, t_0 + R] \times [y_1 - R, y_1 + R] \times \cdots \times [y_n - R, y_n + R] \subseteq V,$$

also $K := [t_0 + R, t_0 + R] \times \overline{B}_R(y_0) \subseteq V$. Da die Menge K kompakt und f stetig ist, ist $f(K)$ eine beschränkte Teilmenge von $\mathbb{R}^n$, also

$$\sup \|f(K)\|_\infty \leq M$$

für eine reelle Zahl $M > 0$. Nach Satz 4.2 hat das Anfangswertproblem $y'(t) = f(t, y(t))$, $y(t_0) = y_0$ eine Lösung $\gamma\colon [t_0 - \varepsilon, t_0 + \varepsilon] \to \mathbb{R}^n$ mit $\varepsilon := \min\{R, R/M\}$. Nun leistet $\gamma|_I$ mit $I :=]t_0 - \varepsilon, t_0 + \varepsilon[$ das Gewünschte. $\square$

Satz 4.4 (Existenzsatz für lineare Differentialgleichungen). *Es seien $J \subseteq \mathbb{R}$ ein nicht entartetes Intervall, $A\colon J \to \mathbb{R}^{n \times n}$ und $b\colon J \to \mathbb{R}^n$ stetige Funktionen und $(t_0, y_0) \in J \times \mathbb{R}^n$. Dann hat das Anfangswertproblem*

$$\begin{cases} y'(t) &= A(t)y(t) + b(t) \\ y(t_0) &= y_0 \end{cases}$$

eine auf ganz J definierte Lösung $\gamma\colon J \to \mathbb{R}^n$.

Beweis. Da nach Kapitel 2 jede lineare Differentialgleichung lokale Eindeutigkeit von Lösungen erfüllt, genügt es zu zeigen, dass die maximale Lösung $\gamma_{\max}\colon I_{\max} \to \mathbb{R}^n$ des Anfangswertproblems auf ganz J definiert ist, also $I_{\max} = J$. Können wir auf jedem Intervall $[a, b] \subseteq J$ mit $a < b$ und $t_0 \in [a, b]$ die Existenz einer Lösung des Anfangswertproblems zeigen, so ist $[a, b] \subseteq I_{\max}$ und somit $I_{\max} = J$, weil J die Vereinigung all solcher Intervalle $[a, b]$ ist. Wir dürfen also o.B.d.A. $J = [a, b]$ annehmen. Wir versehen $\mathbb{R}^n$ mit einer Norm $\|\cdot\|$ und $\mathbb{R}^{n \times n}$ mit der entsprechenden Operatornorm $\|\cdot\|_{\mathrm{op}}$. Da das Intervall $[a, b]$ kompakt und die Funktion

$$[a, b] \to [0, \infty[, \quad t \mapsto \|A(t)\|_{\mathrm{op}}$$

stetig ist, existiert

$$L := \max\{\|A(t)\|_{\mathrm{op}} \colon t \in [a, b]\}.$$

Dann erfüllt $f\colon J \times \mathbb{R}^n \to \mathbb{R}^n$, $(t, y) \mapsto A(t)y + b(t)$ eine globale Lipschitzbedingung mit Lipschitzkonstante L, denn für alle $t \in J$ und $y, z \in \mathbb{R}^n$ ist

$$\begin{aligned} \|f(t,z) - f(t,y)\| &= \|A(t)z + b(t) - A(t)y - b(t)\| = \|A(t)(z-y)\| \\ &\leq \|A(t)\|_{\mathrm{op}}\|z-y\| \leq L\|z-y\|. \end{aligned}$$

Nun ist $C(J, \mathbb{R}^n)$ ein Banachraum bzgl. der Supremumsnorm. Für $\gamma \in C(J, \mathbb{R}^n)$ definieren wir eine Funktion $T(\gamma)\colon J \to \mathbb{R}^n$ via

$$T(\gamma)(t) := y_0 + \int_{t_0}^{t} f(t, \gamma(t))\, dt.$$

Wie im Beweis von Satz 4.2 sehen wir, dass $T(\gamma) \in C^1(J, \mathbb{R}^n)$. Insbesondere ist also T eine Selbstabbildung

$$T\colon C(J, \mathbb{R}^n) \to C(J, \mathbb{R}^n).$$

Wie im Beweis von Satz 4.2 sehen wir, dass T^j für jedes $j \in \mathbb{N}_0$ Lipschitzstetig ist mit Lipschitzkonstante $(\varepsilon L)^j/j!$, wobei $\varepsilon := \max\{b - t_0, t_0 - a\}$. Nach dem Satz von Banach-Weissinger hat T einen Fixpunkt γ und wir sehen wie im Beweis von Satz 4.2, dass γ die gesuchte Lösung des Anfangswertproblems ist. $\square$

5 Lineare Differentialgleichungen

In diesem Kapitel führen wir das Studium linearer Differentialgleichungen erster Ordnung fort und studieren insbesondere die Menge aller maximalen Lösungen. Im Falle einer homogenen linearen Differentialgleichung im $\mathbb{R}^n$ ist die Lösungsmenge (wie wir sehen werden) ein n-dimensionaler Vektorraum, im Fall einer inhomogenen linearen Differentialgleichung ein affiner Raum.

Wir betrachten zunächst eine homogene lineare Differentialgleichung

$$y'(t) = A(t)y(t) \tag{19}$$

im $\mathbb{R}^n$, wobei $A\colon J \to \mathbb{R}^{n\times n}$ eine stetige matrixwertige Funktion auf einem nicht entarteten Intervall $J \subseteq \mathbb{R}$ ist. Wir schreiben L_h für die Menge aller auf ganz J definierten Lösungen $\gamma\colon J \to \mathbb{R}^n$ der Differentialgleichung (19). Dann ist also L_h eine Teilmenge des reellen Vektorraums $C^1(J, \mathbb{R}^n)$ aller $\mathbb{R}^n$-wertigen stetig differenzierbaren Funktionen auf J.

Satz 5.1. *L_h ist ein n-dimensionaler Untervektorraum von $C^1(J,\mathbb{R}^n)$. Für jedes $t_0 \in J$ ist die Punktauswertung*

$$\varepsilon_{t_0}\colon L_h \to \mathbb{R}^n, \quad \gamma \mapsto \gamma(t_0)$$

ein Isomorphismus von Vektorräumen.

Beweis. Die Nullfunktion $0\colon J \to \mathbb{R}^n$, $t \mapsto 0$ ist in L_h, denn $0'(t) = 0 = A(t)0 = A(t)0(t)$ für alle $t \in J$. Sind $\gamma, \eta \in L_h$ und $r, s \in \mathbb{R}$, so ist

$$\begin{aligned}(r\gamma + s\eta)'(t) &= r\gamma'(t) + s\eta'(t) = rA(t)\gamma(t) + sA(t)\eta(t) = A(t)(r\gamma(t) + s\eta(t))\\ &= A(t)((r\gamma + s\eta)(t))\end{aligned}$$

für alle $t \in J$, also ist $r\gamma + s\eta \in L_h$ und somit L_h ein Untervektorraum von $C^1(J,\mathbb{R}^n)$. Für $t_0 \in J$ ist ε_{t_0} eine lineare Abbildung, da

$$\varepsilon_{t_0}(r\gamma + s\eta) = (r\gamma + s\eta)(t_0) = r\gamma(t_0) + s\eta(t_0) = r\varepsilon_{t_0}(\gamma) + s\varepsilon_{t_0}(\eta).$$

Die Abbildung ε_{t_0} ist surjektiv, weil für $y_0 \in \mathbb{R}^n$ das Anfangswertproblem

$$\begin{cases} y'(t) &= A(t)y(t) \\ y(t_0) &= y_0 \end{cases}$$

nach Satz 4.4 eine Lösung $\gamma\colon J \to \mathbb{R}^n$ besitzt; dann ist nämlich $\varepsilon_{t_0}(\gamma) = \gamma(t_0) = y_0$. Aufgrund der Eindeutigket der auf J definierten Lösung des Anfangswertproblems ist γ mit $\varepsilon_{t_0}(\gamma) = y_0$ eindeutig, somit ε_{t_0} auch injektiv. Also ist ε_{t_0} eine bijektive lineare Abbildung und somit ein Isomorphismus von Vektorräumen. Insbesondere ist $L_h \cong \mathbb{R}^n$ und somit L_h (wie $\mathbb{R}^n$) ein n-dimensionaler Vektorraum. □

Definition 5.2. Da L_h ein n-dimensionaler Vektorraum ist, existiert eine Basis $\gamma_1, \ldots, \gamma_n$ des Vektorraums L_h; man nennt eine solche Basis auch ein *Lösungsfundamentalsystem* für die homogene lineare Differentialgleichung $y' = A(t)y$.

Bemerkung 5.3 (a) Gegeben $t_0 \in J$ ist nach Satz 5.1 die Abbildung

$$L_h \to \mathbb{R}^n, \quad \gamma \mapsto \gamma(t_0)$$

ein Isomorphismus von Vektorräumen. Somit bilden Funktionen $\gamma_1, \ldots, \gamma_n \in L_h$ genau dann eine Basis von L_h (also ein Lösungsfundamentalsystem), wenn die Vektoren

$$\gamma_1(t_0), \ldots, \gamma_n(t_0)$$

eine Basis von $\mathbb{R}^n$ bilden, also die Matrix

$$\Phi(t_0) := (\gamma_1(t_0), \dots, \gamma_n(t_0)) \in \mathbb{R}^{n \times n}$$

mit den Spalten $\gamma_1(t_0), \dots, \gamma_n(t_0)$ invertierbar ist.

(b) Ist $\gamma_1, \dots, \gamma_n \in L_h$ ein Lösungsfundamentalsystem der homogenen linearen Differentialgleichung $y' = A(t)y$, so lässt sich jede Lösung $\gamma \in L_h$ als Linearkombination

$$\gamma = r_1\gamma_1 + \cdots + r_n\gamma_n$$

schreiben mit eindeutig festgelegten reellen Zahlen $r_1, \dots, r_n$ (und jede solche Linearkombination ist eine Lösung). Für jedes $t \in J$ ist also

$$\gamma(t) = r_1\gamma_1(t) + \cdots + r_n\gamma_n(t) = \Phi(t)\begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}$$

mit der $(n \times n)$-Matrix $\Phi(t) = (\gamma_1(t), \dots, \gamma_n(t))$.

(c) Gegeben $t_0 \in J$ und $y_0 \in \mathbb{R}^n$ definiert

$$\gamma(t) := \Phi(t)\Phi(t_0)^{-1}y_0 \quad \text{für } t \in J \tag{20}$$

eine Lösung $\gamma \in L_h$ derart, dass $\gamma(t_0) = \Phi(t_0)\Phi(t_0)^{-1}y_0 = y_0$. Also ist (20) eine Formel für die maximale Lösung des Anfangswertproblems $y' = A(t)y$, $y(t_0) = y_0$.

Satz 5.4. *Es sei $t_0 \in J$. Für Funktionen $\gamma_1, \dots, \gamma_n \in L_h$ sind äquivalent*:

(a) $\gamma_1, \dots, \gamma_n$ *ist ein Lösungsfundamentalsystem für die homogene lineare Differentialgleichung $y' = A(t)y$;*

(b) *Die sogenannte* **Wronski-Determinante**

$$W(t_0) := \det(\gamma_1(t_0), \dots, \gamma_n(t_0))$$

ist von Null verschieden.

Beweis. Nach Bemerkung 5.3(a) ist $\gamma_1, \dots, \gamma_n$ genau dann ein Lösungsfundamentalsystem, wenn die $(n \times n)$-Matrix $\Phi(t_0) = (\gamma_1(t_0), \dots, \gamma_n(t_0))$ invertierbar ist, also $W(t_0) = \det\Phi(t_0) \neq 0$. $\square$

Beachte: Ist $W(t_0) \neq 0$ für ein $t_0 \in J$, so ist $\gamma_1, \dots, \gamma_n$ ein Lösungsfundamentalsystem, somit $W(t) \neq 0$ für jedes $t \in J$.

