Inhaltsverzeichnis

Gewöhnliche Differentialgleichungen

Ruben Jakob
Universität Tübingen, Fachbereich Mathematik

11. September 2014

1 Definition und Einführung

Wir bezeichnen in der Vorlesung mit Ω eine offene Teilmenge von $\mathbb{R} \times \mathbb{R}^n$, für ein beliebiges $n \geq 1$, und mit $f : \Omega \longrightarrow \mathbb{R}^n$ eine mindestens stetige Funktion.

Definition 1.0.1. *Eine gewöhnliche Differentialgleichung oder ein System gewöhnlicher Differentialgleichungen erster Ordnung ist eine Gleichung der Form*

$$y' = f(\,\cdot\,, y) \tag{1.1}$$

auf einer offenen Teilmenge $\Omega \subset \mathbb{R} \times \mathbb{R}^n$, und für ein stetiges Vektorfeld $f : \Omega \longrightarrow \mathbb{R}^n$. Eine Lösung von (1.1) ist eine differenzierbare Funktion $y : I \longrightarrow \mathbb{R}^n$ auf einem nicht-degenerierten Intervall $I \subseteq \mathbb{R}$ mit

$$y'(t) = f(t, y(t)) \qquad \forall\, t \in I. \tag{1.2}$$

Insbesondere muss hierbei also $(t, y(t)) \in \Omega$ für jedes $t \in I$ gelten. Für ein vorgegebenes Paar $(t_0, y_0) \in \Omega$ heisst y eine Lösung des Anfangswertproblems (AWP)

$$y' = f(\,\cdot\,, y), \qquad y(t_0) = y_0, \tag{1.3}$$

wenn y eine Lösung von (1.1) mit $y(t_0) = y_0$ und $t_0 \in I$ ist. Eine Beziehung der Form

$$y^{(k)} = f(\,\cdot\,, y, y', y'', \ldots, y^{(k-1)}) \tag{1.4}$$

auf einem offenen Gebiet $\Omega \subset \mathbb{R} \times \mathbb{R}^{kn}$ und für ein stetiges Vektorfeld $f : \Omega \longrightarrow \mathbb{R}^n$, zwischen der k-ten Ableitung $y^{(k)}$ und niedrigeren Ableitungen $y^{(l)}$ einer k-fach differenzierbaren Funktion $y : I \longrightarrow \mathbb{R}^n$ nennen wir eine gewöhnliche Differentialgleichung k-ter Ordnung. Diese kann durch Vorgaben

$$y^{(l)}(t_0) = y_0^l, \qquad l = 0, \ldots, k-1$$

für ein Tupel $(t_0, y_0, y_0^1, y_0^2, \ldots, y_0^{k-1}) \in \Omega$ zur Formulierung eines AWP's k-ter Ordnung verwandt werden.

Desweiteren werden wir Differentialgleichungen betrachten, die von einem Parameter $\lambda \in \mathbb{R}^d$ abhängen, d.h. Gleichungen der Form

$$y' = f(\,\cdot\,, y, \lambda). \tag{1.5}$$

Diese kann mittels des Kunstgriffs $z := (y, \lambda)$ zum System von Gleichungen

$$(y, \lambda)' = (f(\,\cdot\,, y, \lambda), 0)$$

erster Ordnung für die Funktion $z := (y, \lambda)$ äquivalent umformuliert werden.

Beispiele

1) Bereits in der experimentellen und in der klassischen, theoretischen Physik treten skalare Differentialgleichungen 2. Ordnung häufig auf. Beispielsweise erfüllt die Spannung U eines elektrischen Schwingkreises eine Gleichung der Form

$$\mu U'' + \rho U' + C^{-1}U = -C^{-1}f(t)$$

wobei μ die Selbstinduktion der Spule des Schwingkreises, ρ einen eingebauten Widerstand, C die Kapazität des Kondensators des Schwingkreises und $f(t)$ eine von einem Generator periodisch hinzugefügte Spannung bezeichnen mögen.

2) Die Bewegung einer gespannten Feder unter Beachtung einer zur Geschwindigkeit der Feder proportionalen Reibungskraft wird durch die Gleichung

$$mx'' + rx' + kx = 0$$

näherungsweise beschrieben, wobei m die Masse der Feder und $k, r > 0$ vom Federmaterial abhängige Konstanten seien. Bei Vernachlässigung der Reibungskraft, also bei $r = 0$, besagt die obige Gleichung gerade, dass die Kraft mx'' der Feder proportional zur Auslenkung x der Feder anwächst, was man experimentell in der Tat leicht nachweisen kann. Die Lösungen dieser Gleichung sind in diesem Falle, also bei $r = 0$, durch $x(t) = c\,\sin(\omega t + p)$, mit $\omega := \sqrt{k/m}$, für beliebige $c, p \in \mathbb{R}$, gegeben.

3) Newton erkannte bereits, dass ein Punkt mit grosser Masse M, der unbeweglich im Ursprung des $\mathbb{R}^3$ ruhe, auf einen frei beweglichen Punkt $x(t)$ mit relativ kleiner Masse m zur Zeit t eine Anziehungskraft $K(t)$ ausübt, die proportional zu m und M ist, in entgegengesetzter Richtung des Ortsvektors $x(t)$ weist (ihn also in der Tat ,,anzieht") und quadratisch mit der Entfernung $\mid x(t) \mid$ von $x(t)$ abnimmt. Da er auch wusste, dass eine Kraft, die ein Punkt mit Masse m erfährt, gerade das Produkt seiner Masse m mit seiner Beschleunigung $x''(t)$ ist, führte ihn dies zur ,,Gravitations-Differentialgleichung"

$$mx''(t) = -\gamma m M \, \frac{x(t)}{\mid x(t) \mid^3}, \tag{1.6}$$

für eine Konstante $\gamma > 0$. Da man sich mittels der Erhaltung des Drehimpulses $D(t) = mx(t) \times x'(t)$ davon überzeugen kann, dass jede Lösung von (1.6) in einer Ebene verlaufen muss, kann man Polarkoordinaten $x(t) = r(t)\exp(i\phi(t))$ einführen und (1.6) äquivalent in die praktischeren Differentialgleichungen

$$r'' - r(\phi')^2 + \frac{\gamma M}{r^2} = 0, \qquad 2r'\phi' + r\phi'' = 0$$

für die Entfernung $r(t) = \mid x(t) \mid$ und den Winkel $\phi(t)$ von $x(t)$ umformulieren.

4) Die nicht-lineare Differentialgleichung erster Ordnung

$$x'(t) = a(t)x(t) + b(t)x^p(t)$$

$p \in \mathbb{R}$, $p > 1$, taucht in der Biologie bei der Modellierung von Wachstumsprozessen auf und kann durch einen Kunstgriff von Bernoulli äquivalent zur Gleichung

$$y'(t) = (1-p)a(t)y(t) + (1-p)b(t)$$

für die Funktion $y(t) := x(t)^{1-p}$ umformuliert werden. Speziell der einfache Fall $p = 2$, konkret ein AWP der Form

$$x'(t) = k\,x(t) - \beta x^2(t), \qquad x(0) = x_0 > 0, \tag{1.7}$$

für Konstanten $k, \beta > 0$, ist ein sehr naheliegendes, wenn auch simples Modell für die Entwicklung der Anzahl $x(t)$ einer Population, die eine in x lineare „Änderungsrate" $\frac{x'}{x}$ hat und abnimmt, sobald die Population x einen gewissen kritischen Schwellenwert $\xi > 0$ übersteigt. Aus diesen beiden Forderungen ergibt sich nämlich sofort die Differentialgleichung

$$\frac{x'}{x} = \beta(\xi - x) = k\, - \beta x,$$

wenn wir $k := \beta\xi$ setzen, und somit das AWP (1.7). Der Bernoulli-Trick gibt uns nun die Möglichkeit, die Lösung des AWP's (1.7) durch Lösung der Gleichung $y'(t) = -k\,y(t) + \beta$ für $y(t) := 1/x(t)$ sofort als

$$x(t) = \frac{x_0 \xi}{x_0 + (\xi - x_0)e^{-kt}}$$

exakt zu gewinnen, was man durch direktes Nachrechnen bei Beachtung von $k = \beta\xi$ leicht verifizieren kann.

5) Das lineare AWP erster Ordnung

$$y' = ay + b, \qquad y(t_0) = y_0,$$

für auf einem Intervall $I = [t_0, T]$ stetige Funktionen a, b besitzt eine eindeutige Lösung y auf ganz $[t_0, T]$, die mittels der Formel

$$y(t) = y_0 \exp\big(\int_{t_0}^t a(s)\,ds\big) + \int_{t_0}^t b(s)\, \exp\big(\int_s^t a(\tau)\,d\tau\big) ds \tag{1.8}$$

exakt angegeben bzw. berechnet werden kann. Zum Beweis leiten wir die Funktion $\phi(t) := y(t)\exp\big(-\int_{t_0}^t a(s)\,ds\big)$ ab und erhalten bei Verwendung der Produkt- und Kettenregel und des Hauptsatzes der Differential- und Integralrechnung, und bei Beachtung von $y' = ay + b$:

$$\phi'(t) = (y'(t) - a(t)y(t))\exp\big(-\int_{t_0}^t a(s)\,ds\big) = b(t)\exp\big(-\int_{t_0}^t a(s)\,ds\big).$$

Beachten wir nun noch $\phi(t_0) = y(t_0) = y_0$, so erhalten wir per Integration dieser Gleichung über $[t_0, t]$ wieder anhand des Hauptsatzes der Differential- und Integralrechnung:

$$y(t) \exp\big(-\int_{t_0}^{t} a(s)\,ds\big) - y_0 = \int_{t_0}^{t} b(s) \exp\big(-\int_{t_0}^{s} a(\tau)\,d\tau\big)\,ds.$$

Addition dieser Gleichung mit y_0 und anschliessende Multiplikation mit $\exp\big(\int_{t_0}^{t} a(s)\,ds\big)$ liefert die behauptete Formel (1.8). Definieren wir den ,,Evolutionsoperator" $U(\,\cdot\,,\,\cdot\,) : [t_0, T] \times [t_0, T] \to \mathbb{R}^*$ durch

$$U(t, s) := \exp\big(\int_{s}^{t} a(\tau)\,d\tau\big),$$

so lässt sich die Lösungsformel (1.8) auch in der Form

$$y(t) = U(t, t_0)\,y_0 + \int_{t_0}^{t} U(t, s)\,b(s)\,ds$$

angeben, der die Eigenschaften $U(s, s) = id_{\mathbb{R}}$ und $U(t, \tau)\,U(\tau, s) = U(t, s), \forall\, s, t, \tau \in [t_0, T]$, insbesondere also $U(t, s) = U(s, t)^{-1}$, hat und zur Vereinfachung der Notation für die explizite Lösungsformel für die Eigenfunktionen des zeitabhängigen Schrödingeroperators aus der Quantenmechanik benutzt wird.

6) Die nicht-lineare Differentialgleichung erster Ordnung

$$y' = 1 + y^2 \tag{1.9}$$

hat die Lösungen $y(t) = \tan(t + \alpha)$, für $\mid t + \alpha \mid < \pi/2$, wie man anhand von

$$(\arctan(y))' = \frac{y'}{1 + y^2} = 1$$

und somit $\arctan(y(t)) = t + \alpha$, für ein festes $\alpha \in \mathbb{R}$, verifizieren kann. Wir bemerken hierbei, dass die Lösungsfunktion y wegen $\tan(t) \to \pm\infty$ für $t \to \pm\frac{\pi}{2}$ nicht auf ganz $\mathbb{R}$ sondern nur auf beschränkten Intervallen existieren kann ! Wie wir später einsehen werden, ist das quadratische – also nicht mehr lineare – Wachstum von $f(y) := 1 + y^2$ für dieses ,,explosive Verhalten" der Lösungsfunktionen verantwortlich.

7) Das AWP

$$x' = \operatorname{sign}(x)\sqrt{\mid x \mid} =: f(x), \qquad x(t_0) = x_0 \tag{1.10}$$

besitzt im Fall $\mid x_0 \mid > 0$ die eindeutige Lösung

$$x(t) := \begin{cases} \operatorname{sign}(x_0)\big(\sqrt{\mid x_0 \mid} + \frac{t - t_0}{2}\big)^2 & : \text{ für } t > t_0 - 2\sqrt{\mid x_0 \mid} \\ 0 & : \text{ für } t \le t_0 - 2\sqrt{\mid x_0 \mid}. \end{cases}$$

Dies ist also eine entweder nach oben oder nach unten geöffnete Parabel, die zum Zeitpunkt $t_0 - 2\sqrt{| x_0 |}$ die „x-Achse" berührt und dann nach links zu Null fortgesetzt wird. Ist jedoch $x_0 = 0$ und $f(x) := \sqrt{| x |}$, so kann diese Lösung beliebig weit nach rechts verschoben werden, ohne die Anfangsbedingung $x(t_0) = x_0 = 0$ zu verletzen, d.h.

$$x_\alpha(t) := \begin{cases} \left(\frac{t-\alpha}{2}\right)^2 & : \text{ für } t \geq \alpha \\ 0 & : \text{ für } t \leq \alpha, \end{cases}$$

für beliebiges $\alpha \geq t_0$, ist eine Lösungsschar des AWP's

$$x' = \sqrt{| x |}, \qquad x(t_0) = 0$$

mit der Mächtigkeit von $\mathbb{R}$. Der tieferliegende Grund für dieses überraschende Phänomen liegt im Verhalten der Funktion f in 0. f ist in 0 zwar stetig, jedoch nicht (lokal-) Lipschitz-stetig. Die im folgenden Kapitel bewiesenen Sätze von Picard-Lindelöf und Gronwall werden uns zeigen, dass es zu einem AWP (1.3) eine eindeutige, und nicht weiter fortsetzbare Lösung gibt, falls f lokal Lipschitz-stetig ist.

2 Kurzzeit-Existenz, Eindeutigkeit und Maximalität von Lösungen

In diesem Kapitel werden wir die beiden klassischen Sätze von Peano und Picard-Lindelöf beweisen, die die Existenz einer Lösung eines Anfangswertproblems (1.3) für stetige (bzw. lokal Lipschitz-stetige) Vektorfelder f garantieren, und anschliessend ein hinreichendes Kriterium für deren Eindeutigkeit herleiten. Wir beginnen mit

Proposition 2.0.1. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen und $(t_0, y_0) \in \Omega$ fest gewählt. Eine differenzierbare Funktion $y : I \longrightarrow \mathbb{R}^n$, mit $t_0 \in I \subset \mathbb{R}$, ist genau dann eine Lösung des AWP's*

$$y' = f(\,\cdot\,, y), \qquad y(t_0) = y_0, \tag{2.1}$$

wenn $(t, y(t)) \in \Omega$ für jedes $t \in I$ und

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds \qquad \forall t \in I \tag{2.2}$$

gilt.

Proof. Gilt (2.2), so folgen aus der vorausgesetzten Differenzierbarkeit von y und der Stetigkeit von f sowohl die stetige Differenzierbarkeit von y als auch $y'(t) = f(t, y(t))$, für jedes $t \in I$, aus dem (zweiten) Hauptsatz der Differential- und Integralrechnung. Ausserdem erhalten wir aus (2.2) sofort $y(t_0) = y_0$ und somit insgesamt (2.1).
Ist umgekehrt y eine Lösung des AWP's (2.1), so folgt aus $y'(t) = f(t, y(t))$ zunächst die stetige Differenzierbarkeit von y und auch

$$y(t) = y(t_0) + \int_{t_0}^{t} y'(s)\, ds = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

für jedes $t \in I$ aus dem (ersten) Hauptsatz der Differential- und Integralrechnung, also (2.2). $\square$

Theorem 2.0.1. *[Satz von Peano] Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig und $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen. Dann hat für jedes Paar $(t_0, y_0) \in \Omega$ das AWP (2.1) eine Lösung $z : I \longrightarrow \mathbb{R}^n$ auf einem offenen Intervall I, welches t_0 enthält.*

Proof. : Wir wählen zu einem fixierten Paar $(t_0, y_0) \in \Omega$ ein $\rho > 0$, welches $\overline{B_{2\rho}(t_0, y_0)} \subset \Omega$ erfüllt und setzen $M := \sup_{B_{2\rho}(t_0,y_0)} |\, f \,|$, welches anhand der Stetigkeit von f endlich ist. Desweiteren wählen wir $\epsilon > 0$ mit $\epsilon(M+1) < \rho$. Insbesondere folgt hieraus: $[t_0-\epsilon, t_0+\epsilon] \times B_\rho(y_0) \subset B_{2\rho}(t_0, y_0) \subset \Omega$. Die grundlegende Idee des Beweises besteht nun darin, eine

Lösung von (2.1) zunächst auf $[t_0, t_0+\epsilon]$ als gleichmässigen Limes einer bestimmten Folge stückweise linearer Funktionen y_m zu erhalten, welche in y_0 zum Zeitpunkt t_0 starten und in einer diskreten Folge von Zeitpunkten $t_j := t_0 + hj$, für $h := \epsilon/m$ und $j = 1, \ldots, m$, in Richtung von f linear fortgesetzt werden. Präzise heisst dies: Wir unterteilen zu jedem $m \in \mathbb{N}$ das Intervall $[t_0, t_0+\epsilon]$ in die m Intervalle $[t_{j-1}, t_j]$, mit $t_j := t_0 + jh$, für $h := \epsilon/m$ und $j = 0, 1, \ldots, m$, setzen $y_m(t_0) := y_0$ und

$$y_m(t) := y_m(t_{j-1}) + (t - t_{j-1})\, f(t_{j-1}, y_m(t_{j-1})), \qquad \text{für } t \in (t_{j-1}, t_j] \tag{2.3}$$

und $j = 1, \ldots, m$. Zum Nachweis der Wohldefiniertheit von y_m müssen wir zeigen, dass $(t_{j-1}, y_m(t_{j-1})) \in \Omega$ für jedes $j = 1, \ldots, m$ gilt, was bereits aus

$$y_m(t_{j-1}) \in B_\rho(y_0) \qquad \text{für } j = 1, \ldots, m \tag{2.4}$$

folgt, was wir nun per Induktion zeigen. Für $j = 1$ gilt $y_m(t_{j-1}) = y_0 \in B_\rho(y_0)$ per Definition der y_m. Nun nehmen wir an, dass (2.4) für $j = 1, \ldots, k$ und ein $k \leq m$ gelte. Hieraus folgt zunächst, dass $(t_{j-1}, y_m(t_{j-1})) \in [t_0, t_0 + \epsilon] \times B_\rho(y_0) \subset B_{2\rho}(t_0, y_0)$ und somit $\mid f(t_{j-1}, y_m(t_{j-1})) \mid \leq M$ gilt, woraus wir

$$\mid y_m(t) - y_m(t_{j-1}) \mid \leq Mh \qquad \text{für } t \in [t_{j-1}, t_j] \tag{2.5}$$

aus der Definition (2.3) der y_m herleiten. Insbesondere erhalten wir hieraus $\mid y_m(t_i) - y_m(t_{i-1}) \mid \leq Mh$, für $i = 1, \ldots, j$, und somit

$$\mid y_m(t_j) - y_0 \mid \leq \sum_{i=1}^{j} \mid y_m(t_i) - y_m(t_{i-1}) \mid \leq Mjh \leq M\epsilon < \rho \tag{2.6}$$

für jedes $j = 1, \ldots, k$ erhalten. Also ist der Induktionsschritt von $j - 1$ auf j geglückt, und wir erhalten (2.4) per Induktion für jedes $j = 1, \ldots, m$. Aus der Definition (2.3) der y_m folgt sofort:

$$y_m'(t) = f(t_{j-1}, y_m(t_{j-1})) \qquad \forall\, t \in (t_{j-1}, t_j), \tag{2.7}$$

insbesondere also $\mid y_m' \mid < M$ auf jedem offenen Teilintervall (t_{j-1}, t_j). Für ein Paar $t < s \in [t_{k-1}, t_k]$, für ein k, erhält man somit sofort aus dem Mittelwertsatz:

$$\mid y_m(t) - y_m(s) \mid \leq M(s - t).$$

Falls t und s in verschiedenen Teil-Intervallen von $[t_0, t_0 + \epsilon]$ liegen, also $t \in [t_{k-1}, t_k]$ und $s \in [t_{l-1}, t_l]$, für $k < l$, so können wir wieder aus dem Mittelwertsatz folgern:

$$\begin{aligned} &\mid y_m(t) - y_m(s) \mid \\ \leq &\mid y_m(t) - y_m(t_k) \mid + \mid y_m(t_k) - y_m(t_{k+1}) \mid + \ldots + \mid y_m(t_{l-2}) - y_m(t_{l-1}) \mid + \mid y_m(t_{l-1}) - y_m(s) \mid \\ &\leq M((t_k - t) + (t_{k+1} - t_k) + \ldots + (t_{l-1} - t_{l-2}) + (s - t_{l-1})) = M(s - t). \end{aligned} \tag{2.8}$$