Beispiel 5.5 (Ein dreieckiges System). Für die homogene lineare Differentialgleichung

$$y'(t) = \begin{pmatrix} 1 & 0 \\ 2t & 1 \end{pmatrix} y(t) \tag{21}$$

(mit $J := \mathbb{R}$) wollen wir ein Lösungsfundamentalsystem finden. Hierzu lösen wir zunächst das Anfangswertproblem mit $t_0 = 0$, $y(0) = y_0$ in zwei Fällen:

Ist $y_0 = e_2 = (0,1)^T$, so ist das Anfangswertproblem (21) für $y(t) = \psi(t) = (\psi_1(t), \psi_2(t))$ äquivalent zu

$$\psi_1'(t) = \psi_1(t), \qquad \psi_1(0) = 0 \tag{22}$$

$$\psi_2'(t) = 2t\,\psi_1(t) + \psi_2(t), \quad \psi_2(0) = 1. \tag{23}$$

Die eindeutige auf $\mathbb{R}$ definierte Lösung $\psi_1 \colon \mathbb{R} \to \mathbb{R}$ von (22) ist die Nullfunktion $\psi_1 = 0$ (siehe Satz 3.3). Setzen wir dieses ψ_1 in (23) ein, so erhalten wir das Anfangswertproblem

$$\psi_2'(t) = \psi_2(t), \quad \psi_2(0) = 1$$

mit der eindeutigen auf $\mathbb{R}$ definierten Lösung $\psi_2 \colon \mathbb{R} \to \mathbb{R}$, $\psi_2(t) = e^t$. Wir haben also

$$\psi \colon \mathbb{R} \to \mathbb{R}^2, \quad \psi(t) = \begin{pmatrix} 0 \\ e^t \end{pmatrix}.$$

Ist nun $y_0 = e_1 = (1,0)^T$, so ist das Anfangswertproblem (21) für $y(t) = \phi(t) = (\phi_1(t), \phi_2(t))$ äquivalent zu

$$\phi_1'(t) = \phi_1(t), \qquad \phi_1(0) = 1 \tag{24}$$

$$\phi_2'(t) = 2t\,\phi_1(t) + \phi_2(t), \quad \phi_2(0) = 0. \tag{25}$$

Die eindeutige auf $\mathbb{R}$ definierte Lösung $\phi_1 \colon \mathbb{R} \to \mathbb{R}$ von (24) ist $\phi_1 = e^t$. Setzen wir dieses ϕ_1 in (25) ein, so erhalten wir das Anfangswertproblem

$$\phi_2'(t) = 2te^t + \phi_2(t), \quad \phi_2(0) = 0$$

mit der eindeutigen auf $\mathbb{R}$ definierten Lösung $\phi_2 \colon \mathbb{R} \to \mathbb{R}$, $\phi_2(t) = t^2\,e^t$ (vgl. Satz 3.4). Wir haben also

$$\phi \colon \mathbb{R} \to \mathbb{R}^2, \quad \phi(t) = \begin{pmatrix} e^t \\ t^2\,e^t \end{pmatrix}.$$

Da $\phi(0) = e_1$ und $\psi(0) = e_2$, erhalten wir für die Wronski-Determinante mit $t_0 = 0$

$$W(0) = \det(e_1, e_2) = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

Da $W(0) \neq 0$, ist ϕ, ψ ein Lösungsfundamentalsystem für die homogene lineare Differentialgleichung (21).

Es sei nun $A\colon J \to \mathbb{R}^{n\times n}$ eine stetige matrixwertige Funktion und zudem $b\colon J \to \mathbb{R}^n$ eine stetige vektorwertige Funktion; wir betrachten die inhomogene lineare Differentialgleichung erster Ordnung

$$y'(t) = A(t)y(t) + b(t) \tag{26}$$

in $\mathbb{R}^n$. Es sei L_i die Menge aller auf ganz J definierten Lösungen $\gamma\colon J \to \mathbb{R}^n$ der inhomogenen linearen Differentialgleichung (26). Weiter sei $L_h \subseteq C^1(J, \mathbb{R}^n)$ der Untervektorraum der auf J definierten Lösungen der zugehörigen homogenen linearen Differentialgleichung

$$y'(t) = A(t)y(t).$$

Dann gilt:

Satz 5.6. *L_i ist ein n-dimensionaler affiner Unterraum des reellen Vektorraums $C^1(J, \mathbb{R}^n)$. Für jedes $y_p \in L_i$ gilt*

$$L_i = y_p + L_h. \tag{27}$$

Beweis. Es ist $L_i \neq \emptyset$, weil es nach Satz 4.4 auf ganz J definierte Lösungen der betrachteten inhomogenen linearen Differentialgleichung gibt. Für jedes $\gamma \in C^1(J, \mathbb{R}^n)$ ist

$$\alpha(\gamma)\colon J \to \mathbb{R}^n, \quad \alpha(\gamma)(t) := \gamma'(t) - A(t)\gamma(t)$$

eine stetige Funktion und die Abbildung

$$\alpha\colon C^1(J, \mathbb{R}^n) \to C(J, \mathbb{R}^n), \quad \gamma \mapsto \alpha(\gamma)$$

ist offenbar linear. Da $\ker(\alpha) = L_h$ und $L_i = \{\gamma \in C^1(J, \mathbb{R}^n)\colon \alpha(\gamma) = b\}$, können wir elementare Resultate über lineare Gleichungen anwenden: Wählen wir $\gamma_p \in L_i$, so ist

$$L_i = \gamma_p + L_h$$

(da sich je zwei Lösungen einer inhomogenen Gleichung $\alpha(\gamma) = b$ nur durch Addition einer Lösung der homogenen Gleichung unterscheiden und umgekehrt Addieren einer Lösung der homogenen Gleichung zu γ_p wieder auf eine Lösung der inhomogenen Gleichung führt). □

In traditioneller Sprechweise kann man (27) auch so lesen:

Die allgemeine Lösung einer inhomogenen linearen Differentialgleichung ist die Summe einer speziellen ("partikulären") Lösung der inhomogenen Gleichung und der allgemeinen Löung der zugehörigen homogenen linearen Differentialgleichung.

Satz 5.7 (Variation der Konstanten). *Es seien $J \subseteq \mathbb{R}$ ein nicht entartetes Intervall, $A\colon J \to \mathbb{R}^{n\times n}$, $t \mapsto A(t)$ eine stetige matrixwertige Funktion, $b\colon J \to \mathbb{R}^n$ eine stetige vektorwertige Funktion und $t_0 \in J$, $y_0 \in \mathbb{R}^n$. Weiter sei $\gamma_1, \ldots, \gamma_n \in C^1(J, \mathbb{R}^n)$ ein Lösungsfundamentalsystem der homogenen linearen Differentialgleichung*

$$y'(t) = A(t)y(t)$$

und

$$\Phi(t) := (\gamma_1(t), \ldots, \gamma_n(t)) \in \mathrm{GL}_n(\mathbb{R}).$$

Dann ist $\gamma\colon J \to \mathbb{R}^n$,

$$\gamma(t) = \Phi(t)\left(\Phi(t_0)^{-1}y_0 + \int_{t_0}^{t} \Phi(s)^{-1}b(s)\,ds\right) \tag{28}$$

die auf ganz J definierte Lösung des Anfangswertproblems

$$y'(t) = A(t)y(t) + b(t), \quad y(t_0) = y_0.$$

Beweis. Für die Lösung $\gamma\colon J \to \mathbb{R}^n$ des obigen Anfangswertproblems machen wir den Ansatz

$$\gamma(t) = \Phi(t)c(t)$$

mit einer C^1-Funktion $c\colon J \to \mathbb{R}^n$ (und werden sehen, dass dies funktioniert!) Die Anfangsbedingung ist äquivalent zu

$$y_0 = \gamma(t_0) = \Phi(t_0)c(t_0),$$

also

$$c(t_0) = \Phi(t_0)^{-1}(y_0). \tag{29}$$

Nun ist

$$\gamma'(t) = \Phi'(t)c(t) + \Phi(t)c'(t) = A(t)\Phi(t)c(t) + \Phi(t)c'(t),$$

Einsetzen in die DGL führt auf die äquivalente Bedingung

$$A(t)\Phi(t)c(t) + \Phi(t)c'(t) = A(t)\Phi(t)c(t) + b(t),$$

also

$$\Phi(t)c'(t) = b(t),$$

also

$$c'(t) = \Phi(t)^{-1}b(t).$$

Für die Stammfunktion erfordert dies

$$c(t) = c(t_0) + \int_{t_0}^{t} \Phi(s)^{-1}b(s)\,ds = \Phi(t_0)^{-1}y_0 + \int_{t_0}^{t} \Phi(s)^{-1}b(s)\,ds,$$

wobei $c(t_0)$ aus (29) eingesetzt wurde. Also ist $\gamma(t) = \Phi(t)c(t)$ von der Form (28). $\square$

Bemerkung 5.8 Ist $\gamma_1, \ldots, \gamma_n \in L_h \subseteq C^1(J, \mathbb{R}^n)$ ein Lösungsfundamentalsystem von

$$y'(t) = A(t)y(t)$$

mit einer stetigen matrixwertigen Funktion $A\colon J \to \mathbb{R}^{n\times n}$ und $I \subseteq J$ ein nicht entartetes Intervall, so ist $\gamma_1|_I, \ldots, \gamma_n|_I$ ein Lösungsfundamentalsystem von $y'(t) = A|_I(t)y(t)$.

Denn offenbar sind $\gamma_1|_I, \ldots, \gamma_n|_I$ auf I definierte Lösungen der letzteren Differentialgleichung. Wählen wir $t_0 \in I$, so erhalten wir für die Wronski-Determinate zudem

$$\det(\gamma_1|_I(t_0), \ldots, \gamma_n|_I(t_0)) = \det(\gamma_1(t_0), \ldots, \gamma_n(t_0)) \neq 0.$$

6 Lineare Systeme erster Ordnung mit konstanten Koeffizienten

Wir betrachten nun homogene lineare Differentialgleichungen $y'(t) = Ay(t)$ mit einer festen Matrix $A \in \mathbb{R}^{n\times n}$. Am Ende des Kapitels gehen wir auch kurz auf inhomogene lineare Differentialgleichungen der Form

$$y'(t) = Ay(t) + b(t)$$

ein, die wiederum *konstante Koeffizienten* haben in dem Sinn, dass $A \in \mathbb{R}^{n\times n}$ eine feste Matrix ist (und nicht wie bisher eine allgemeine stetige matrixwertige Funktion). Das Hauptergebnis des Kapitels ist der folgende Satz.

Satz 6.1. *Es seien $A \in \mathbb{R}^{n\times n}$, $t_0 \in \mathbb{R}$ und $y_0 \in \mathbb{R}^n$. Dann ist*

$$\gamma\colon \mathbb{R} \to \mathbb{R}^n, \quad t \mapsto e^{(t-t_0)A}y_0$$

die maximale Lösung des Anfangswertproblems

$$y'(t) = Ay(t), \qquad y(t_0) = y_0.$$

Zudem gilt: Seien $\gamma_1(t), \ldots, \gamma_n(t)$ für $t \in \mathbb{R}$ die Spalten der Matrix $e^{(t-t_0)A}$. Dann bilden $\gamma_1, \ldots, \gamma_n\colon \mathbb{R} \to \mathbb{R}^n$ ein Lösungsfundamentalystem für die homogene lineare DGL $y'(t) = Ay(t)$.

Wir benutzen hierbei die Matrix-Exponentalfunktion.