Dies beweist insbesondere die gleichgradige Stetigkeit der Funktionenfolge $\{y_m\}$ auf $[t_0, t_0 + \epsilon]$. Desweiteren folgt aus (2.5) zusammen mit der Dreiecks-Ungleichung:

$$\mid (t, y_m(t)) - (t_{j-1}, y_m(t_{j-1})) \mid < (M + 1)h, \tag{2.9}$$

für jedes $t \in [t_{j-1}, t_j]$ und jedes $j = 1, \ldots, m$. Zusammen mit (2.6) und $(M+1)\epsilon < \rho$ folgt somit

$$| (t, y_m(t)) - (t_0, y_0) | < (M+1)h + \sqrt{\epsilon^2 + M^2\epsilon^2} < \rho + \sqrt{(M^2+1)\epsilon^2} < 2\rho,$$

also $(t, y_m(t)) \in B_{2\rho}(t_0, y_0)$ für jedes $t \in [t_0, t_0+\epsilon]$ und jedes m. Wenden wir auf (2.9) f an und beachten wir noch (2.7), so liefert die gleichmässige Stetigkeit von f auf $\overline{B_{2\rho}(t_0, y_0)}$, dass es zu jedem $\delta > 0$ ein hinreichend kleines $h_\delta > 0$ bzw. ein hinreichend grosses $m_\delta \in \mathbb{N}$ gibt, sodass

$$| f(t, y_m(t)) - y_m'(t) | < \delta \qquad \forall\, t \in (t_{j-1}, t_j), \tag{2.10}$$

für $h < h_\delta$ bzw. für jedes $m > m_\delta$, und für jedes $j = 1, \ldots, m$, gilt. Da jedes y_m auf jedem abgeschlossenen Intervall $[t_{j-1}, t_j]$ linear, also insbesondere die Ableitung y_m' der Einschränkung (!) von y_m auf jedes abgeschlossene Intervall $[t_{j-1}, t_j]$ konstant und insbesondere stetig auf ganz $[t_{j-1}, t_j]$ ist, dürfen wir auf jedem Intervall $[t_{j-1}, t_j]$ den ersten Hauptsatz der Differential- und Integralrechnung auf y_m anwenden und erhalten durch Summation der resultierenden Gleichungen:

$$y_m(t) - y_0 = y_m(t) - y_m(t_0) = \int_{t_0}^{t} y_m'(s)\, ds$$

für jedes $t \in [t_0, t_0+\epsilon]$. Somit folgt zusammen mit (2.10) für jedes $t \in [t_0, t_0+\epsilon]$:

$$\begin{aligned} | y_m(t) - y_0 - \int_{t_0}^{t} f(s, y_m(s))\, ds | &= | \int_{t_0}^{t} y_m'(s) - f(s, y_m(s))\, ds | \\ &\le \int_{t_0}^{t} | y_m'(s) - f(s, y_m(s)) |\, ds < \delta\epsilon, \end{aligned} \tag{2.11}$$

falls $m > m_\delta$ ist. Wir bewiesen bereits sowohl die Beschränktheit als auch die gleichgradige Stetigkeit der Funktionenfolge $\{y_m\}$ auf $[t_0, t_0+\epsilon]$, sodass uns der Satz von Arzela-Ascoli die Existenz einer Teilfolge $\{y_{m_k}\}$ und einer auf $[t_0, t_0+\epsilon]$ stetigen Funktion y mit $\sup_{[t_0,t_0+\epsilon]} | y_{m_k} - y | \to 0$ für $k \to \infty$ garantiert. Da f auf $\overline{B_{2\rho}(t_0, y_0)}$ gleichmässig stetig ist und $(t, y_{m_k}(t)) \in B_{2\rho}(t_0, y_0)$ für jedes $t \in [t_0, t_0+\epsilon]$ gilt, folgt hieraus zunächst $\sup_{[t_0,t_0+\epsilon]} | f(\,\cdot\,, y_{m_k}) - f(\,\cdot\,, y) | \to 0$ und somit anhand von (2.11) im Limes für $\delta \searrow 0$ und $m_k \to \infty$:

$$y(t) = y_0 + \int_{t_0}^{t} f(s, y(s))\, ds \qquad \forall\, t \in [t_0, t_0+\epsilon].$$

Anhand von Proposition 2.0.1 bedeutet diese Gleichung gerade, dass die Funktion y eine Lösung des AWP's (2.1) auf $[t_0, t_0+\epsilon]$ ist.
Unterteilt man das Intervall $[t_0-\epsilon, t_0]$ ebenfalls in Teil-Intervalle $[t_{j+1}, t_j]$, mit $t_j := t_0 - jh$, für $h := \epsilon/m$ und $j = 0, 1, \ldots, m$, und setzt $\tilde{y}_m(t_0) := y_0$ und

$$\tilde{y}_m(t) := \tilde{y}_m(t_j) + (t - t_j)\, f(t_j, \tilde{y}_m(t_j)), \qquad \text{für } t \in [t_{j+1}, t_j) \tag{2.12}$$

und $j = 0, 1, \ldots, m-1$, so erhält man genau wie oben eine Teilfolge $\{\tilde{y}_{m_k}\}$, die gleichmässig gegen eine Lösung $\tilde{y}$ des AWP's (2.1) auf $[t_0 - \epsilon, t_0]$ konvergiert. Wegen

$$\lim_{t \nearrow t_0} \tilde{y}'(t) = \lim_{t \nearrow t_0} f(t, \tilde{y}(t)) = f(t_0, \tilde{y}(t_0)) = f(t_0, y_0) = \lim_{t \searrow t_0} f(t, y(t)) = \lim_{t \searrow t_0} y'(t)$$

erhält man somit durch

$$z(t) := \begin{cases} y(t) & : \text{ für } t \in [t_0, t_0 + \epsilon] \\ \tilde{y}(t) & : \text{ für } t \in [t_0 - \epsilon, t_0] \end{cases}$$

eine auf ganz $[t_0 - \epsilon, t_0 + \epsilon]$ stetig differenzierbare Lösung des AWP's (2.1), wie behauptet. □

Theorem 2.0.2. *[Lemma von Gronwall, erste Version] Es sei $y : [a, b] \longrightarrow \mathbb{R}^n$ differenzierbar und $\kappa : [a, b] \longrightarrow [0, \infty)$ stetig mit*

$$\langle y'(t), y(t) \rangle \leq \kappa(t) \mid y(t) \mid^2 \qquad \text{für } t \in [a, b],$$

was z.B. im Falle

$$\mid y'(t) \mid \leq \kappa(t) \mid y(t) \mid \qquad \text{für } t \in [a, b]$$

erfüllt ist. Dann folgt:

$$\mid y(t) \mid \leq \mid y(a) \mid \exp\left(\int_a^t \kappa(s)\, ds\right) \qquad \text{für alle } t \in [a, b].$$

Proof. Wir definieren die Funktion

$$\phi(t) := \mid y(t) \mid^2 \exp\left(-2 \int_a^t \kappa(s)\, ds\right)$$

und erhalten aus der Produktregel und dem Hauptsatz der Differential- und Integralrechnung sofort:

$$\phi'(t) = (2 \langle y'(t), y(t) \rangle - 2\kappa(t) \mid y(t) \mid^2) \exp\left(-2 \int_a^t \kappa(s)\, ds\right) \leq 0.$$

Somit ist ϕ monoton nicht-wachsend, insbesondere gilt also $\phi(t) \leq \phi(a)$ für jedes $t \in [a, b]$, woraus sofort die Behauptung folgt. □

Theorem 2.0.3. *[Lemma von Gronwall, zweite Version] Es sei $z : [a, b] \longrightarrow \mathbb{R}$ eine stetige Lösung der Integralungleichung*

$$z(t) \leq K + L \int_a^t z(s)\, ds \qquad \text{für } t \in [a, b] \tag{2.13}$$

für Konstanten $K, L \geq 0$. Dann folgt:

$$z(t) \leq K \exp(L(t - a)) \qquad \text{für alle } t \in [a, b].$$

Proof. Wir wählen ein $\epsilon > 0$ beliebig klein und definieren die Funktion

$$\phi(t) := (K + \epsilon) \exp(L(t - a)) \qquad \text{für alle } t \in [a, b].$$

Dann gilt $\phi' = L\phi$ auf $[a, b]$ und ausserdem $\phi(a) = K + \epsilon$, sodass der Hauptsatz der Differential- und Integralrechnung

$$\phi(t) = K + \epsilon + L \int_a^t \phi(s)\, ds$$

liefert. Nun zeigen wir, dass

$$z(t) < \phi(t) \qquad \text{für alle } t \in [a, b] \tag{2.14}$$

gilt. Angenommen, diese Behauptung wäre falsch, so existierte wegen $z(a) \leq K < K + \epsilon = \phi(a)$ nach (2.13) und wegen der Stetigkeit von z und ϕ ein $t_1 \in (a, b]$ mit

$$z(t) < \phi(t) \qquad \text{für alle } t \in [a, t_1) \qquad \text{jedoch} \qquad z(t_1) = \phi(t_1).$$

Zusammen mit (2.13) folgte hieraus

$$z(t_1) \leq K + L \int_a^{t_1} z(s)\, ds < K + \epsilon + L \int_a^{t_1} \phi(s)\, ds = \phi(t_1),$$

was im Widerspruch zu $z(t_1) = \phi(t_1)$ steht. Somit gilt die Behauptung (2.14), für jedes $\epsilon > 0$, sodass wir in der Tat

$$z(t) \leq K \exp(L(t - a)) \qquad \text{für alle } t \in [a, b]$$

anhand der Definition von ϕ und bei $\epsilon \searrow 0$ erhalten. □

Theorem 2.0.4. *[Eindeutigkeits-Satz] Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, und f erfülle*

$$\langle y_1 - y_2, f(t, y_1) - f(t, y_2) \rangle \leq L \mid y_1 - y_2 \mid^2 \tag{2.15}$$

für ein $L \in [0, \infty)$ und für beliebige Paare $(t, y_1), (t, y_2) \in \Omega$, was insbesondere gilt, falls f Lipschitz-stetig mit Lipschitz-Konstante L in y ist, also falls

$$\mid f(t, y_1) - f(t, y_2) \mid \leq L \mid y_1 - y_2 \mid \tag{2.16}$$

für beliebige Paare $(t, y_1), (t, y_2) \in \Omega$ gilt. Bezeichne z die Lösung des AWP's (2.1) auf einem Intervall $I = [t_0 - \epsilon, t_0 + \epsilon]$ um t_0 aus Theorem 2.0.1, so gilt für jede Lösung y von (2.1) auf einem offenen Intervall $J := (t_0 - \rho, t_0 + \rho) \subset I$: $y \equiv z$ auf $[t_0, t_0 + \rho)$. Unter der stärkeren Lipschitz-Bedingung (2.16) an f folgt $y \equiv z$ auf ganz J.

Proof. Wir definieren zu zwei beliebigen Lösungen y_1, y_2 von (2.1) auf einem offenen Intervall $J := (t_0 - \rho, t_0 + \rho) \subset I$ die Funktion $\triangle := \mid y_1 - y_2 \mid^2$ auf J und erhalten aus der Voraussetzung (2.15) an f:

$$\begin{aligned} \triangle'(t) = 2 \langle y_1(t) - y_2(t), y_1'(t) - y_2'(t) \rangle &= 2 \langle y_1(t) - y_2(t), f(t, y_1) - f(t, y_2) \rangle \\ &\leq 2L \mid y_1(t) - y_2(t) \mid^2 = 2L\triangle(t) \end{aligned}$$

und somit $\triangle'(t)\triangle(t) \leq 2L\triangle^2(t)$, für jedes $t \in J$. Aus Theorem 2.0.2, also aus der ersten Version des Lemmas von Gronwall, folgt somit wegen $t_0 \in J$ und $\triangle(t_0) = 0$:

$$0 \leq \triangle(t) \leq \triangle(t_0) \exp(2L(t-t_0)) \equiv 0 \qquad \text{für alle } t \in [t_0, t_0+\rho),$$

also $\triangle \equiv 0$ und somit $y_1 \equiv y_2 \equiv z$ auf $[t_0, t_0+\rho)$.
Nun erfülle f die stärkere Lipschitz-Bedingung (2.16). Ist y eine beliebige Lösung von (2.1) auf einem Intervall $J := (t_0-\rho, t_0+\rho) \subset I$, so löst die Funktion $\tilde{y}(t) := y(2t_0-t)$ wegen

$$\tilde{y}'(t) = -y'(2t_0-t) = -f(2t_0-t, y(2t_0-t)) = -f(2t_0-t, \tilde{y}(t))$$

und wegen $\tilde{y}(t_0) = y(t_0) = y_0$ das AWP (2.1) zur Funktion $\tilde{f}(t,y) := -f(2t_0-t, y)$, welche ebenfalls (2.16) erfüllt. Sind nun y_1, y_2 zwei beliebige Lösungen von (2.1) auf J, so sind $\tilde{y}_1, \tilde{y}_2$ zwei Lösungen von (2.1) auf J zur Funktion $\tilde{f}$, die (2.16) erfüllt. Somit folgt für $\tilde{y}_1, \tilde{y}_2$ aus dem soeben Bewiesenen: $\tilde{y}_1 \equiv \tilde{y}_2$ auf $[t_0, t_0+\rho)$, d.h. $y_1(2t_0-t) = y_2(2t_0-t)$ für alle $t \in [t_0, t_0+\rho)$, und somit $y_1(t) = y_2(t) = z(t)$ für alle $t \in (t_0-\rho, t_0]$. Insgesamt folgt also $y_1 \equiv y_2 \equiv z$ auf ganz J unter der Bedingung (2.16) an f. □

Bemerkung 2.0.1. *Man bemerke hierbei, dass man für nur stetiges f im Allgemeinen keine Eindeutigkeit der Lösung des AWP's* (2.1) *auf einem beliebig kleinen Intervall um t_0 erwarten darf, wie das siebte Beispiel aus Kapitel 1 bereits zeigte.*

Theorem 2.0.5 (Existenz- und Eindeutigkeitssatz von Picard-Lindelöf)**.** *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, und ausserdem lokal Lipschitz-stetig in y, d.h. zu jedem Punkt $(t_0, y_0) \in \Omega$ existiert eine Umgebung $B_\rho(t_0, y_0) \subset \Omega$ und eine Konstante $L_\rho \in [0, \infty)$, sodass*

$$| f(t,y_1) - f(t,y_2) | \leq L_\rho \mid y_1 - y_2 \mid \tag{2.17}$$

für beliebige Paare $(t, y_1), (t, y_2) \in B_\rho(t_0, y_0)$ gelte. Dann hat das AWP (2.1) *auf einem hinreichend kleinen Intervall $I = [t_0-\epsilon, t_0+\epsilon]$ um t_0 eine eindeutige Lösung z.*

Proof. Wir betrachten einen beliebigen Punkt $(t_0, y_0) \in \Omega$ und wählen $\rho > 0$ hinreichend klein, sodass (2.17) auf $B_\rho(t_0, y_0)$ und $\overline{B_\rho(t_0, y_0)} \subset \Omega$ gilt, und setzen desweiteren

$$M := \sup_{(t,y) \in B_\rho(t_0,y_0)} | f(t,y) | < \infty.$$

Anschliessend wählen wir ein hinreichend kleines $\epsilon > 0$, welches

$$[t_0-\epsilon, t_0+\epsilon] \times \overline{B_{M\epsilon}(y_0)} \subset B_\rho(t_0, y_0)$$

und $L_\rho \epsilon < 1$ erfüllt. Nun betrachten wir die Funktionenmenge

$$\mathcal{F} := \{ y : [t_0-\epsilon, t_0+\epsilon] \longrightarrow \overline{B_{M\epsilon}(y_0)} \mid y \text{ ist stetig und } y(t_0) = y_0 \},$$

statten diese mit der Supremumsnorm $\| y \| := \sup_{t \in [t_0-\epsilon, t_0+\epsilon]} \mid y(t) \mid$ aus und werden nun zeigen, dass durch

$$\Phi(y)(t) := y_0 + \int_{t_0}^{t} f(s, y(s))\, ds$$

eine ,,Kontraktion" $\Phi : \mathcal{F} \longrightarrow \mathcal{F}$, d.h. eine stetige Selbstabbildung von $\mathcal{F}$ mit

$$\| \Phi(y_1) - \Phi(y_2) \| \leq q \ \| y_1 - y_2 \| \qquad \text{für alle} \qquad y_1, y_2 \in \mathcal{F} \tag{2.18}$$

und für ein $q \in [0,1)$ geliefert wird. Zunächst folgt für jedes $y \in \mathcal{F}$ per Definition von Φ sofort die Stetigkeit der Funktion $\Phi(y)(t)$ und ebenso $\Phi(y)(t_0) = y_0$ und

$$| \Phi(y)(t) - y_0 | = | \int_{t_0}^{t} f(s, y(s))\, ds | \leq M \mid t - t_0 \mid \leq M\epsilon \qquad \forall\, t \in [t_0 - \epsilon, t_0 + \epsilon],$$

sodass wir also bereits $\Phi(y) \in \mathcal{F}$ für jedes $y \in \mathcal{F}$ nachgewiesen haben. Nun rechnen wir weiter mittels der lokalen Lipschitz-Bedingung (2.17):

$$\begin{aligned} | \Phi(y_1)(t) - \Phi(y_2)(t) | &= | \int_{t_0}^{t} f(s, y_1(s)) - f(s, y_2(s))\, ds | \\ &\leq | t - t_0 | \sup_{s \in [t_0 - \epsilon, t_0 + \epsilon]} | f(s, y_1(s)) - f(s, y_2(s)) | \\ &\leq L_\rho\, \epsilon \sup_{s \in [t_0 - \epsilon, t_0 + \epsilon]} | y_1(s) - y_2(s) | = L_\rho\, \epsilon \ \| y_1 - y_2 \| \ . \end{aligned}$$

Wenn wir hierin auf der linken Seite das Supremum über alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$ bilden, so erhalten wir mit $q := L_\rho \epsilon < 1$ insgesamt also die Behauptung in (2.18). Eine Funktion $z \in \mathcal{F}$ ist nach Proposition 2.0.1 genau dann eine Lösung des AWP's (2.1) auf $[t_0 - \epsilon, t_0 + \epsilon]$, wenn sie

$$z(t) = y_0 + \int_{t_0}^{t} f(s, z(s))\, ds = \Phi(z)(t) \qquad \forall\, t \in [t_0 - \epsilon, t_0 + \epsilon]$$

erfüllt, d.h. genau dann wenn z ein Fixpunkt von Φ ist. Tatsächlich liefert die Kontraktionseigenschaft (2.18) von Φ die Existenz solch eines Fixpunktes in $\mathcal{F}$, wie wir nun direkt nachweisen werden. Dazu konstruieren wir rekursiv eine Folge von Funktionen $\{y_k\} \subset \mathcal{F}$ durch $y_0(t) \equiv y_0$, $y_1 := \Phi(y_0), \ldots, y_{k+1} := \Phi(y_k), \ldots$ und rechnen mittels k-facher Anwendung von (2.18) auf Φ:

$$\| y_{k+1} - y_k \| = \| \Phi(y_k) - \Phi(y_{k-1}) \| \leq q \ \| y_k - y_{k-1} \| \leq q^k \ \| y_1 - y_0 \|$$

und somit für beliebiges $l \geq k$:

$$\| y_l - y_k \| \leq \sum_{i=k+1}^{l} \| y_i - y_{i-1} \| \leq \sum_{i=k+1}^{\infty} q^{i-1} \ \| y_1 - y_0 \| = \frac{q^k}{1-q} \ \| y_1 - y_0 \| \tag{2.19}$$

da $q < 1$ ist. Ebenso folgt wegen $q < 1$ aus dieser Abschätzung, dass $\{y_k\}$ eine Cauchy-Folge in $\mathcal{F}$ bezüglich $\| \cdot \|$ ist, und wir werden sofort sehen, dass diese in der Tat bzgl. $\| \cdot \|$ gegen ein $z \in \mathcal{F}$ konvergiert. Zunächst folgt aus (2.19) anhand der Vollständigkeit des $\mathbb{R}^n$ die Existenz eines punktweisen Limes

$$z(t) := \lim_{k \to \infty} y_k(t) \in \overline{B_{M\epsilon}(y_0)}$$

in jedem $t \in [t_0 - \epsilon, t_0 + \epsilon]$. Desweiteren sehen wir:

$$| z(t) - y_k(t) |= \lim_{l\to\infty} | y_l(t) - y_k(t) |\leq \limsup_{l\to\infty} \| y_l - y_k \|\leq \frac{q^k}{1-q} \| y_1 - y_0 \|,$$

für jedes $t \in [t_0 - \epsilon, t_0 + \epsilon]$. Also konvergiert die Folge $\{y_k\}$ bzgl. $\| \cdot \|$, d.h. gleichmässig gegen z, woraus die Stetigkeit von z auf $[t_0 - \epsilon, t_0 + \epsilon]$ und damit auch $z \in \mathcal{F}$ folgt. Aus (2.18) ergibt sich somit auch

$$\| \Phi(z) - \Phi(y_k) \|\leq q \ \| z - y_k \|\to 0,$$

also die gleichmässige Konvergenz der Funktionenfolge $\{\Phi(y_k)\}$ gegen die Funktion $\Phi(z)$. Insgesamt erhalten wir also:

$$\Phi(z) \longleftarrow \Phi(y_k) = y_{k+1} \longrightarrow z \qquad \text{bzgl. } \ \| \cdot \|$$

und somit wie gewünscht $z = \Phi(z)$.
Nun zeigen wir noch die Eindeutigkeit dieser Lösung z auf $[t_0 - \epsilon, t_0 + \epsilon]$. Zunächst beweisen wir hierfür, dass jede Lösung des AWP's (2.1) auf $[t_0 - \epsilon, t_0 + \epsilon]$ in der Funktionenmenge $\mathcal{F}$ liegen muss. Angenommen, dies gälte nicht, dann existierte eine Lösung y^* auf $[t_0 - \epsilon, t_0 + \epsilon]$, für welche

$$0 < T := \inf\{| t - t_0 | \mid t \in [t_0 - \epsilon, t_0 + \epsilon] \quad \text{und} \quad y^*(t) \notin \overline{B_{M\epsilon}(y_0)}\} < \epsilon$$

gälte. Anhand der Stetigkeit von y^* folgte hieraus entweder

$$y^*(t) \in \overline{B_{M\epsilon}(y_0)} \qquad \text{für alle } t \in [t_0, t_0 + T] \quad \text{und} \quad y^*(t_0 + T) \in \partial B_{M\epsilon}(y_0) \tag{2.20}$$

oder

$$y^*(t) \in \overline{B_{M\epsilon}(y_0)} \qquad \text{für alle } t \in [t_0 - T, t_0] \quad \text{und} \quad y^*(t_0 - T) \in \partial B_{M\epsilon}(y_0). \tag{2.21}$$