Definition 6.2. Ist $A \in \mathbb{R}^{n\times n}$ eine reelle $(n \times n)$-Matrix, so definieren wir eine reelle $(n \times n)$-Matrix e^A via

$$e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k. \tag{30}$$

Allgemeiner definieren wir $e^A \in \mathbb{C}^{n\times n}$ via (30) für jede komplexe $(n \times n)$-Matrix A. Die Abbildung

$$\exp\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$$

(und die entsprechende Abbildung $\mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$) nennt man die *Matrix-Exponentialfunktion.*

Man beachte, dass die Exponentialreihe in (30) für alle $A \in \mathbb{C}^{n\times n}$ konvergiert. Um dies einzusehen, sei $\|\cdot\|$ eine Norm auf $\mathbb{C}^n$ und $\|\cdot\|_{\mathrm{op}}\colon \mathbb{C}^{n\times n} \to [0, \infty[$ die entsprechende Operatornorm,

$$\|A\|_{\mathrm{op}} := \sup\{\|Ax\|\colon x \in \mathbb{C}^n\colon \|x\| \leq 1\} \in [0, \infty[\,.$$

Wie jede Norm (die ja alle äquivalent sind) definiert $\|\cdot\|_{\mathrm{op}}$ die Topologie von $C^{n\times n} \cong \mathbb{C}^{n^2}$. Die Operatornom ist submultiplikativ: Es gilt

$$\|AB\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}}\|B\|_{\mathrm{op}}$$

für alle $A, B \in \mathbb{C}^{n\times n}$, woraus $\|A^k\|_{\text{op}} \leq \|A\|_{\text{op}}^k$ für alle $k \in \mathbb{N}$ folgt. Weiter gilt $\|\mathbf{1}\|_{\text{op}} = 1$ und somit $\|A^k\|_{\text{op}} = \|\mathbf{1}\|_{\text{op}} = 1 = \|A\|_{\text{op}}^k$ auch für $k = 0$. Nach dem Vorigen ist

$$\sum_{k=0}^{\infty} \left\| \frac{1}{k!} A^k \right\|_{\text{op}} \leq \sum_{k=0}^{\infty} \frac{1}{k!} (\|A\|_{\text{op}})^k = e^{\|A\|_{\text{op}}} < \infty.$$

Die Reihe $\sum_{k=0}^{\infty} \frac{1}{k!} A^k$ ist somit absolut konvergent und folglich konvergent, da $(\mathbb{C}^{n\times n}, \|\cdot\|_{\text{op}})$ ein Banachraum ist (siehe Prof. Glöckners Analysis 2-Skript vom Sommersemester 2019 in Panda, Satz 9.18).

Wir müssen die Matrixexponentialfunktion zunächst besser verstehen: Zum einen, um Satz 6.1 beweisen zu können; zum anderen, um die im Satz vorkommende Matrix $e^{(t-t_0)A}$ konkreter ausrechnen zu können.

Satz 6.3 (Eigenschaften der Matrixexponentialfunktion). *Es seien $A, B \in \mathbb{C}^{n\times n}$. Dann gilt:*

(a) *Für die Nullmatrix $\mathbf{0} \in \mathbb{R}^{n\times n}$ ist $e^{\mathbf{0}} = \mathbf{1} \in \mathbb{R}^{n\times n}$ die Einheitsmatrix.*

(b) (Diagonalmatrizen). *Für all $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ gilt*

$$e^{\operatorname{diag}(\lambda_1,\ldots,\lambda_n)} = \operatorname{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}).$$

(c) (Blockdiagonalmatrizen). *Ist $A = \operatorname{diag}(A_1, \ldots, A_r)$ eine Blockdiagonalmatrix, wobei $n = n_1 + \cdots + n_r$ und $A_j \in \mathbb{C}^{n_j \times n_j}$ für $j \in \{1, \ldots, r\}$, so ist*

$$e^{\operatorname{diag}(A_1,\ldots,A_r)} = \operatorname{diag}(e^{A_1}, \ldots, e^{A_r}).$$

(d) *Ist $T \in \mathrm{GL}_n(\mathbb{C})$, so gilt $e^{TBT^{-1}} = Te^B T^{-1}$.*

(e) *Vertauschen die Matrizen A und B (d.h. gilt $AB = BA$), so ist $e^{A+B} = e^A e^B$.*

(f) *Für alle $s, t \in \mathbb{R}$ ist $e^{(t+s)A} = e^{tA} e^{sA}$.*

(g) *Es ist $e^A \in \mathrm{GL}_n(\mathbb{C})$.*

(h) *Die Matrixexponentialfunktion $\exp\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $C \mapsto e^C$ ist differenzierbar an der Stelle $\mathbf{0}$ mit $\exp'(\mathbf{0}) = \mathrm{id}$, also $\exp'(\mathbf{0})(C) = C$ für alle $C \in \mathbb{C}^{n\times n}$.*

(i) *Es ist* $\frac{d}{dt}e^{tA} = Ae^{tA} = e^{tA}A$.

(j) *Ist* A *nilpotent mit* $A^m = \mathbf{0}$, *so ist* $e^A = \sum_{k=0}^{m-1} \frac{1}{k!}A^k$.

Bemerkung 6.4. Satz 6.3 ermöglicht uns (zumindest im Prinzip), für jede Matrix $A \in \mathbb{C}^{n\times n}$ und jedes $t \in \mathbb{R}$ die Matrix

$$e^{tA}$$

explizit zu berechnen, wie folgt: Nach dem Satz über die Jordansche Normalform aus der Linearen Algebra 2 existiert eine Matrix $B \in \mathbb{C}^{n\times n}$ in Jordanscher Normalform derart, dass

$$A = TBT^{-1}$$

mit einer invertierbaren Matrix $T \in \mathrm{GL}_n(\mathbb{C})$. Es ist also B eine Blockdiagonalmatrix

$$B = \mathrm{diag}(J_1, \dots, J_r)$$

mit gewissen Jordan-Blöcken $J_1, \dots, J_r$. Dann ist

$$tA = T(tB)T^{-1}$$

mit $tB = \mathrm{diag}(tJ_1, \dots, tJ_r)$ und nach Satz 6.3 (d) und (c) folglich

$$e^{tA} = Te^{tB}T^{-1} = T\,\mathrm{diag}(tJ_1, \dots, tJ_r)T^{-1}.$$

Wir müssen also nur noch e^{tJ} berechnen können für einen Jordan-Block $J \in \mathbb{C}^{m\times m}$ der Form

$$J = \begin{pmatrix} \lambda & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 0 & \lambda & 1 \\ 0 & \cdots & 0 & 0 & 0 & \lambda \end{pmatrix} = \lambda\mathbf{1} + N$$

mit $\lambda \in \mathbb{C}$ und

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{m\times m}.$$

In den Potenzen N^k rutscht die Nebendiagonale mit Einsen in jedem Schritt eins weiter in die rechte obere Ecke; ist etwa $m = 4$, so ist

$$N = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad N^4 = \mathbf{0}$$

und somit

$$e^{tN} = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 \\ 0 & 1 & t & t^2/2 \\ 0 & 0 & 1 & t \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Für allgemeines $m \in \mathbb{N}$ ist analog

$$e^{tN} = \sum_{k=0}^{m-1} \frac{1}{k!} t^k N^k = \begin{pmatrix} 1 & t & t^2/2 & t^3/6 & \cdots & t^{m-1}/(m-1)! \\ 0 & 1 & t & t^2/2 & \cdots & t^{m-2}/(m-2)! \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 1 & t \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Da die Matrix $t\lambda\mathbf{1}$ (als Vielfaches der Einheitsmatrix) mit tN vertauscht, folgt mit Satz 6.2 (b) und (e) weiter

$$e^{tJ} = e^{t\lambda\mathbf{1}+tN} = e^{t\lambda\mathbf{1}}e^{tN} = e^{t\lambda}e^{tN} = e^{t\lambda} \begin{pmatrix} 1 & t & t^2/2 & t^3/6 & \cdots & t^{m-1}/(m-1)! \\ 0 & 1 & t & t^2/2 & \cdots & t^{m-2}/(m-2)! \\ \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & \cdots & 0 & 0 & 1 & t \\ 0 & \cdots & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Mithilfe der Jordanschen Normalform können wir e^{tA} also berechnen (und insbesondere e^A), oder jedenfalls im Prinzip – da wir die Transformationsmatrix T und die Jordansche Normalform B von A natürlich nicht immer explizit kennen (dies erfordert ja Kenntnis der Eigenwerte und somit Kenntnis der Nullstellen des charakteristischen Polynoms von A, eines Polynoms $n$ten Grades!)

Beweis von Satz 6.3. (a) Da $\mathbf{0}^k = \mathbf{0}$ für alle $k \in \mathbb{N}$, ist $e^{\mathbf{0}} = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{0}^k = \frac{1}{0!}\mathbf{0}^0 = \mathbf{1}$.

(b) Für jedes $k \in \mathbb{N}_0$ ist $(\operatorname{diag}(\lambda_1, \ldots, \lambda_n))^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k)$, somit

$$\sum_{k=0}^{N} \frac{1}{k!} (\operatorname{diag}(\lambda_1, \ldots, \lambda_n))^k = \operatorname{diag}\left(\sum_{k=0}^{N} \frac{1}{k!} \lambda_1^k, \ldots, \sum_{k=0}^{N} \frac{1}{k!} \lambda_n^k \right).$$

Für $N \to \infty$ folgt die Formel aus (b).

(c) Wird analog bewiesen; man ersetze $\lambda_1, \ldots, \lambda_n$ durch die Diagonalblöcke $A_1, \ldots, A_r$ im Beweis von (b).

(d) Für $N \in \mathbb{N}_0$ ist

$$\sum_{k=0}^{N} \frac{1}{k!} \underbrace{(TBT^{-1})^k}_{=TB^kT^{-1}} = T \left(\sum_{k=0}^{N} \frac{1}{k!} B^k \right) T^{-1}.$$

Es folgt

$$\begin{aligned} Te^BT^{-1} &= T \left(\lim_{N\to\infty} \sum_{k=0}^{N} \frac{1}{k!} B^k \right) T^{-1} = \lim_{N\to\infty} T \left(\sum_{k=0}^{N} \frac{1}{k!} B^k \right) T^{-1} \\ &= \lim_{N\to\infty} \sum_{k=0}^{N} \frac{1}{k!} (TBT^{-1})^k = e^{TBT^{-1}}. \end{aligned}$$

(e) Die Reihen $\sum_{k=0}^{\infty} \frac{1}{k!} A^k$ und $\sum_{k=0}^{\infty} \frac{1}{k!} B^k$ sind absolut konvergent. Da die Matrixmultiplikation

$$\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$$

bilinear ist, können wir die Cauchysche Produktformel anwenden (in der Form von Lemma 6.5). Die Reihe

$$\sum_{k=0}^{\infty} c_k \quad \text{mit} \quad c_k := \sum_{j=0}^{k} \frac{1}{j!} A^j \frac{1}{(k-j)!} B^{k-j}$$

ist daher absolut konvergent und es ist

$$e^A e^B = \sum_{k=0}^{\infty} \frac{1}{k!} A^k \sum_{k=0}^{\infty} \frac{1}{k!} B^k = \sum_{k=0}^{\infty} c_k = e^{A+B},$$

da

$$c_k = \frac{1}{k!}\sum_{j=0}^{k} \frac{k!}{j!(k-j)!}A^j B^{k-j} = \frac{1}{k!}\sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \end{array}\right) A^j B^{k-j} = \frac{1}{k!}(A+B)^k$$

unter Benutzung des binomischen Lehrsatzes (der in jeder assoziativen reellen Algebra für $(A+B)^k$ mit kommutierenden Elementen A und B gültig ist, per Induktion nach $k \in \mathbb{N}$ wie in der Analysis 1).

(f) Da die Matrizen tA und sA miteinander vertauschen, gilt nach (e) $e^{(t+s)A} = e^{tA+sA} = e^{tA}e^{sA}$.

(g) Da A und $-A$ vertauschen, gilt nach (e) und (a) $\mathbf{1} = e^{\mathbf{0}} = e^{A+(-A)} = e^A e^{-A}$. Also ist e^A invertierbar mit Inverser e^{-A}.