Wir nehmen der Einfachheit halber den ersten Fall an und nennen $t^* := t_0 + T$. Desweiteren setzten wir zu Beginn des Beweises $M := \sup_{(t,y)\in B_\rho(t_0,y_0)} | f(t, y) |$ und wählten $\epsilon > 0$ derart klein, dass $[t_0 - \epsilon, t_0 + \epsilon] \times \overline{B_{M\epsilon}(y_0)} \subset B_\rho(t_0, y_0)$ gilt. Dies liefert zusammen mit (2.20) und $y^*(t_0) = y_0$ anhand des Hauptsatzes der Differential- und Integralrechnung:

$$\begin{aligned} M\epsilon =| y^*(t^*) - y_0 |=| y^*(t^*) - y^*(t_0) | &=| \int_{t_0}^{t^*} (y^*)'(s)\, ds |=| \int_{t_0}^{t^*} f(s, y^*(s))\, ds | \\ &\leq \int_{t_0}^{t^*} | f(s, y^*(s)) | \ ds \leq M\, T < M\epsilon, \end{aligned}$$

also $M\epsilon < M\epsilon$, was offenbar ein Widerspruch ist. Insbesondere beweist dies also graph$(y) \subset B_\rho(t_0, y_0)$ für jede Lösung y von (2.1) auf $[t_0-\epsilon, t_0+\epsilon]$. Nun wählten wir $\rho > 0$ zu Beginn des Beweises so klein, dass f die Lipschitz-Bedingung (2.17) auf $B_\rho(t_0, y_0)$ erfüllt. Somit folgt aus Satz 2.0.4, hier angewandt auf $\Omega := B_\rho(t_0, y_0)$, dass eine beliebige Lösung y von (2.1) auf $[t_0 - \epsilon, t_0 + \epsilon]$ mit dem oben konstruierten Fixpunkt z von $\Phi : \mathcal{F} \longrightarrow \mathcal{F}$ übereinstimmen muss. $\square$

Definition 2.0.2. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen. Eine Lösung $y : I \longrightarrow \mathbb{R}^n$ der Differentialgleichung $y' = f(\,\cdot\,, y)$ heisst* rechtsmaximal*, falls es keine (!) Lösung $z : J \longrightarrow \mathbb{R}^n$ dieser Differentialgleichung auf einem Intervall J mit $I \subset J$ gibt, sodass $y(t) = z(t)$ für alle $t \in I$ gilt und es ein $\bar{t} \in J$ mit $\sup I < \bar{t}$ gibt. Analog heisse eine Lösung $y : I \longrightarrow \mathbb{R}^n$ der Differentialgleichung $y' = f(\,\cdot\,, y)$* linksmaximal*, falls es keine (!) Lösung $z : J \longrightarrow \mathbb{R}^n$ dieser Differentialgleichung auf einem Intervall J mit $I \subset J$ gibt, sodass $y(t) = z(t)$ für alle $t \in I$ gilt und es ein $\bar{t} \in J$ mit $\bar{t} < \inf I$ gibt. Wir nennen eine Lösung* maximal*, falls sie sowohl rechts- als auch linksmaximal ist.*

Theorem 2.0.6. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen. Für eine rechtsmaximale Lösung y der Differentialgleichung $y' = f(\,\cdot\,, y)$ ist das Definitionsintervall rechtsoffen, und das Paar $(t, y(t))$ verlässt für $t \nearrow \sup I$ jedes Kompaktum K von Ω, d.h. präzise: Zu jedem Kompaktum $K \subset \Omega$ existiert ein $t_+ \in I$ mit*

$$(t, y(t)) \notin K \qquad \text{für alle} \qquad t \geq t_+. \tag{2.22}$$

Weiter kann jede Lösung von $y' = f(\,\cdot\,, y)$ zu einer (rechts-)maximalen Lösung fortgesetzt werden.

Proof. Wir nehmen zunächst an, es existierte eine rechtsmaximale Lösung y von $y' = f(\,\cdot\,, y)$ und dass $b := \sup I \in I \subset \mathbb{R}$ gälte. Dann gilt $(b, y(b)) \in \Omega$, und nach dem Existenzsatz von Peano existiert eine Lösung $z : [b, b+\epsilon] \longrightarrow \mathbb{R}^n$ des AWP's $z' = f(\cdot, z)$, $z(b) = y(b)$. Setzen wir y durch $y(t) := z(t)$ für $t \in [b, b+\epsilon]$ fort, so gilt $y'(t) = f(t, y(t))$ für $t \in I \cup [b, b+\epsilon] \setminus \{b\}$ und ausserdem:

$$\lim_{t \nearrow b} y(t) = y(b) = z(b) = \lim_{t \searrow b} z(t)$$

also dass die Fortsetzung y in $t = b$ stetig ist und weiterhin:

$$\lim_{t \nearrow b} y'(t) = \lim_{t \nearrow b} f(t, y(t)) = f(b, y(b)) = f(b, z(b)) = \lim_{t \searrow b} f(t, z(t)) = \lim_{t \searrow b} z'(t),$$

woraus sich ergibt, dass die links- und rechtsseitigen Ableitungen der Fortsetzung y in $t = b$ existieren, übereinstimmen und mit $f(b, y(b)) = f(b, z(b))$ zusammenfallen. Insgesamt folgt hieraus, dass die Fortsetzung y auf ganz $I \cup [b, b+\epsilon]$ stetig differenzierbar ist und der Differentialgleichung $y' = f(\,\cdot\,, y)$ auf $I \cup [b, b+\epsilon]$ genügt, was der Rechtsmaximalität der ursprünglichen Lösung y widerspricht. Also muss die obige Annahme „$\sup I \in I \subset \mathbb{R}$" falsch und somit das Existenz-Intervall I von y rechtsoffen sein.
Nun betrachten wir ein Kompaktum $K \subset \Omega$ und nehmen die Existenz einer Folge $t_m \nearrow \sup I$ mit $(t_m, y(t_m)) \in K$ für alle m an. Anhand der Kompaktheit von K können wir nach Übergang zu einer Teilfolge

$$(t_m, y(t_m)) \to (t_0, y_0)$$

für ein Paar $(t_0, y_0) \in K$ bei $m \to \infty$ annehmen. Insbesondere folgt hieraus: $t_0 = \sup I \in \mathbb{R}$. Wir wählen nun ein $\rho > 0$ so klein, dass $[t_0 - \rho, t_0 + \rho] \times \overline{B_{2\rho}(y_0)} \subset \Omega$ gilt und setzen

$$M := \sup_{[t_0-\rho, t_0+\rho] \times \overline{B_{2\rho}(y_0)}} |\, f \,| < \infty.$$

Nach Definition von t_0 und y_0 gilt $y(t_m) \in B_\rho(y_0)$ und $0 \le t_0 - t_m < \frac{\rho}{M+1}$ für alle hinreichend grossen m. Wir zeigen nun, dass

$$y(t) \in B_{2\rho}(y_0) \qquad \text{für alle} \qquad t \in [t_m, t_0), \tag{2.23}$$

für ein derart gross gewähltes, festes m gilt. Angenommen, dies sei falsch. Dann setzen wir

$$\bar{t} := \inf\{t \in [t_m, t_0) \mid y(t) \notin B_{2\rho}(y_0)\}$$

und erhalten aus $y(t_m) \in B_\rho(y_0)$ und wegen der Stetigkeit von y:

$$t_m < \bar{t} < t_0, \qquad y(t) \in B_{2\rho}(y_0) \qquad \forall\, t \in [t_m, \bar{t}) \qquad \text{und} \qquad y(\bar{t}) \in \partial B_{2\rho}(y_0).$$

Hieraus erhalten wir

$$| y'(t) | = | f(t, y(t)) | \le M \qquad \forall\, t \in [t_m, \bar{t})$$

und somit per Integration über $[t_m, \bar{t})$ zusammen mit $0 \le t_0 - t_m < \frac{\rho}{M+1}$ und $y(t_m) \in B_\rho(y_0)$:

$$\rho \le | y(\bar{t}) - y(t_m) | \le M \mid \bar{t} - t_m | < \rho,$$

was unmöglich ist, und (2.23) ist damit bewiesen. Hiermit erhalten wir also nun sogar

$$| y'(t) | = | f(t, y(t)) | \le M \qquad \forall\, t \in [t_m, t_0)$$

und somit per Integration über $[t_m, t]$ für beliebiges $t \in [t_m, t_0)$ wie soeben:

$$| y(t) - y(t_m) | \le M \mid t - t_m \mid .$$

Dies zeigt uns:

$$\limsup_{t \nearrow t_0} | y(t) - y_0 | \le \limsup_{t \nearrow t_0} | y(t) - y(t_m) | + | y(t_m) - y_0 | \le M \mid t_0 - t_m \mid + \mid y(t_m) - y_0 \mid .$$

Lassen wir nun m gegen ∞ streben, so ergibt sich hieraus wegen $(t_m, y(t_m)) \to (t_0, y_0)$:

$$\lim_{t \nearrow t_0} y(t) = y_0. \tag{2.24}$$

Definieren wir nun $y(t_0) := y_0$, so haben wir die ursprüngliche rechtsmaximale Lösung y wegen (2.24) stetig auf $I \cup \{t_0\}$ fortgesetzt, und ausserdem gilt:

$$\lim_{t \nearrow t_0} y'(t) = \lim_{t \nearrow t_0} f(t, y(t)) = f(t_0, y_0) = f(t_0, y(t_0)).$$

Hieraus folgt insbesondere, dass y in t_0 differenzierbar ist, da wir nun

$$y'(t_0) = \lim_{h \nearrow 0} \frac{y(t_0 + h) - y(t_0)}{h} = \lim_{t \nearrow t_0} y'(t) = f(t_0, y(t_0))$$

schliessen können. Somit ist y eine Lösung von $y' = f(\,\cdot\,, y)$ auf $I \cup \{t_0\}$ im Widerspruch zur Rechtsmaximalität der Lösung y auf I. Also muss die Behauptung (2.22) korrekt sein.

Schliesslich ist zu zeigen, dass jede Lösung y von $y' = f(\,\cdot\,, y)$ zu einer rechtsmaximalen Lösung dieser Differentialgleichung fortgesetzt werden kann. Betrachten wir hierzu eine nicht rechtsmaximale Lösung y, so muss $b := \sup I < \infty$ gelten, und es muss eine Lösung z auf einem Intervall J mit $I \subset J$ und $\bar{t} > t$ für alle $t \in I$ und ein $\bar{t} \in J$ existieren. Wäre J rechtsabgeschlossen, so folgte aus dem bereits Bewiesenen, dass z nicht rechtsmaximal wäre und somit zu einer Lösung $\tilde{z}$ auf ein rechtsoffenes Intervall $\tilde{J}$ echt nach rechts fortgesetzt werden könnte. Daher dürfen wir a priori J als rechtsoffen annehmen und somit die Existenz einer Zahl $q \in \mathbb{Q} \cap J$ mit $q > \bar{t} \geq \sup I$. Nun wählen wir eine Aufzählung der rationalen Zahlen $\mathbb{Q}$, d.h. eine bijektive Abbildung $\mathbb{N} \longrightarrow \mathbb{Q}$, die jeder natürlichen Zahl m genau eine rationale Zahl q_m zuordnet. Wir setzen nun

$$m_1 := \min\{m \in \mathbb{N} \mid q_m > b,\ y \text{ kann als Lösung auf } I \cup [b, q_m) \text{ fortgesetzt werden}\}$$

und wählen eine Fortsetzung y_{m_1} von y auf das Intervall $I \cup [b, q_{m_1})$. Wir nehmen nun an, dass y nicht (!) zu einer rechtsmaximalen Lösung fortgesetzt werden kann. Somit könnte insbesondere die Fortsetzung y_{m_1} nicht rechtsmaximal sein. Ersetzen wir also y durch y_{m_1}, so erhalten wir wie oben einen Index $m_2 > m_1$ durch

$$m_2 := \min\{m \in \mathbb{N} \mid q_m > q_{m_1} > b,\ y_{m_1} \text{ kann als Lösung auf } I \cup [b, q_m) \text{ fortgesetzt werden}\}. \tag{2.25}$$

Da wir weiterhin annehmen, dass y nicht zu einer rechtsmaximalen Lösung fortgesetzt werden kann, kann auch deren Fortsetzung y_{m_2} nicht rechtsmaximal sein. Fortsetzung dieses Verfahrens liefert somit eine Folge von Indizes

$$m_1 < m_2 < \ldots < m_k < \ldots, \quad \text{rationale Zahlen} \quad q_{m_1} < q_{m_2} < \ldots < q_{m_k} < \ldots$$

mit $b < q_{m_1}$ und dazu Funktionen $y_{m_k} : I \cup [b, q_{m_k}) \to \mathbb{R}^n$, welche die Gleichung $y' = f(\,\cdot\,, y)$ auf $I \cup [b, q_{m_k})$ lösen und sich jeweils gegenseitig fortsetzen. Definieren wir nun also $\bar{b} := \sup_{k \in \mathbb{N}} q_{m_k}$, so können wir eine Funktion z auf dem Intervall $J := I \cup [b, \bar{b})$ durch die Vorschrift $z(t) := y_{m_k}(t)$ für $t \in I \cup [b, q_{m_k})$, $\forall\, k \in \mathbb{N}$, wohldefinieren, welche die Gleichung $y' = f(\,\cdot\,, y)$ auf ganz J löst und alle Lösungen y_{m_k} simultan fortsetzt. Angenommen, diese Lösung z wäre nicht rechtsmaximal, so könnte z wiederum auf ein Intervall $J \cup [\bar{b}, q^*)$ für ein $q^* = q_{m^*} \in \mathbb{Q}$ mit $q^* > \bar{b}$ zu einer Lösung $\tilde{z}$ echt fortgesetzt werden. Da $m_k \geq k$ für jedes k gilt, existiert ein k^* mit $m_{k^*} > m^*$. Da nun $\tilde{z}$ insbesondere $y_{m_{k^*-1}}$ als Lösung von $y' = f(\,\cdot\,, y)$ auf das Intervall $J \cup [\bar{b}, q_{m^*}) = I \cup [b, q_{m^*})$ echt fortsetzt, folgt aus der Definition des folgenden „minimalen" Index m_{k^*} – wie oben in (2.25) für die Definition von $m_2 > m_1$ – dass doch $m_{k^*} \leq m^*$ gelten müsste, was im Widerspruch zu $m_{k^*} > m^*$ steht. Somit ist die oben konstruierte Funktion z eine rechtsmaximale Lösung von $y' = f(\,\cdot\,, y)$. □

Wir sahen bereits in Beispiel 6, dass die Längen der Intervalle von maximalen Lösungen der nicht-linearen Differentialgleichung

$$y' = f(\,\cdot\,, y) := 1 + y^2$$

exakt π betragen. Insbesondere kann keine maximale Lösung dieser Gleichung auf ganz $\mathbb{R}$ existieren ! Der folgende Satz zeigt uns, dass dieses Phänomen nicht eintreten kann, falls f in y ein höchstens lineares Wachstum hat.

Theorem 2.0.7. *Es sei* $f : \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}^n$ *stetig und von höchstens linearem Wachstum in* y *auf kompakten Zeit-Intervallen* $[-R, R]$, *d.h. zu jedem* $R > 0$ *existiere eine (von* f *abhängige) Konstante* $M_R < \infty$, *sodass*

$$\mid f(t,y) \mid \leq M_R\,(1+ \mid y \mid) \qquad \text{für alle } (t,y) \in [-R,R]\times\mathbb{R}^n \tag{2.26}$$

gelte. Dann sind alle maximalen Lösungen $y : I \longrightarrow \mathbb{R}^n$ *von* $y' = f(\,\cdot\,, y)$ *„global", d.h.* $I = \mathbb{R}$.

Proof. Wir betrachten eine maximale Lösung $y : I \longrightarrow \mathbb{R}^n$, $I \neq \emptyset$, nehmen im ersten Fall: $b := \sup I < \infty$ an und wählen ein $t_0 \in (\inf I, b)$. Wir wählen ein $R > \max\{\mid t_0 \mid, \mid b \mid\}$, betrachten die Funktion $\triangle := 1+ \mid y \mid^2$ auf I und rechnen mit $2\,ab \leq a^2 + b^2$ und (2.26):

$$\begin{aligned} \triangle'(t) &= 2\langle y(t), y'(t)\rangle \leq 2\,M_R \mid y(t) \mid (1+ \mid y(t) \mid) \\ &\leq M_R\,(2 + 3 \mid y(t) \mid^2) \leq 3M_R\,\triangle(t) \qquad \text{für alle } t \in [t_0, b). \end{aligned} \tag{2.27}$$

Also folgt $\triangle'(t)\triangle(t) \leq 3M_R\,\triangle^2(t)$ für alle $t \in [t_0, b)$, und das Lemma von Gronwall, Theorem 2.0.2, liefert:

$$1+ \mid y(t) \mid^2 = \triangle(t) \leq \triangle(t_0)\,e^{3\,M_R\,(b-t_0)} =: \Gamma < \infty \qquad \text{für alle } t \in [t_0, b). \tag{2.28}$$

Betrachten wir also die kompakte Teilmenge $K := [t_0, b] \times \overline{B_\Gamma(0)}$ von $\mathbb{R}\times\mathbb{R}^n$, so besagt (2.28), dass das Paar $(t, y(t))$ das Kompaktum K für kein $t \in [t_0, b)$ verlässt, also dass die maximale Lösung y die in Theorem 2.0.6 bewiesene Eigenschaft (2.22) nicht besitzt, was unmöglich ist. Dies beweist $\sup I = \infty$.
Im zweiten Fall nehmen wir $b := \inf I > -\infty$ an, wählen ein $t_0 \in (b, \sup I)$ und betrachten anstatt y die Funktion $\tilde{y}(t) := y(2t_0 - t)$ für $t \in [t_0, 2t_0 - b)$ und sehen wie zum Ende des Beweises von Theorem 2.0.4, dass diese

$$\tilde{y}'(t) = -y'(2t_0 - t) = -f(2t_0 - t, \tilde{y}(t)) = \tilde{f}(t, \tilde{y}(t)) \qquad \text{für alle } t \in [t_0, 2t_0 - b)$$

mit $\tilde{f}(t, y) := -f(2t_0 - t, y)$ erfüllt. Man bemerke, dass die Rechtsmaximalität von y zur Linksmaximalität von $\tilde{y}$ und die Linksmaximalität von y zur Rechtsmaximalität von $\tilde{y}$ äquivalent ist. Für $t \in [-R, R]$ ist $\mid 2t_0 - t \mid < 3R$, wegen $\mid t_0 \mid < R$, und somit erhalten wir aus (2.26) für $\tilde{f}(t, y) = -f(2t_0 - t, y)$:

$$\mid \tilde{f}(t,y) \mid = \mid f(2t_0 - t, y) \mid \leq M_{3R}\,(1+ \mid y \mid) \qquad \text{für alle } (t,y) \in [-R,R]\times\mathbb{R}^n.$$

Somit erhalten wir wie oben, dass das Paar $(t, \tilde{y}(t)) \in \tilde{K} := [t_0, 2t_0 - b] \times \overline{B_{\tilde{\Gamma}}(0)}$, hier für $\tilde{\Gamma} := (1+ \mid y(t_0) \mid^2)e^{3\,M_{3R}\,(t_0-b)} < \infty$, für alle $t \in [t_0, 2t_0 - b)$ gelten müsste, was wiederum Theorem 2.0.6 widerspricht, da y linksmaximal und somit $\tilde{y}$ rechtsmaximal ist. Also schliessen wir ebenfalls: $\inf I = -\infty$. □

Zum Abschluss dieses Kapitels leiten wir noch aus Theorem 2.0.2 eine erste Variante der „stetigen Abhängigkeit" der eindeutigen Lösungen von AWP's (2.1) von deren Anfangswerten ab, bei Lipschitz-stetigem f:

Theorem 2.0.8. *[Erster Stetigkeits-Satz] Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig auf einer offenen Teilmenge $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ und Lipschitz-stetig in y, erfülle also*

$$| f(t, y_1) - f(t, y_2) | \leq L \mid y_1 - y_2 \mid \tag{2.29}$$

für ein $L \in [0, \infty)$ und für beliebige Paare $(t, y_1), (t, y_2) \in \Omega$. Seien nun $y_0^1, y_0^2 \in \Omega$ zwei festgewählte Anfangswerte und y_1, y_2 die beiden (eindeutigen) Lösungen aus Theorem 2.0.5 der entsprechenden AWP's

$$y' = f(\,\cdot\,, y), \qquad y(t_0) = y_0^i, \qquad i = 1, 2$$

auf einem hinreichend kleinen Intervall $[t_0 - \epsilon, t_0 + \epsilon]$, so gilt für deren gegenseitigen Abstand:

$$| y_1(t) - y_2(t) | \leq | y_0^1 - y_0^2 \mid e^{L|t-t_0|} \tag{2.30}$$

für alle $t \in [t_0 - \epsilon, t_0 + \epsilon]$.