(h) Wir wählen eine Norm $\|\cdot\|$ auf $\mathbb{C}^n$ und schreiben $\|\cdot\|_{\mathrm{op}}$ für die zugehörige Operatornorm auf $\mathbb{C}^{n\times n}$. Sei $\mathrm{id}\colon \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ die identische Abbildung. Für $C \in \mathbb{C}^{n\times n}$ ist dann

$$\exp(C) = \sum_{k=0}^{\infty} \frac{1}{k!}C^k = \mathbf{0} + C + \sum_{k=2}^{\infty} \frac{1}{k!}C^k = \exp(\mathbf{0}) + \mathrm{id}(C) + R(C)$$

mit

$$R(C) := \sum_{k=2}^{\infty} \frac{1}{k!}C^k.$$

Wir haben zu zeigen, dass

$$\lim_{C\to\mathbf{0}} \frac{R(C)}{\|C\|_{\mathrm{op}}} = \mathbf{0},$$

oder äquivalent dazu, dass

$$\lim_{C\to\mathbf{0}} \frac{\|R(C)\|_{\mathrm{op}}}{\|C\|_{\mathrm{op}}} = 0. \tag{31}$$

Da die reelle Exponentialfunktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto e^x$ an der Stelle 0 differenzierbar ist mit Ableitung $e^0 = 1$ und Funktionswert $e^0 = 1$, haben wir eine affin-lineare Approximation

$$e^x = 1 + x + \rho(x)$$

mit $\lim_{x\to 0} \frac{|\rho(x)|}{|x|} = 0$. Gegeben $\varepsilon > 0$ gibt es also ein $\delta > 0$ derart, dass

$$|\rho(x)| \le \varepsilon|x| \quad \text{für alle } x \in \mathbb{R} \text{ mit } |x| \le \delta.$$

Hierbei ist

$$\rho(x) = e^x - 1 - x = \sum_{k=2}^{\infty} \frac{1}{k!} x^k.$$

Für alle $C \in \mathbb{C}^{n\times n}$ mit $\|C\|_{\mathrm{op}} \leq \delta$ folgt nun

$$\|R(C)\|_{\mathrm{op}} \leq \sum_{k=2}^{\infty} \frac{1}{k!} \|C^k\|_{\mathrm{op}} \leq \sum_{k=2}^{\infty} \frac{1}{k!} (\|C\|_{\mathrm{op}})^k = \rho(\|C\|_{\mathrm{op}}) \leq \varepsilon \|C\|_{\mathrm{op}}.$$

Also gilt (31).

(i) Nach (g) ist die Funktion $\mathbb{R} \to \mathbb{C}^{n\times n}$, $t \mapsto e^{tA}$ an der Stelle $t = 0$ differenzierbar mit

$$\frac{d}{dt}\Big|_{t=0} e^{tA} = \mathrm{id}(A) = A.$$

Nun ist nach (f)

$$e^{(t+s)A} = e^{tA} e^{sA} = e^{sA} e^{tA}.$$

Also existiert die Ableitung $\frac{d}{dt} e^{tA}$ an der Stelle t und ist gegeben durch

$$\frac{d}{dt} e^{tA} = \frac{d}{ds}\Big|_{s=0} e^{(t+s)A} = \frac{d}{ds}\Big|_{s=0} (e^{tA} e^{sA}) = e^{tA} \frac{d}{ds}\Big|_{s=0} e^{sA} = e^{tA} A.$$

Ebenso ist $\frac{d}{dt} e^{tA} = \frac{d}{ds}\big|_{s=0} (e^{sA} e^{tA}) = A e^{tA}$.

(j) Ist $A^m = \mathbf{0}$, so ist auch $A^k = \mathbf{0}$ für alle $k \geq m$ und folglich $e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = \sum_{k=0}^{m-1} \frac{1}{k!} A^k$. $\square$

Es sei daran erinnert, dass jede bilineare Abbildung $\beta\colon E_1 \times E_2 \to F$ automatisch stetig ist, wenn E_1, E_2 und F endlich-dimensionale Vektorräume sind (vgl. Prof. Glöckners Skript zur Analysis 2 im SoSe 2019 in Panda, Satz 10.6(b)). Sind $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_F$ Normen auf E_1, E_2 bzw. F, so ist also

$$\|\beta\|_{\mathrm{op}} := \sup\{\|\beta(x_1, x_2)\|_F \colon (x_1, x_2) \in E_1 \times E_2 \text{ mit } \|x_1\|_1, \|x_2\|_2 \leq 1\} < \infty.$$

Insbesondere ist die bilineare Matrixmultiplikation $\beta\colon \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$, $\beta(A, B) := AB$ stetig, was allerdings auch direkt klar ist, denn $\|\beta\|_{\mathrm{op}} \leq 1$ wegen

$$\|\beta(A, B)\|_{\mathrm{op}} = \|AB\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} \|B\|_{\mathrm{op}} \leq 1$$

für alle $A, B \in \mathbb{C}^{n\times n}$ mit $\|A\|_{\mathrm{op}}, \|B\|_{\mathrm{op}} \leq 1$ (vgl. auch Satz 10.5 im genannten Skript). Wir wenden folgendes Lemma nur auf $E_1 := E_2 := F := \mathbb{C}^{n\times n}$ an.

Lemma 6.5 (Cauchysche Produktformel). *Es seien* $(E_1, \|\cdot\|_1)$, $(E_2, \|\cdot\|_2)$ *und* $(F, \|\cdot\|_F)$ *reelle Banachräume und* $\beta\colon E_1 \times E_2 \to F$ *eine stetige bilineare Abbildung. Weiter sei* $\sum_{k=0}^\infty a_k$ *eine absolut konvergente Reihe in* E_1 *und* $\sum_{k=0}^\infty b_k$ *eine absolut konvergente Reihe in* E_2. *Für alle* $k \in \mathbb{N}_0$ *sei*

$$c_k := \sum_{j=0}^k \beta(a_j, b_{k-j}).$$

Dann gilt:

(a) *Die Reihe* $\sum_{k=0}^\infty c_k$ *ist in* F *absolut konvergent.*

(b) *Es ist* $\beta\Big(\sum_{k=0}^\infty a_k, \sum_{k=0}^\infty b_k\Big) = \sum_{k=0}^\infty c_k$.

Beweis. Die Reihen $\sum_{k=0}^\infty \|a_k\|_1$ und $\sum_{k=0}^\infty \|b_k\|_2$ sind oer Voraussetzung absolut konvergent. Nach der aus der Analysis 1 bekannten Cauchyschen Produktformel für reelle Zahlenfolgen ist auch die Zahlenreihe

$$\sum_{k=0}^\infty d_k \quad \text{mit} \quad d_k := \sum_{j=0}^k \|a_j\|_1 \|b_{k-j}\|_2$$

absolut konvergent, mit

$$P := \sum_{k=0}^\infty d_k = \Big(\sum_{k=0}^\infty \|a_k\|_1\Big)\Big(\sum_{k=0}^\infty \|b_k\|_2\Big) = \lim_{m\to\infty} \sum_{k=0}^m \|a_k\|_1 \sum_{\ell=0}^m \|b_\ell\|_2.$$

Die Reihe $\sum_{k=0}^\infty c_k$ ist in F nach dem Majorantenkriterium absolut konvergent, da

$$\|c_k\|_F = \left\|\sum_{j=0}^k \beta(a_j, b_{k-j})\right\|_F \le \sum_{j=0}^k \|\beta(a_j, b_{k-j})\|_F \le \|\beta\|_{\mathrm{op}} \sum_{j=0}^k \|a_j\|_1 \|b_{k-j}\|_2 = \|\beta\|_{\mathrm{op}} d_k$$

mit $\sum_{k=0}^\infty \|\beta\|_{\mathrm{op}} d_k = \|\beta\|_{\mathrm{op}} \sum_{k=0}^\infty d_k < \infty$. Gegeben $\varepsilon > 0$ gibt es ein $m_0 \in \mathbb{N}$ derart, dass

$$\Big|P - \sum_{k=0}^m d_k\Big| < \varepsilon/2 \quad \text{und} \quad \Big|P - \sum_{k=0}^m \sum_{\ell=0}^m \|a_k\|_1 \|b_\ell\|_2\Big| \le \varepsilon/2$$

für all $m \geq m_0$ und somit

$$\begin{aligned}
\varepsilon &\geq \Big|\sum_{k,\ell=0}^{m} \|a_k\|_1 \|b_\ell\|_2 - \sum_{k=0}^{m} d_k\Big| \\
&= \Big|\sum_{k,\ell=0}^{m} \|a_k\|_1 \|b_\ell\|_2 - \sum_{k+\ell\leq m} \|a_k\|_1 \|b_\ell\|_2\Big| \\
&= \Big|\sum_{(k,\ell)\in\Delta_m} \|a_k\|_1 \|b_\ell\|_2\Big| = \sum_{(k,\ell)\in\Delta_m} \|a_k\|_1 \|b_\ell\|_2
\end{aligned}$$

mit

$$\Delta_m := \{(k,\ell) \in \{0,1,\ldots,m\}^2 \colon k+\ell > m\}.$$

Also ist

$$\begin{aligned}
&\Big\|\beta\Big(\sum_{k=0}^{m} a_k, \sum_{\ell=0}^{m} b_\ell\Big) - \sum_{k=0}^{m} c_k\Big\|_F \\
&\qquad = \Big\|\sum_{k,\ell=0}^{m} \beta(a_k,b_\ell) - \sum_{k=0}^{m} c_k\Big\|_F \\
&\qquad = \Big\|\sum_{k,\ell=0}^{m} \beta(a_k,b_\ell) - \sum_{k+\ell\leq m} \beta(a_k,b_\ell)\Big\|_F \\
&\qquad = \Big\|\sum_{(k,\ell)\in\Delta_m} \beta(a_k,b_\ell)\Big\|_F \\
&\qquad \leq \sum_{(k,\ell)\in\Delta_m} \|\beta(a_k,b_\ell)\|_F \\
&\qquad \leq \|\beta\|_{\mathrm{op}} \sum_{(k,\ell)\in\Delta_m} \|a_k\|_1 \|b_\ell\|_2 \leq \|\beta\|_{\mathrm{op}}\, \varepsilon
\end{aligned}$$

und für $m \to \infty$ folgt

$$\Big\|\beta\Big(\sum_{k=0}^{\infty} a_k, \sum_{\ell=0}^{\infty} b_\ell\Big) - \sum_{k=0}^{\infty} c_k\Big\|_F \leq \|\beta\|_{\mathrm{op}}\, \varepsilon,$$

was beliebig klein gemacht werden kann. Die linke Seite ist also 0 und somit gilt $\beta\big(\sum_{k=0}^{\infty} a_k, \sum_{\ell=0}^{m} b_\ell\big) = \sum_{k=0}^{m} c_k$. $\square$

Beweis von Satz 6.1. Nach Satz 6.3(a) ist $\gamma(0) = e^{0A}y_0 = \mathbf{1}y_0 = y_0$. Aus Satz 6.3(i) folgt, dass γ differenzierbar ist, mit

$$\gamma'(t) = \frac{d}{dt}\big(e^{(t-t_0)A}y_0\big) = \big(\frac{d}{dt}e^{(t-t_0)A}\big)y_0 = Ae^{(t-t_0)A}y_0 = A\gamma(t).$$

Dies ist eine stetige Funktion von t (weil γ als differenzierbare Funktion stetig ist), folglich γ eine C^1-Funktion und die Lösung des im Satz genannten Anfangswertproblems.