Proof. Wie im Beweis von Theorem 2.0.4 definieren wir zu y_1, y_2 die Funktion $\triangle := | y_1 - y_2 |^2$ auf $I := [t_0 - \epsilon, t_0 + \epsilon]$ und erhalten aus der Voraussetzung (2.29) an f:

$$\begin{aligned}\triangle'(t) = 2\,\langle y_1(t) - y_2(t), y_1'(t) - y_2'(t)\rangle &= 2\,\langle y_1(t) - y_2(t), f(t, y_1) - f(t, y_2)\rangle \\ &\leq 2L \mid y_1(t) - y_2(t) \mid^2 = 2L\triangle(t)\end{aligned}$$

und somit $\triangle'(t)\triangle(t) \leq 2L\triangle^2(t)$, für jedes $t \in I$. Aus Theorem 2.0.2, also aus der ersten Version des Lemmas von Gronwall, folgt somit

$$\triangle(t) \leq \triangle(t_0)\,\exp(2L(t - t_0)) \qquad \text{für alle } t \in [t_0, t_0 + \epsilon],$$

was wegen $\triangle(t_0) = | y_0^1 - y_0^2 |^2$ gerade die Abschätzung (2.30) für alle $t \in [t_0, t_0 + \epsilon]$ ist. Durch Betrachtung der Funktionen $\tilde{y}_i(t) := y_i(2t_0 - t)$, $i = 1, 2$, erhalten wir hieraus dann ähnlich wie zum Ende des Beweises von Theorem 2.0.4 auch die Abschätzung (2.30) (anhand der Voraussetzung (2.29) an f) für alle $t \in [t_0 - \epsilon, t_0]$. $\square$

3 Lineare Systeme

3.1 Grundlagen

In diesem Kapitel erarbeiten wir die grundlegenden Sätze über homogene, lineare Systeme von Differentialgleichungen erster Ordnung, d.h. über Systeme der Form

$$y'(t) = A(t) \cdot y(t) \qquad \text{für} \qquad t \in [a, b], \tag{3.1}$$

für eine stetige, matrix-wertige Funktion $A : [a, b] \longrightarrow \mathrm{Mat}_{n,n}(\mathbb{R})$. Wir werden diese u.a. anwenden, um homogene, lineare Differentialgleichungen $n-$ter Ordnung, also Differentialgleichungen der Form

$$y^{(n)} + a_{n-1}\, y^{(n-1)} + a_{n-2}\, y^{(n-2)} + \ldots + a_0\, y = 0 \qquad \text{auf} \qquad [a, b]$$

für stetige Funktionen $a_0, \ldots, a_{n-1} : [a, b] \to \mathbb{R}$ zu untersuchen, bzw. für nur konstante Koeffizienten(-funktionen) $a_0, \ldots, a_{n-1}$ explizit und vollständig zu lösen. Bis zum Ende der Vorlesung versehen wir den Raum $\mathrm{Mat}_{n,n}(\mathbb{R})$ der quadratischen, reellen Matrizen mit der euklidischen Norm, setzen also

$$\mid A \mid := \big(\sum_{i,j=1} a_{ij}^2 \big)^{1/2} \tag{3.2}$$

für jedes $A = (a_{ij}) \in \mathrm{Mat}_{n,n}(\mathbb{R})$.

Proposition 3.1.1. *Es sei $A : [a, b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$, für $-\infty < a < b < \infty$, eine stetige Funktion.*

1) *Dann existieren die maximalen Lösungen y von (3.1) auf ganz $[a, b]$, d.h. sind „global", und bilden einen linearen $n-$dimensionalen Unterraum, „$\mathrm{Lös}(A)$", des Raumes $C^1([a, b], \mathbb{R}^n)$ der auf $[a, b]$ stetig differenzierbaren Funktionen.*

2) *Zu fixiertem $t_0 \in [a, b]$ und einem Startvektor $\xi \in \mathbb{R}^n$ existiert genau eine maximale Lösung y_ξ des AWP's*

$$y'(t) = A(t) \cdot y(t), \qquad y(t_0) = \xi \tag{3.3}$$

auf $[a, b]$. Die somit konstruierte Abbildung $y : \mathbb{R}^n \to C^1([a, b], \mathbb{R}^n)$ liefert einen Isomorphismus des $\mathbb{R}^n$ auf den Lösungsraum $\mathrm{Lös}(A)$ von (3.1).

Proof. Wir setzen A mittels des Satzes von Tietze auf ganz $\mathbb{R}$ mit $\sup_{t\in\mathbb{R}} \mid A(t) \mid = \max_{t\in[a,b]} \mid A(t) \mid$ fort und sehen $\mid A(t)\cdot y \mid \leq (\max_{t\in[a,b]} \mid A(t) \mid) \mid y \mid \forall\, t \in \mathbb{R}$. Somit dürfen wir die Theoreme 2.0.1 und 2.0.7 auf das stetige Vektorfeld $f(t,y) := A(t)\cdot y$ anwenden und erhalten, dass das AWP (3.3) zu jedem Startwert $\xi \in \mathbb{R}^n$ eine Lösung besitzt, die ausserdem global ist, also auf ganz $[a,b]$ existent ist. Desweiteren ist die Nullfunktion eine Lösung von (3.1), und für je zwei globale Lösungen, y_1, y_2, ist jede reelle Linearkombination $\alpha\, y_1 + \beta y_2$ wieder eine (globale) Lösung von (3.1). Somit bildet $\text{Lös}(A)$ einen linearen Unterraum von $C^1([a,b],\mathbb{R}^n)$. Desweiteren erfüllt $f(t,y) := A(t)\cdot y$ mit $L := \max_{t\in[a,b]} \mid A(t) \mid$

$$\mid f(t,y_1) - f(t,y_2) \mid \leq L \mid y_1 - y_2 \mid \tag{3.4}$$

für beliebige Paare $(t,y_1),(t,y_2)$ aus $[a,b]\times\mathbb{R}^n$, also gerade die Lipschitz-Bedingung (2.16) aus Satz 2.0.4. Somit ist also zu fixiertem $t_0 \in [a,b]$ und beliebig vorgegebenem Startvektor $\xi \in \mathbb{R}^n$ jede maximale Lösung y_ξ des AWP's (3.3) eindeutig, sodass die Abbildung $\xi \mapsto y_\xi$ wohldefiniert ist. Ausserdem sehen wir

$$y_{\alpha\xi+\beta\eta}(t_0) = \alpha\,\xi + \beta\,\eta = (\alpha\, y_\xi + \beta\, y_\eta)(t_0),$$

und wir erhalten anhand der Eindeutigkeit einer jeden maximalen Lösung, dass die Lösung $y_{\alpha\xi+\beta\eta}$ zum Startwert $\alpha\xi+\beta\eta$ mit der Linearkombination $\alpha\, y_\xi + \beta\, y_\eta$ der Lösungen y_ξ, y_η übereinstimmen muss, also

$$y_{\alpha\xi+\beta\eta} = \alpha\, y_\xi + \beta\, y_\eta$$

und somit die Linearität der Abbildung y. Ist nun z eine Lösung von (3.1), also $z \in \text{Lös}(A)$, so erhalten wir durch $y_{z(t_0)}$ eine Funktion aus $\text{Lös}(A)$, die wegen $y_{z(t_0)}(t_0) = z(t_0)$ und wieder wegen Satz 2.0.4 mit z auf ganz $[a,b]$ übereinstimmen muss. Also ist y surjektiv auf $\text{Lös}(A)$. Und gilt $y_\xi \equiv 0$ auf $[a,b]$, so folgt insbesondere: $\xi = y_\xi(t_0) = 0$, also dass die lineare Abbildung $\xi \mapsto y_\xi$ injektiv ist. y liefert somit einen kanonischen Isomorphismus des $\mathbb{R}^n$ auf $\text{Lös}(A)$. □

Definition 3.1.1. *Eine stetig differenzierbare Abbildung $Y : [a,b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$ heisse eine Fundamentalmatrix von* (3.1)*, falls deren Spalten(-Funktionen) eine Basis von* $\text{Lös}(A)$ *bilden.*

Proposition 3.1.2. *Es sei $A : [a,b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$, für $-\infty < a < b < \infty$, eine stetige Funktion. Dann ist $Y : [a,b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$ genau dann eine Fundamentalmatrix von* (3.1)*, wenn Y die Matrix-Differentialgleichung*

$$Y'(t) = A(t)\cdot Y(t), \qquad \text{für alle} \qquad t \in [a,b] \tag{3.5}$$

löst und $Y(t_0)$ in einem beliebig fixierten Zeitpunkt $t_0 \in [a,b]$ invertierbar ist.