Mit $y_0 := e_j$ sehen wir insbesondere, dass $\gamma_j(t) := e^{(t-t_0)A}e_j$ eine Lösung $\gamma_j\colon \mathbb{R} \to \mathbb{R}^n$ des Anfangswertproblems $y'(t) = Ay(t)$, $y(t_0) = e_j$ definiert, für alle $j \in \{1, \ldots, n\}$. Also sind $\gamma_1, \ldots, \gamma_n\colon \mathbb{R} \to \mathbb{R}^n$ Lösungen der homogenen linearen Differentialgleichung $y'(t) = Ay(t)$. Diese bilden ein Lösungsfundamentalsystem, denn die Wronski-Determinante hat für t_0 den Wert

$$W(t_0) = \det(\gamma_1(t_0), \ldots, \gamma_n(t_0)) = \det(e_1, \ldots, e_n) = \det(\mathbf{1}) = 1 \neq 0. \qquad \square$$

Bemerkung 6.6. Wir untersuchen nun, wie sich die mit Variation der Konstanten bewiesene Lösungsformel für Anfangswertprobleme der Form $y(t) = A(t)y(t) + b(t)$ aus Satz 5.7 vereinfachen lässt im Falle einer inhomogenen linearen Differentialgleichung *mit konstanten Koeffizienten.*

Sei $J \subseteq \mathbb{R}$ ein nicht entartetes Intervall, $b\colon J \to \mathbb{R}^n$ eine stetige vektorwertige Funktion, $A \in \mathbb{R}^{n\times n}$ eine feste Matrix, $t_0 \in J$ und $y_0 \in \mathbb{R}^n$. Setzen wir $\gamma_j(t) := e^{(t-t_0)A}e_j$ für $j \in \{1, \ldots, n\}$ und $t \in J$, so bilden die Funktonen $\gamma_1, \ldots, \gamma_n\colon J \to \mathbb{R}^n$ ein Lösungsfundamentalsystem der homogenen linearen Differentialgleichung $y'(t) = Ay(t)$, welche wir als eine Differentialgleichung mit rechter Seite

$$f\colon J \times \mathbb{R}^n \to \mathbb{R}^n, \quad (t, y) \mapsto Ay$$

auffassen (siehe Satz 6.1 und Bemerkung 5.8). Für jedes $t \in J$ ist dann

$$\Phi(t) := (\gamma_1(t), \ldots, \gamma_n(t)) = e^{(t-t_0)A}$$

und insbesondere $\Phi(t_0) = e^{0A} = \mathbf{1}$. Die Lösungsformel aus Satz 5.7 für die auf J definierte Lösung $\gamma\colon J \to \mathbb{R}^n$ des Anfangswertproblems

$$y'(t) = Ay(t), \quad y(t_0) = y_0$$

nimmt somit folgende Gestalt an:

$$\gamma(t) = e^{(t-t_0)A}\left(y_0 + \int_{t_0}^{t} e^{-(\tau-t_0)A}b(\tau)\,d\tau\right) = e^{(t-t_0)A}y_0 + \int_{t_0}^{t} e^{(t-\tau)A}b(\tau)\,d\tau,$$

wobei $e^{(t-t_0)A}e^{(t_0-\tau)A} = e^{(t-t_0)A+(t_0-\tau)A} = e^{(t-\tau)A}$ gerechnet wurde.

7 Differentialgleichungen höherer Ordnung, insb. lineare mit konstanten Koeffizienten

In diesem Kapitel führen wir Differentialgleichungen $m$ter Ordnung und den Begriff einer Lösung einer solchen Differentialgleichung ein. Eine Differentialgleichung höherer Ordnung lässt sich immer als ein System von Differentialgleichungen erster Ordnung "umschreiben", so dass wir auf die Theorie der vorigen Kapitel zurückgreifen können. Von besonderer Wichtigkeit in den Anwendungen sind lineare Differentialgleichungen in $\mathbb{R}$ mit konstanten Koeffizienten, denen die zweite Hälfte des Kapitels gewidmet ist.

Definition 7.1 Es seien $n, m \in \mathbb{N}$ und $f\colon U \to \mathbb{R}^n$ eine Abbildung auf einer Teilmenge $U \subseteq \mathbb{R} \times (\mathbb{R}^n)^m$. Eine *Differentialgleichung mter Ordnung in* $\mathbb{R}^n$ mit rechter Seite f ist eine Gleichung der Form

$$y^{(m)}(t) \;=\; f(t, y(t), y'(t), \ldots, y^{(m-1)}(t)). \tag{32}$$

Eine *Lösung* der Differentialgleichung (32) ist eine C^m-Funktion $\gamma\colon I \to \mathbb{R}^n$ auf einem nicht entarteten Intervall $I \subseteq \mathbb{R}$ derart, dass

$$(t, \gamma(t), \ldots, \gamma^{(m-1)}(t)) \in U \quad \text{für alle } t \in I$$

und

$$\gamma^{(m)}(t) = f(t, \gamma(t), \ldots, \gamma^{(m-1)}(t)).$$

Ist $\gamma\colon I \to \mathbb{R}^n$ eine Lösung von (32), so können wir C^1-Funktionen

$$\eta_1, \ldots, \eta_m\colon I \to \mathbb{R}^n$$

definieren via

$$\begin{array}{rcl} \eta_1 & := & \gamma \\ \eta_2 & := & \gamma' \\ \vdots & & \vdots \\ \eta_m & := & \gamma^{(m-1)}. \end{array}$$

Dann gilt also

$$\eta_j' = \eta_{j+1} \quad \text{für alle } \; j \in \{1, \ldots, m-1\}. \tag{33}$$

Weiter gilt

$$\eta_m'(t) = (\gamma^{(m-1)})'(t) = \gamma^{(m)}(t) = f(t, \gamma(t), \ldots, \gamma^{(m-1)}(t)) = f(t, \eta_1(t), \ldots, \eta_m(t)). \tag{34}$$

Nach (33) und (34) erfüllt $\eta = (\eta_1, \ldots, \eta_m)\colon I \to (\mathbb{R}^n)^m \cong \mathbb{R}^{nm}$ also das System von Differentialgleichungen erster Ordnung

$$\begin{cases} u_1'(t) & = & u_2(t) \\ u_2'(t) & = & u_3(t) \\ \vdots & & \vdots \\ u_{m-1}'(t) & = & u_m(t) \\ u_m'(t) & = & f(t, u_1(t), \ldots, u_m(t)). \end{cases}$$

Kompakter hingeschrieben erfüllt η somit die Differentialgleichung

$$u'(t) = g(t, u(t)) \tag{35}$$

mit $u = (u_1, \ldots, u_m)$ und

$$g\colon U \to (\mathbb{R}^m)^n \cong \mathbb{R}^{nm}, \quad g(t, u_1, \ldots, u_m) := \begin{pmatrix} u_2 \\ u_3 \\ \vdots \\ u_m \\ f(t, u_1, \ldots, u_m). \end{pmatrix}.$$

Sei umgekehrt $\eta = (\eta_1, \ldots, \eta_m)\colon I \to (\mathbb{R}^n)^m$ eine Lösung der Differentialgleichung (35) mit den C^1-Funktionen $\eta_1, \ldots, \eta_m\colon I \to \mathbb{R}^n$. Per Induktion ist dann η_j eine C^{1+m-j}-Funktion für alle $j = m, m-1, \ldots, 1$. In der Tat ist ja η_m eine C^1-Funktion und ist $j \in \{1, \ldots, m-1\}$ und η_{j+1} eine $C^{1+m-j-1}$-Funktion, so ist

$$\eta_j' = \eta_{j+1}$$

eine $C^{1+m-j-1}$-Funktion und somit η_j eine $C^{(1+m-j-1)+1}$-Funktion, also C^{1+m-j}. Insbesondere ist $\gamma := \eta_1$ eine C^m-Funktion und per Induktion ist

$$\gamma^{(j)} = \eta_{j+1}$$

für alle $j \in \{0, 1, \ldots, m-1\}$. Folglich ist

$$\gamma^{(m)}(t) = (\gamma^{(m-1)})'(t) = \eta_m'(t) = f(t, \eta_1(t), \ldots, \eta_m(t)) = f(t, \gamma(t), \ldots, \gamma^{(m-1)}(t)),$$

also γ eine Lösung der Differentialgleichung (32) $m$ter Ordnung.

Wir haben gezeigt:

Satz 7.2 (Umschreiben als System erster Ordnung)

(a) *Für jede Lösung* $\gamma\colon I \to \mathbb{R}^n$ *von* (32) *ist* $\eta := (\gamma, \gamma', \ldots, \gamma^{(m-1)})$ *eine Lösung von* (35).

(b) *Für jede Lösung* $\eta = (\eta_1, \ldots, \eta_m)\colon I \to (\mathbb{R}^n)^m$ *von* (35) *ist* $\gamma := \eta_1$ *eine Lösung von* (32). $\square$

Definition 7.3 Eine Differentialgleichung der Form[6]

$$y^{(m)}(t) + a_{m-1}(t)y^{(m-1)}(t) + \cdots + a_1(t)y'(t) + a_0(t)y(t) = b(t)$$

mit stetigen Funktionen $a_0, \ldots, a_{m-1}\colon J \to \mathbb{R}$ und $b\colon J \to \mathbb{R}$ auf einem nicht entarteten Intervall $J \subseteq \mathbb{R}$ heißt *lineare Differentialgleichung* m*ter Ordnung* in $\mathbb{R}$. Ist b die Nullfunktion, so heißt die lineare Differentialgleichung *homogen*, andernfalls *inhomogen*. Sind die Funktionen $a_0, \ldots, a_{m-1}$ konstant, die DGL also von der Form

$$y^{(m)}(t) + a_{m-1}y^{(m-1)}(t) + \cdots + a_1y'(t) + a_0y(t) = b(t)$$

mit reellen Zahlen $a_0, a_1, \ldots, a_{m-1}$, so spricht man von einer linearen Differentialgleichung $m$ter Ordnung *mit konstanten Koeffizienten*.

Satz 7.4 *Seien* $J \subseteq \mathbb{R}$ *ein nicht entartetes Intervall und* $a_0, \ldots, a_{m-1}\colon J \to \mathbb{R}$ *sowie* $b\colon J \to \mathbb{R}$ *stetige Funktionen. Weiter sei* $L_h \subseteq C^m(J, \mathbb{R})$ *die Menge aller auf* J *definierten Lösungen* $\gamma\colon J \to \mathbb{R}$ *der homogenen linearen Differentialgleichung*

$$y^{(m)}(t) + a_{m-1}(t)y^{(m-1)}(t) + \cdots + a_1(t)y'(t) + a_0(t)y(t) = 0. \qquad (36)$$

Dann ist L_h *ein* m*-dimensionaler Untervektorraum von* $C^m(J, \mathbb{R})$*. Für jedes* $t_0 \in J$ *ist die Abbildung*

$$L_h \to \mathbb{R}^m, \quad \gamma \mapsto (\gamma(t_0), \gamma'(t_0), \ldots, \gamma^{(m-1)}(t_0)) \qquad (37)$$

[6] Also $y^{(m)}(t) = -a_{m-1}(t)y^{(m-1)} - \cdots - a_1(t)y'(t) - a_0(t)y(t) + b(t)$.

ein Isomorphismus von reellen Vektorräumen. Die Menge $L_i \subseteq C^m(J,\mathbb{R})$ aller auf ganz J definierten Lösungen der inhomogenen linearen DGL

$$y^{(m)}(t) + a_{m-1}(t)y^{(m-1)}(t) + \cdots + a_1(t)y'(t) + a_0(t)y(t) = b(t) \qquad (38)$$

ist nicht leer und von der Form $L_i = \gamma_p + L_h$ mit einem beliebigen $\gamma_p \in L_i$.

Beweis. (a) Sei $\widetilde{L}_h \subseteq C^1(J,\mathbb{R}^m)$ die Menge der auf J definierten Lösungen der homogenen linearen Differentialgleichung

$$\begin{array}{rcl} u_1'(t) & = & u_2(t) \\ & \vdots & \\ u_{m-1}'(t) & = & u_m(t) \\ u_m'(t) & = & -a_{m-1}(t)u_m(t) - \cdots - a_0(t)u_1(t) \end{array}$$

erster Ordnung in $\mathbb{R}^n$, also der Differentialgleichung

$$u'(t) = A(t)u(t)$$

mit der stetigen matrixwertigen Funktion $A\colon J \to \mathbb{R}^{m\times m}$,

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ & & & \ddots & & \\ 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 \\ -a_0(t) & -a_1(t) & \cdots & -a_{m-3}(t) & -a_{m-2}(t) & -a_{m-1}(t) \end{pmatrix}. \qquad (39)$$

Dies ist das zur linearen homogenen Differentialgleichung (36) gehörige System (35). Nach Satz 5.1 ist $\widehat{L}_h$ ein m-dimensionaler Untervektorraum von $C^1(J,\mathbb{R}^m)$ und die Abbildung

$$\varepsilon_{t_0}\colon \widetilde{L}_h \to \mathbb{R}^m, \quad \eta \mapsto \eta(t_0)$$

ist ein Isomorphismus von Vektorräumen. Nach Satz 7.2 ist die Abbildung

$$\Psi\colon L_h \to \widetilde{L}_h, \quad \gamma \mapsto (\gamma, \gamma', \ldots, \gamma^{(m-1)})$$

eine Bijektion mit Umkehrabbildung $\widetilde{L}_h \to L_h$, $\eta = (\eta_1, \ldots, \eta_m) \mapsto \eta_1$. Offenbar ist Ψ eine lineare Abbildung und somit ein Isomorphismus reeller Vektorräume. Also ist

$$\dim_{\mathbb{R}}(L_h) = \dim_{\mathbb{R}} \widetilde{L}_h = m$$

und $\varepsilon_{t_0} \circ \Phi\colon \mathbb{R}^m$ ist ein Isomorphismus; dies ist die Abbildung aus (37).

(b) Ist $u = (u_1, \ldots, u_n)$ eine auf J definierte Lösung der inhomogenen linearen Differentiagleichung

$$u'(t) = A(t)u(t) + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix},$$

so ist $u_1 \in L_i$ nach Satz 7.2. Also ist $L_i \neq \emptyset$. Ist $\gamma_p \in L_i$, so ist $\gamma_p + L_h \subseteq L_i$ (wie man sofort nachrechnet) und wir haben Gleichheit der zwei Mengen, denn ist $\gamma \in L_i$, so ist offenbar $\gamma - \gamma_p \in L_h$ und $\gamma = \gamma_p + (\gamma - \gamma_p)$. $\square$

Ein m-Tupel von Lösungen $\gamma_1, \ldots, \gamma_m \in L_h$ wird ein *Lösungsfundamentalsystem* der homogenen linearen Differentialgleichung (36) $m$ter Ordnung genannt, wenn $\gamma_1, \ldots, \gamma_m$ eine Basis von L_h ist.

Bemerkung 7.5. (a) Gegeben $t_0 \in J$ und $y_1, \ldots, y_m \in \mathbb{R}$ ist eine Lösung $\gamma\colon J \to \mathbb{R}$ der linearen Differentialgleichung (38) eindeutig festgelegt durch Vorgabe der Anfangsbedingungen

$$\gamma(t_0) = y_1,\ \gamma'(t_0) = y_2, \ldots, \gamma^{(m-1)}(t_0) = y_m.$$

Im Falle einer homogenen linearen Differentialgleichung (36) ist dies gerade die Bijektvität der Abbildung (37). Sind im inhomogenen Fall $\gamma, \theta \in L_i$ Lösungen mit

$$\gamma^{(j)}(t_0) = \eta^{(j)}(t_0) \quad \text{für alle } j \in \{0, 1, \ldots, m-1\},$$

so ist $\gamma - \theta \in L_h$ und $(\gamma - \theta)^{(j)}(t_0) = 0$ für alle $j \in \{0, 1 \ldots, m-1\}$, somit $\gamma - \theta = 0$ und folglich $\gamma = \theta$.

(b) Insbesondere ist eine auf J definierte Lösung γ einer linearen Differentialgleichung zweiter Ordnung eindeutig festgelegt durch Vorgabe der Anfangsposition $\gamma(t_0)$ und der Anfangsgeschwindigkeit $\gamma'(t_0)$.

(c) Gegeben stetige Funktionen $a_0, \ldots, a_{m-1}\colon J \to \mathbb{C}$ (an Stelle von reellwertigen Funktionen) sei $L_h^{\mathbb{C}}$ die Menge aller Lösungen $\gamma\colon J \to \mathbb{C}^m \cong \mathbb{R}^{2m}$ der Differentialgleichung $m$ter Ordnung

$$y^{(m)}(t) + a_{m-1}(t)y^{(m-1)}(t) + \cdots + a_1(t)y'(t) + a_0 y(t) = 0$$

in $\mathbb{C}^m \cong \mathbb{R}^{2m}$. Man rechnet sofort nach, dass $L_h^{\mathbb{C}}$ ein komplexer Untervektorraum des komplexen Vektorraums $C^m(J, \mathbb{C})$ ist. Dann definiert (39) eine stetige matrixwertige Funktion $A\colon J \to \mathbb{C}^{m\times m}$ und es ist

$$u'(t) = A(t)u(t) \tag{40}$$

eine homogene lineare Differentalgleichung erster Ordnung in $\mathbb{C}^m \cong \mathbb{R}^{2m}$. Man rechnet sofort nach, dass die Menge $\widetilde{L}_h^{\mathbb{C}}$ aller auf ganz J definierten Lösungen $\eta\colon J \to \mathbb{C}^m$ der Differentialgleichung (40) ein komplexer Untervektorraum des komplexen Vektorraums $C^1(J, \mathbb{C}^m)$ ist. Als reeller Untervektorraum von $C^1(J, \mathbb{C}^m) \cong C^1(J, \mathbb{R}^{2m})$ hat $\widetilde{L}_h^{\mathbb{C}}$ nach Satz 5.1 die Dimension $2m$; somit[7] ist die Dimension von $\widetilde{L}_h^{\mathbb{C}}$ als komplexer Vektorraum gleich m. Nach Satz 7.2 ist die Abbildung

$$\Psi\colon L_h^{\mathbb{C}} \to \widetilde{L}_h^{\mathbb{C}}, \quad \gamma \mapsto (\gamma, \gamma', \ldots, \gamma^{(m-1)})$$

eine Bijektion und somit ein Isomorphismus komplexer Vektorräume, weil Ψ offenbar $\mathbb{C}$-linear ist. Also hat $L_h^{\mathbb{C}}$ als komplexer Vektorraum die Dimension m (ebenso wie $\widetilde{L}_h^{\mathbb{C}}$).

Satz 7.6 *Seien $\gamma_1, \ldots, \gamma_m$ auf J definierte Lösungen einer homogenen linearen Differentialgechung* (36) *und sei $t_0 \in J$. Dann sind äquivalent*:

(a) *$\gamma_1, \ldots, \gamma_m$ bilden ein Lösungsfundamentalsystem für* (36);

(b) *Die Spaltenvektoren*

$$\begin{pmatrix} \gamma_1(t) \\ \gamma_1'(t) \\ \vdots \\ \gamma_1^{(m-1)}(t) \end{pmatrix}, \ldots, \begin{pmatrix} \gamma_m(t) \\ \gamma_m'(t) \\ \vdots \\ \gamma_m^{(m-1)}(t) \end{pmatrix} \tag{41}$$

sind linear unabhängig in $\mathbb{R}^m$, d.h. die $(m \times m)$-Matrix

$$\Phi(t_0) := \begin{pmatrix} \gamma_1(t_0) & \cdots & \gamma_m(t_0) \\ \gamma_1'(t_0) & \cdots & \gamma_m'(t_0) \\ \vdots & \vdots & \vdots \\ \gamma_1^{(m-1)}(t_0) & \cdots & \gamma_m^{(m-1)}(t_0) \end{pmatrix}$$

ist invertierbar;

[7] Ist $b_1, \ldots, b_m$ eine $\mathbb{C}$-Basis eines $\mathbb{C}$-Vektorraums V, so ist $b_1, \ldots, b_m, ib_1, \ldots, ib_m$ eine $\mathbb{R}$-Basis von V als reeller Vektorraum, wie man sofort nachrechnet.

(c) *Es gilt*

$$W(\gamma_1, \ldots, \gamma_m)(t_0) := \det \Phi(t_0) \neq 0.$$

Beweis. Da die Abbildung aus (37) ein Isomorphismus ist, bilden $\gamma_1, \ldots, \gamma_m$ genau dann eine Basis von L_h, wenn die Spaltenvektoren aus (41) eine Basis von $\mathbb{R}^m$ bilden. Aus der linearen Algebra ist bekannt, dass diese Bedingung zur Invertierbarkeit der Matrix $\Phi(t_0)$ mit den gegebenen Spalten äquivalent ist und somit auch zu $\det \Phi(t_0) \neq 0$. □

Man nennt $W(\gamma_1, \ldots, \gamma_m)(t_0)$ die zugehörige *Wronski-Determinante* zur Zeit t_0.

Das Hauptergebnis dieses Kapitels ist der folgende Satz.

Satz 7.7. *Es seien $m \in \mathbb{N}$ und $a_0, a_1, \ldots, a_{m-1} \in \mathbb{R}$. Weiter seien*

$$\lambda_1, \ldots, \lambda_k \in \mathbb{R}$$

und

$$\alpha_1 + i\beta_1, \ \ldots, \ \alpha_\ell + i\beta_\ell,$$
$$\alpha_1 - i\beta_1, \ \ldots, \ \alpha_\ell - i\beta_\ell$$

mit $\alpha_1, \ldots, \alpha_\ell \in \mathbb{R}$ und $\beta_1, \ldots, \beta_\ell \in \mathbb{R} \setminus \{0\}$ die paarweise verschiedenen Nullstellen des Polynoms

$$p \colon \mathbb{C} \to \mathbb{C}, \quad p(t) := t^m + a_{m-1}t^{m-1} + \cdots + a_1 t + a_0. \tag{42}$$

Es sei n_j die Vielfachheit der Nullstelle λ_j von p für $j \in \{1, \ldots, k\}$ und m_j die Vielfachheit der Nullstellen $\alpha_j + i\beta_j$ sowie $\alpha_j - i\beta_j$ für $j \in \{1, \ldots, \ell\}$. Dann bilden die folgenden reellwertigen Funktionen von $t \in \mathbb{R}$ zusammen ein Lösungsfundamentalsystem für die homogene lineare Differentialgleichung

$$y^{(m)} + a_{m-1}y^{(m-1)} + \cdots + a_1 y' + a_0 y \ = \ 0 \tag{43}$$

mit konstanten Koeffizienten:

- $t^a e^{\lambda_j t}$ *für $j \in \{1, \ldots, k\}$ und $a \in \{0, 1, \ldots, n_j - 1\}$;*
- $t^a e^{\alpha_j t} \sin(\beta_j t)$ *für $j \in \{1, \ldots, \ell\}$ und $a \in \{0, 1, \ldots, m_j - 1\}$;*
- $t^a e^{\alpha_j t} \cos(\beta_j t)$ *für $j \in \{1, \ldots, \ell\}$ und $a \in \{0, 1, \ldots, m_j - 1\}$.*

Das Polynom p aus (42) wird das *charakteristische Polynom* der homogenen linearen Differentialgleichung (43) mit konstanten Koeffizienten genannt.

Bevor wir Satz 7.7 beweisen, schauen wir uns zwei für die Physik relevante Beispiele an.

Beispiel 7.8 Die Differentialgleichung

$$y'' + \omega^2 y = 0$$

mit $\omega > 0$ beschreibt in der Physik einen harmonischen Oszillator. Dies ist eine homogene lineare Differentialgleichung 2. Ordnung. Das charakteristische Polynom

$$p(z) = z^2 + \omega^2$$

hat die Nullstellen $i\omega$ und $-i\omega$, also ein Paar komplex konjugierter nicht reeller Nullstellen der Vielfachheit 1. Es ist $i\omega = \alpha + i\beta$ mit $\alpha = 0$, $\beta = \omega$. Nach Satz 7.7 bilden also die Funktionen

$$t \mapsto e^{0t}\sin(\omega t) = \sin(\omega t) \quad \text{und} \quad t \mapsto \cos(\omega t)$$

ein Lösungsfundamentalsystem.

Beispiel 7.9. Seien $k, \omega > 0$. In der Physik beschreibt die homogene lineare Differentalgleichung

$$y'' + ky' + \omega^2 y = 0$$

einen gedämpften Oszillator. Das charakteristische Polynom lautet

$$p(z) = z^2 + kz + \omega^2$$

und wir haben

$$0 = p(z) = z^2 + kz + \omega^2 = (z + k/2)^2 + \omega^2 - k^2/4$$

genau dann, wenn

$$(z + k/2)^2 = k^2/4 - \omega^2. \tag{44}$$

1. Fall (Schwingfall): Ist $k^2/4 - \omega < 0$, so können wir (44) schreiben als

$$(z + k/2)^2 = -(\omega^2 - k^2/2)$$

mit $\omega^2 + k^2/2 > 0$. Wir haben also ein Paar komplex konjugierter, nicht reeller Nullstellen,

$$z_1 = -k/2 + i\sqrt{\omega^2 - k^2/4}, \quad z_2 = -k/2 - i\sqrt{\omega^2 - k^2/4}$$

der Vielfachheit 1. Nach Satz 7.7 bilden also die Funktionen

$$t \mapsto e^{-\frac{k}{2}t}\sin(t\sqrt{\omega^2 - k^2/4}) \quad \text{und} \quad t \mapsto e^{-\frac{k}{2}t}\cos(t\sqrt{\omega^2 - k^2/4})$$

ein Lösungsfundamentalsystem.

2. Fall (Kriechfall). Ist $k/2/4 - \omega^2 > 0$, so hat p die zwei reellen Nullstellen

$$\lambda_1 = -k/2 - \sqrt{k^2/4 - \omega^2} \quad \text{und} \lambda_2 = -k^2/2 + \sqrt{k^2/4 - \omega^2}$$

der Vielfachheit 1. Offenbar ist $\lambda_1 < 0$ und es ist auch $\lambda_2 < 0$, weil $\sqrt{k^2/4 - \omega^2} < \sqrt{k^2/4} = k/2$. Nach Satz 7.7 bilden die Funktionen

$$t \mapsto e^{t\lambda_1} \quad \text{und} \quad t \mapsto e^{t\lambda_2}$$

ein Lösungsfundamentalsystem.

3. Fall (aperiodischer Grenzfall). Ist $k^2/4 - \omega^2 = 0$, so hat p die doppelte Nullstelle

$$\lambda = -k/2.$$

Nach Satz 7.7 bilden also die Funktionen

$$t \mapsto e^{-\frac{k}{2}t} \quad \text{und} \quad t \mapsto t\,e^{-\frac{k}{2}t}$$

ein Lösungsfundamentalsystem.

Die folgenden drei Hilfssätze bereiten den Beweis von Satz 7.7 vor.

Lemma 7.10. *Es sei E ein komplexer Vektorraum und $\tau\colon E \to E$ eine antilineare Involution, also eine reell lineare Selbstabbildung derart, dass*

$$\tau(ix) = -i\tau(x) \quad \text{für alle } x \in E$$

und $\tau \circ \tau = \mathrm{id}_E$. Dann ist

$$V := \{x \in E \colon \tau(x) = x\}$$

ein reeller Untervektorraum von E und $E = V \oplus iV$ als reeller Vektorraum.

Beweis. Wegen $\tau(0) = 0$ ist $0 \in V$; sind $x, y \in V$ und $r, s \in \mathbb{R}$, so ist

$$\tau(rx + sy) = r\tau(x) + s\tau(y) = rx + sy,$$

somit $rx + sy \in V$. Also ist V ein reeller Untervektorraum von E (Alternativ: V ist der Eigenraum zum Eigenwert 1 der reell-linearen Selbstabbildung $\tau \colon E \to E$, also ein reeller Untervektorraum von E). Gegeben $x \in E$ sind

$$\mathrm{Re}(x) := \frac{1}{2}(x + \tau(x)) \in V \quad \text{und} \quad \mathrm{Im}(x) := \frac{1}{2i}(x - \tau(x)) \in V,$$

denn

$$\tau(\mathrm{Re}(x)) = \frac{1}{2}(\tau(x) + \tau^2(x)) = \frac{1}{2}(\tau(x) + x) = \mathrm{Re}(x)$$

unter Benutzung von $\tau \circ \tau = \mathrm{id}_E$ und

$$\tau(\mathrm{Im}(x)) = -\frac{1}{2i}(\tau(x) - x) = \mathrm{Im}(x).$$

Weiter ist

$$x = \frac{1}{2}(x + \tau(x)) + i\frac{1}{2i}(x - \tau(x)) = \mathrm{Re}(x) + i\,\mathrm{Im}(x).$$

Also ist $E = V + iV$. Ist $x \in V \cap iV$, so ist $x = iy$ für ein $y \in V$ und somit

$$iy = x = \tau(x) = \tau(iy) = -i\tau(y) = -iy,$$

woraus $y = 0$ und $x = iy = 0$ folgt. Also ist $E = V \oplus iV$ als reeller Vektorraum. $\square$

Für $z \in \mathbb{C}$ und $x \in E$ ist übrigens

$$\tau(zx) = \overline{z}\tau(x),$$

denn ist $z = \alpha + i\beta$ mit $\alpha, \beta \in \mathbb{R}$, so ist

$$\tau(zx) = \tau(\alpha x) + \tau(i\beta x) = \tau(\alpha x) - i\tau(\beta x) = \alpha\tau(x) - i\beta\tau(x) = (\alpha - i\beta)\tau(x) = \overline{z}\tau(x).$$

Beispiel. Ein Beispiel für die Situation von Lemma 7.10 ist $E := C^\infty(\mathbb{R}, \mathbb{C})$ und $\tau \colon C^\infty(\mathbb{R}, \mathbb{C}) \to C^\infty(\mathbb{R}, \mathbb{C})$, $f \mapsto \overline{f}$ mit $\overline{f}(t) := \overline{f(t)}$ (unter Benutzung komplexer Konjugation). In diesem Fall ist $V = C^\infty(\mathbb{R},\mathbb{R})$ und offenbar $C^\infty(\mathbb{R},\mathbb{C}) = C^\infty(\mathbb{R},\mathbb{R}) \oplus i\, C^\infty(\mathbb{R},\mathbb{R})$ als reeller Vektorraum.

Lemma 7.11. *Es seien E, τ und V mit $E = V \oplus iV$ wie in Lemma* 7.10. *Weiter sei*

$$x_1,\ \ldots,\ x_n,$$
$$a_1 + ib_1,\ \ldots,\ a_m + ib_m,$$
$$a_1 - ib_1,\ \ldots,\ a_m - ib_m$$

eine Basis des komplexen Vektorraums E, wobei $x_1,\ldots,x_n,a_1,\ldots,a_m,b_1,\ldots,b_m \in V$. Dann ist

$$x_1,\ldots,x_n,a_1,\ldots,a_m,b_1,\ldots,b_m$$

eine Basis des reellen Vektorraums V.

Beweis. Die Dimension des komplexen Vektorraums E ist $d = n + 2m$. Betrachten wir E lediglich als reellen Vektorraum, so ist seine Dimension $2d$. Da $E = V \oplus iV$ als reeller Vektorraum, wobei $V \cong iV$ als reeller Vektorraum via $x \mapsto ix$, ist

$$2d = 2\dim_{\mathbb{R}}(V).$$

Der reelle Vektorraum V hat also die Dimension $d = n + 2m$.

Zur Abkürzung sei $y_j := a_j + ib_j$ für $j \in \{1,\ldots,m\}$; dann ist also $\tau(y_j) = a_j - ib_j$. Gegeben $x \in V$ ist

$$x = r_1x_1 + \cdots + r_nx_n + z_1y_1 + \cdots + z_my_m + w_1\tau(y_1) + \cdots + w_m\tau(y_m)$$

mit eindeutigen Koeffizienten $x_1,\ldots,x_n,z_1,\ldots,z_m,w_1,\ldots,w_m \in \mathbb{C}$, weil ja $x_1,\ldots,x_n,y_1,\ldots,y_m,\tau(y_1),\ldots,\tau(y_m)$ eine Basis des komplexen Vektorraums E ist. Wir schreiben $z_j = \alpha_j/2 - i\beta_j/2$ mit $\alpha_j,\beta_j \in \mathbb{R}$ für $j \in \{1,\ldots,m\}$. Da $x \in V$, ist

$$x = \tau(x) = \overline{r_1}x_1 + \cdots + \overline{r_n}x_n + \overline{z_1}\tau(y_1) + \cdots + \overline{z_m}\tau(y_m) + \overline{w_1}y_1 + \cdots + \overline{w_m}y_m;$$

wegen der Eindeutigkeit der Koeffizienten folgt $r_j = \overline{r_j}$ und somit $r_j \in \mathbb{R}$ für alle $j \in \{1,\ldots,m\}$; weiter folgt

$$w_j = \overline{z_j} \quad \text{für alle } j \in \{1,\ldots,m\}.$$

Dann ist

$$\begin{aligned} z_j y_j + \overline{z_j}\tau(y_j) &= z_j y_j + \tau(z_j y_j) = 2\,\mathrm{Re}(z_j y_j) = 2\,\mathrm{Re}(z_j)\,\mathrm{Re}(y_j) - 2\,\mathrm{Im}(z_j)\,\mathrm{Im}(y_j) \\ &= \alpha_j a_j + \beta_j b_j. \end{aligned}$$

Also ist

$$x = r_1 x_1 + \cdots + r_n x_n + \alpha_1 a_1 + \cdots + \alpha_m a_m + \beta_1 b_1 + \cdots + \beta_m b_m.$$

Nach dem Vorigen wird der reelle Vektorraum V durch die $n+2m$ Vektoren

$$x_1, \ldots, x_n, a_1, \ldots, a_m, b_1, \ldots, b_m$$

aufgespannt. Da $\dim_{\mathbb{R}}(V) = n+2m$, bilden die genannten Vektoren automatisch eine Basis des reellen Vektorraums V. $\square$

Der folgende Begriff ist aus der linearen Algebra bekannt.

Definition. Ist A ein Endomorphismus eines komplexen Vektorraums E und $\lambda \in \mathbb{C}$, so ist

$$\ker(A - \lambda\,\mathrm{id}_E) \subseteq \ker(A - \lambda\,\mathrm{id}_E)^2 \subseteq \cdots$$

eine aufsteigende Folge von Untervektorräumen von E, also

$$E_\lambda := \bigcup_{j \in \mathbb{N}} \ker(A - \lambda\,\mathrm{id}_E)^j$$

ein Untervektorraum von E; man nennt E_λ den *Hauptraum* (oder auch *verallgemenerten Eigenraum*) von A zu $\lambda \in \mathbb{C}$.

Bemerkung 7.12. Sind A und B Endomorphismen eines Vektorraums E mit $A \circ B = B \circ A$, so ist

$$B(\ker(A)) \subseteq \ker(A).$$

Ist nämlich $x \in \ker(A)$, so ist $ABx = BAx = B0 = 0$ und somit $Bx \in \ker(A)$.

Aus der linearen Algebra ist weiter bekannt:

Lemma 7.13 *Ist A ein Endomorphismus eines endlich-dimensionalen komplexen Vektorraums E und $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$ die Menge der (paarweise verschiedenen) komplexen Eigenwerte $\lambda_1, \dots, \lambda_m$ von A, so ist*

$$E = E_{\lambda_1} \oplus \cdots \oplus E_{\lambda_m}. \qquad \square$$

Lemma 7.14. *Es seien E ein komplexer Vektorraum, $F \subseteq E$ ein Untervektorraum, $v_1, \dots, v_m \in F$ linear unabhängige Vektoren und $v_{m+1} \in E$ mit $v_{m+1} \notin F$. Dann sind auch $v_1, \dots, v_{m+1}$ linear unabhängig.*

Beweis. Seien $\lambda_1, \dots, \lambda_{m+1} \in \mathbb{C}$ mit $\sum_{k=1}^{m+1} \lambda_k v_k = 0$. Wäre $\lambda_{m+1} \neq 0$, so wäre

$$v_{m+1} = -\sum_{k=1}^{m} \frac{\lambda_k}{\lambda_{m+1}} v_k \in F,$$

im Widerspruch zur Voraussetzung. Also ist $\lambda_{m+1} = 0$ und somit $\sum_{k=1}^{m} \lambda_k v_k = 0$, woraus $\lambda_1 = \cdots = \lambda_m = 0$ folgt. $\square$

Bemerkung. Ist E ein komplexer Vektorraum und $T\colon E \to E$ ein Endomorphismus, so können wir einem durch

$$p(z) = \sum_{j=0}^{n} b_j z^j$$

mit $n \in \mathbb{N}_0$ und $b_0, b_1, \dots, b_n \in \mathbb{C}$ gegebenen Polynom den Endomorphismus

$$p[T] := \sum_{j=0}^{n} b_j T^j$$

zuordnen. Aus der linearen Algebra ist bekannt, dass die so erhaltene Abbildung

$$\mathbb{C}[X] \to \mathrm{End}_{\mathbb{C}}(E), \quad p \mapsto p[T]$$

komplex linear und ein Ringhomomorphismus ist (der sogenannte *Auswertungshomomorphismus*).

Lemma 7.15. (Haupträume des Ableitungsoperators). *Wir betrachten den Endomorphismus*

$$D\colon C^\infty(\mathbb{R}, \mathbb{C}) \to C^\infty(\mathbb{R}, \mathbb{C}), \quad f \mapsto f'$$

des $\mathbb{C}$-Vektorraums $E := C^\infty(\mathbb{R}, \mathbb{C})$. Sei $\lambda \in \mathbb{C}$. Dann ist für jedes $m \in \mathbb{N}$

$$\ker(D - \lambda \operatorname{id}_E)^m$$

ein m-dimensionaler komplexer Untervektorraum von E und die Funktionen

$$\mathbb{R} \to \mathbb{C}, \quad t \mapsto t^a e^{\lambda t}$$

mit $a \in \{0, 1, \ldots, m-1\}$ bilden eine Basis von $\ker(D - \lambda \operatorname{id})^m$.

Beweis. Da

$$(D - \lambda \operatorname{id}_E)^m = \sum_{k=0}^{m} \binom{m}{k} \lambda^{m-k} D^k,$$

ist für jedes $m \in \mathbb{N}$ der Kern $\ker(D - \lambda \operatorname{id}_E)^m$ die Menge aller auf ganz $\mathbb{R}$ definierten komplexwertigen Lösungen γ der homogenen linearen Differentialgleichung

$$\sum_{k=0}^{m} \binom{m}{k} \lambda^{m-k} y^{(k)} = 0$$

$m$ter Ordnung und somit ein m-dimensionaler komplexer Untervektorraum von E (nach Bemerkung 7.5(c)).[8]

Für $a \in \mathbb{N}_0$ sei nun $f_a \colon \mathbb{R} \to \mathbb{C}$ die durch $f_a(t) = t^a e^{\lambda t}$ gegebene Funktion. Wir zeigen per Induktion nach $m \in \mathbb{N}$, dass $f_0, f_1, \ldots, f_{m-1}$ eine Basis von

$$\ker(D - \lambda \operatorname{id}_E)^m$$

ist und $f_{m-1} \notin \ker(D - \lambda \operatorname{id}_E)^{m-1}$.

Induktionsanfang $m = 1$: Da $\frac{d}{dt} e^{\lambda t} = \lambda e^{\lambda t}$, ist $D f_0 = \lambda f_0$ und somit $f_0 \in \ker(D - \lambda \operatorname{id}_E)$ enthalten. Da f_0 nicht die Nullfunktion ist und $\ker(D - \lambda \operatorname{id}_E)$ nach dem Vorigen ein $\mathbb{C}$-Vektorraum der Dimension 1 ist, bildet die Funktion eine Basis. Da $(D - \lambda \operatorname{id}_E)^0 = \operatorname{id}_E$, ist $f_0 \notin \ker(D - \lambda \operatorname{id}_E)^0 = \{0\}$ enthalten.

Induktionsschritt: Gegeben $m \in \mathbb{N}$ sei $f_0, \ldots, f_{m-1}$ eine Basis für $\ker(D - \lambda \operatorname{id}_E)^m$ und $f_{m-1} \notin \ker(D - \lambda \operatorname{id}_E)^{m-1}$. Nach der Produktregel ist

$$\frac{d}{dt}(t^m e^{\lambda t}) = m t^{m-1} e^{\lambda t} + \lambda t^m e^{\lambda t},$$

[8] Aus $\gamma^{(m)} = -\sum_{k=0}^{m-1} \binom{m}{k} \lambda^{m-k} \gamma^{(k)}$ folgt per Induktion, dass $\gamma^{(m)}$ eine C^j-Funktion (und somit γ eine C^{m+j}-Funktion) ist für jedes $j \in \mathbb{N}$; also jede Lösung γ automatisch eine C^∞-Funktion.

also $(D-\lambda\,\mathrm{id}_E)(f_m) = mf_{m-1}$. Folglich ist

$$(D-\lambda\,\mathrm{id}_E)^{m+1}(f_m) = (D-\lambda\,\mathrm{id}_E)^m(D-\lambda\,\mathrm{id}_E)(f_m) = m(D-\lambda\,\mathrm{id}_E)^m(f_{m-1}) = 0,$$

also $f_m \in \ker(D-\lambda\,\mathrm{id}_E)^{m+1}$. Zudem ist

$$(D-\lambda)^m(f_m) = (D-\lambda)^{m-1}(D-\lambda\,\mathrm{id}_E)(f_m) = m(D-\lambda\,\mathrm{id}_E)^{m-1}(f_{m-1}) \neq 0,$$

da $f_{m-1} \notin \ker(D-\lambda\,\mathrm{id}_E)^{m-1}$. Also ist $f_m \notin \ker(D-\lambda\,\mathrm{id}_E)^m$. Wenden wir Lemma 7.14 auf $f_0 \ldots, f_m \in \ker(D-\lambda\,\mathrm{id}_E)^m$ und $f_{m+1} \in \ker(D-\lambda\,\mathrm{id}_E)^{m+1}$ an, so finden wir, dass die Vektoren $f_0, \ldots, f_m$ in $\ker(D-\lambda\,\mathrm{id}_E)^m$ linear unabhängig sind und somit eine Basis des $(m+1)$-dimensionalen Vektorraums $\ker(D-\lambda\,\mathrm{id}_E)^m$ bilden. $\square$

Bemerkung. Insbesondere bilden nach Lemma 7.15 die Funktionen $t \mapsto t^a e^{\lambda t}$ mit $a \in \mathbb{N}_0$ eine Basis des Hauptraums E_λ des Ableitungsoperators D.

Beweis von Satz 7.7. Sei p das durch

$$p(z) = z^m + \sum_{j=0}^{m-1} a_j z^j$$

gegebene charakteristische Polynom der vorgelegten homogenen linearen Differentialgleichung $m$ter Ordnung. Sei $E := C^\infty(\mathbb{R}, \mathbb{C})$. Mit dem Ableitungsoperator $D\colon C^\infty(\mathbb{R}, \mathbb{C}) \to C^\infty(\mathbb{R}, \mathbb{C})$ ist dann

$$p[D] = D^m + a_{m-1}D^{m-1} + \cdots + a_1 D + a_0\,\mathrm{id}_E$$

in $\mathrm{End}_{\mathbb{C}}(E)$, also $H := \ker(p[D]) \subseteq C^\infty(\mathbb{R}, \mathbb{C})$ der (oben $L_h^{\mathbb{C}}$ genannte) Vektorraum aller auf ganz $\mathbb{R}$ definierten komplexwertigen Lösungen der vorgelegten Differentialgleichung. Es sei σ die Menge aller komplexen Nullstellen von p. Gegeben $\lambda \in \sigma$ sei $n(\lambda) \in \mathbb{N}$ die Vielfachheit der Nullstelle λ von p. Dann ist

$$p(z) = q_\lambda(z)(z-\lambda)^{m(\lambda)}$$

mit einem Polynom q_λ und somit

$$p[D] = q_\lambda[D] \circ (D-\lambda\,\mathrm{id}_E)^{n(\lambda)},$$

also

$$H(\lambda) := \ker(D-\mathrm{id}_E)^{n(\lambda)} \subseteq \ker p[D] = H.$$

Da $D \circ p[D] = p[D] \circ D$, ist $D(H) \subseteq H$ (siehe Bemerkung 7.12). Also lässt sich D zu einem Endomorphismus von H einschränken. Da $H(\lambda)$ Teilmenge des Hauptraums H_λ von $D|_H^H$ ist, ist die Summe

$$\bigoplus_{\lambda \in \sigma} H(\lambda) \subseteq H$$

nach Lemma 7.13 direkt. Nun ist $\dim_{\mathbb{C}}(H(\lambda)) = n(\lambda)$ nach Lemma 7.15, also

$$\dim_{\mathbb{C}}\left(\bigoplus_{\lambda \in \sigma} H(\lambda)\right) = \sum_{\lambda \in \sigma} n(\lambda) = m = \dim_{\mathbb{C}}(H),$$

woraus $H = \bigoplus_{\lambda \in \sigma} H(\lambda)$ (und $H_\lambda = H(\lambda)$) folgt. Nach Lemma 7.15 bilden die durch

$$t \mapsto t^a e^{\lambda t}$$

mit $a \in \{0, 1, \ldots, n(\lambda) - 1\}$ gegebenen Funktionen eine $\mathbb{C}$-Basis für $H(\lambda)$. Sei $\tau\colon C^\infty(\mathbb{R}, \mathbb{C}) \to C^\infty(\mathbb{R}, \mathbb{C})$ die komplexe Konjugation von Funktionen. Da $\overline{f'(t)} = (\overline{f})'(t)$, ist

$$D \circ \tau = \tau \circ D$$

und somit

$$\tau(H) \subseteq H$$

nach Bemerkung 7.12. Also schränkt sich τ ein zu einem Endomorphismus $\tau|_H^H$ von H, betrachtet als reeller Vektorraum (und dies ist eine antilineare Involution). Es gibt $k, \ell \in \mathbb{N}_0$, reeller Zahlen $\lambda_1, \ldots, \lambda_k$ und nicht-reelle komplexe Zahlen $\mu_1 = \alpha_1 + i\beta_1, \ldots, \mu_\ell = \alpha_\ell + i\beta_\ell$ derart, dass

$$\lambda_1, \ldots, \lambda_k, \mu_1, \ldots, \mu_\ell, \overline{\mu_1}, \ldots, \overline{\mu_\ell}$$

die paarweise verschiedenen Nullstellen von p sind. Für $j \in \{1, \ldots, k\}$ sei n_j die Vielfachheit der Nullstelle λ_j; für $j \in \{1, \ldots, \ell\}$ sei m_j die Vielfachheit der Nullstelle μ_j (und somit auch der Nullstelle $\overline{\mu_j}$). Nach dem obigen bilden die reellwertigen Funktionen

$$t \mapsto t^a e^{\lambda_j t} \tag{45}$$

mit $j \in \{1, \ldots, k\}$ und $a \in \{0, 1, \ldots, n_j - 1\}$ sowie die komplexwertigen Funktionen

$$t \mapsto t^a e^{\alpha_j t} e^{i\beta_j t} \tag{46}$$

und die komplex konjugierten Funktionen

$$t \mapsto t^a e^{\alpha_j t} e^{-i\beta t}$$

für $j \in \{1, \dots, \ell\}$ und $a \in \{0, \dots, m_j\}$ eine $\mathbb{C}$-Basis für H. Nach Lemma 7.11 bilden die Funktionen aus (45) sowie die Real- und Imaginärteile der Funktionen aus (46) eine $\mathbb{R}$-Basis des reell m-dimensionalen Vektorraums

$$\{f \in H \colon f = \overline{f}\} = L_h.$$

Dies war zu zeigen $\square$