Proof. Es sind genau dann alle Spalten $y_1, y_2, \ldots, y_n$ von Y in $\text{Lös}(A)$ enthalten, wenn die Matrix(-Funktion) $Y = (y_1,\ldots,y_n)$ die Differentialgleichung (3.5) auf $[a,b]$ löst. Wir wählen nun ein $t_0 \in [a,b]$ beliebig und bezeichnen wieder mit $y : \mathbb{R}^n \to C^1([a,b],\mathbb{R}^n)$ die

Isomorphie des $\mathbb{R}^n$ auf Lös(A) aus (3.3). Für die Spaltenfunktionen einer Lösung Y von (3.5) gilt:

$$y_i = y_{y_i(t_0)}. \tag{3.6}$$

Da die Zuordnung $\xi \mapsto y_\xi$ aus (3.3) ein Isomorphismus ist, bilden somit seine n Bilder $y_{y_1(t_0)}, y_{y_2(t_0)}, \ldots, y_{y_n(t_0)}$ genau dann eine Basis von Lös(A), falls die n Startvektoren $y_1(t_0), \ldots, y_n(t_0)$ eine Basis des $\mathbb{R}^n$ bilden. Wegen (3.6) und anhand Definition 3.1.1 bedeutet dies: $Y = (y_1, \ldots, y_n)$ ist genau dann eine Fundamentalmatrix von (3.1), falls $Y(t_0) = (y_1, \ldots, y_n)(t_0)$ invertierbar ist. □

Proposition 3.1.3. *Es sei $A : [a, b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$, für $-\infty < a < b < \infty$, eine stetige Funktion. Dann erfüllt die Determinante einer Lösung $Y : [a, b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$ von (3.5) die Differentialgleichung*

$$\det(Y)' = \mathrm{Spur}(A)\,\det(Y) \qquad \text{auf} \qquad [a, b]. \tag{3.7}$$

Insbesondere folgt hieraus, dass die Determinante einer Fundamentalmatrix $Y : [a, b] \to \mathrm{Mat}_{n,n}(\mathbb{R})$ von (3.1) $\det(Y)(t) \neq 0$ und

$$\det(Y)(t) = \det(Y(t_0))\,\exp\Big(\int_{t_0}^{t} \mathrm{Spur}(A)(s)\,ds\Big) \qquad \text{für alle} \qquad t_0 \le t \in [a, b] \tag{3.8}$$

erfüllt.

Proof. Wir betrachten anstatt Y deren Transponierte Y^T, deren Spalten also die Zeilen z_i von Y sind, und sehen anhand von (3.5), d.h. anhand von $(Y^T)'(t) = Y^T(t) \cdot A^T(t)$:

$$z_i'(t) = \sum_{j=1}^{n} a_{ij}(t)\, z_j(t), \qquad \text{für alle} \qquad t \in [a, b]. \tag{3.9}$$

Anhand der Linearität der Determinante in jeder ihrer Spalten berechnet man die folgende ,,Produkt-Ableitungsformel"

$$\frac{d}{dt} \det(\zeta_1, \zeta_2, \ldots, \zeta_n)(t) = \sum_{i=1}^{n} \det(\zeta_1(t), \ldots, \zeta_i'(t), \ldots, \zeta_n(t)),$$

für n beliebige, stetig differenzierbare Funktionen $\zeta_i : [a, b] \to \mathbb{R}^n$. Somit erhalten wir aus (3.9):

$$\begin{aligned}
\det(Y)'(t) &= \det(Y^T)'(t) = \det((z_1, z_2, \ldots, z_n))'(t) \\
&= \sum_{i=1}^{n} \det((z_1, \ldots, z_{i-1}, z_i', z_{i+1}, \ldots, z_n))(t) \\
&= \sum_{i=1}^{n} \det((z_1, \ldots, z_{i-1}, \sum_{j=1}^{n} a_{ij}\, z_j, z_{i+1}, \ldots, z_n))(t) \\
&= \sum_{i,j=1}^{n} a_{ij}(t)\, \det((z_1, \ldots, z_{i-1}, z_j, z_{i+1}, \ldots, z_n))(t) \\
&= \sum_{i=1}^{n} a_{ii}(t)\, \det((z_1, \ldots, z_i, \ldots, z_n))(t) \\
&= \mathrm{Spur}(A(t))\, \det(Y^T)(t) = \mathrm{Spur}(A(t))\, \det(Y)(t).
\end{aligned}$$

Somit folgt die letzte Behauptung (3.8) zu einer Fundamentalmatrix Y von (3.1) zusammen mit Proposition 3.1.2 und Beispiel 5 von Kapitel 1. □

Als eine interessante Anwendung studieren wir nun homogene, lineare Differentialgleichungen n-ter Ordnung, d.h. Differentialgleichungen der Form

$$y^{(n)} + a_{n-1}\, y^{(n-1)} + a_{n-2}\, y^{(n-2)} + \ldots + a_0\, y = 0 \qquad \text{auf} \qquad [a, b], \tag{3.10}$$

für n-fach stetig differenzierbare Funktionen $y : [a, b] \to \mathbb{R}$, wobei wir hierbei annehmen, dass die Koeffizientenfunktionen $a_0, \ldots, a_{n-1} : [a, b] \to \mathbb{R}$ stetig seien. Man überlege sich zunächst, dass y genau dann eine Lösung der skalaren Gleichung (3.10) auf $[a, b]$ ist, wenn die $\mathbb{R}^n-$wertige C^1-Funktion $z := (y, y', y'', \ldots, y^{(n-1)})$ eine Lösung des Systems 1. Ordnung

$$z' = A \cdot z \qquad \text{auf} \qquad [a, b] \tag{3.11}$$

ist, wobei A die matrix-wertige Funktion

$$\begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & 0 & 0 & \ldots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0 & -a_1 & \ldots & -a_{n-2} & -a_{n-1} \end{pmatrix} \tag{3.12}$$

bezeichne. Somit muss der Lösungsraum „Lös$(a_0, a_2, \ldots, a_{n-1})$" von (3.10) isomorph zum linearen Lösungsraum Lös(A) des Gleichungssystems (3.11) 1. Ordnung sein, also anhand von Proposition 3.1.1 isomorph zum $\mathbb{R}^n$. Präziser erhalten wir aus dem zweiten Teil von Proposition 3.1.1, dass die Isomorphie des $\mathbb{R}^n$ auf Lös(A) und somit auf Lös$(a_0, a_2, \ldots, a_{n-1})$ durch die Zuordnung $z(t_0) = (y(t_0), \ldots, y^{(n-1)}(t_0)) \mapsto z = (y, y', y'', \ldots, y^{(n-1)})$ gegeben ist. D.h. zu fixiertem $t_0 \in [a, b]$ existiert zu jedem vorgegebenen Tupel $(y_0, y_0^1, \ldots, y_0^{(n-1)}) \in \mathbb{R}^n$ genau eine Lösung y von (3.10), deren ersten $n-1$

Ableitungen zum Zeitpunkt t_0 gerade die vorgegebenen Werte $y(t_0) = y_0$, $y'(t_0) = y_0^1$, $\ldots$, $y^{(n-1)}(t_0) = y_0^{(n-1)}$ hat.

Im Hinblick auf die Propositionen 3.1.2 und 3.1.3 legt dies die folgende Definition nahe:

Definition 3.1.2. *n C^n-Funktionen $y_1, y_2, \ldots, y_n : [a,b] \to \mathbb{R}$ heissen ein ,,Fundamentalsystem" von (3.10), falls sie eine Basis des n-dimensionalen Lösungsraums von (3.10) bilden.*
Desweiteren definieren wir die ,,Wronski-Determinante" W von n Lösungen $y_1, y_2, \ldots, y_n$ von (3.10) als Determinante der Matrix

$$\begin{pmatrix} y_1 & y_2 & \ldots & y_n \\ y_1' & y_2' & \ldots & y_n' \\ \vdots & \vdots & \ldots & \vdots \\ y_1^{(n-2)} & y_2^{(n-2)} & \ldots & y_n^{(n-2)} \\ y_1^{(n-1)} & y_2^{(n-1)} & \ldots & y_n^{(n-1)} \end{pmatrix} \tag{3.13}$$

Theorem 3.1.1. *n C^n-Funktionen $y_1, y_2, \ldots, y_n : [a,b] \to \mathbb{R}$ bilden genau dann ein Fundamentalsystem von (3.10), wenn die Wronski-Determinante W von $\{y_1, y_2, \ldots, y_n\}$ in einem beliebigen Zeitpunkt $t \in [a,b]$ nicht verschwindet. $W(t)$ lässt sich vermöge der Formel*

$$W(t) = W(t_0)\, \exp(-\int_{t_0}^{t} a_{n-1}(s)\, ds) \qquad \text{für alle} \qquad t_0 \le t \in [a,b] \tag{3.14}$$

zu beliebig fixiertem $t_0 \in [a,b]$ berechnen. Ist insbesondere der Koeffizient a_{n-1} zu $y^{(n-1)}$ in (3.10) konstant, so gilt:

$$W(t) = W(t_0)\, \exp(-a_{n-1}(t - t_0)) \qquad \text{für alle} \qquad t_0 \le t \in [a,b].$$

Proof. Anhand der Äquivalenz von (3.10) und (3.11) löst die Matrix Y aus (3.13) zu n Lösungen $y_1, y_2, \ldots, y_n$ von (3.10) die Matrix-Differentialgleichung $Y' = A \cdot Y$, wobei die Matrix A in (3.12) durch die Koeffizienten $a_0, a_1, \ldots, a_{n-1}$ gegeben ist. n Funktionen $y_1, y_2, \ldots, y_n$ bilden genau dann ein Fundamentalsystem von (3.10), falls die Spalten der Matrix Y eine Basis des zu (3.10) äquivalenten Gleichungssystems (3.11) 1. Ordnung bilden, also falls – per Definition 3.1.1 – die Matrix Y eine Fundamentalmatrix des Systems 1. Ordnung (3.11) ist. Da $Y' = A \cdot Y$ gilt, ist dies nach Proposition 3.1.2 wiederum äquivalent zur Invertierbarkeit von $Y(t)$ in einem beliebigen Zeitpunkt $t \in [a,b]$, und dies ist wiederum äquivalent zu $W(t) \neq 0$ in einem beliebigen Zeitpunkt $t \in [a,b]$, wegen $W = \det Y$. Schliesslich erhalten wir aus $Y' = A \cdot Y$ zusammen mit $\mathrm{Spur}(A) = -a_{n-1}$ anhand von Proposition 3.1.3 die letzte Behauptung:

$$W(t) = \det Y(t) = \det Y(t_0)\, \exp\big(\int_{t_0}^{t} \mathrm{Spur}(A)(s)\, ds\big)$$

$$= W(t_0)\, \exp\big(-\int_{t_0}^{t} a_{n-1}(s)\, ds\big) \qquad \text{für alle} \qquad t_0 \le t \in [a,b].$$

□

3.2 Lineare Differentialgleichungen n-ter Ordnung mit konstanten Koeffizienten

In diesem Abschnitt geben wir den gesamten Lösungsraum einer linearen Differentialgleichung (3.10) $n-$ter Ordnung explizit für den Fall an, dass die Koeffizienten $a_0, a_1, \ldots, a_{n-1}$ konstante reelle Zahlen sind. Zunächst lässt sich (3.10) auch in der Form

$$P(\frac{d}{dt})(y) = 0 \qquad \text{auf} \qquad [a, b]$$

für den ,,polynomialen" Differentialoperator

$$P(\frac{d}{dt}) := \frac{d^n}{dt^n} + a_{n-1}\frac{d^{n-1}}{dt^{n-1}} + a_{n-2}\frac{d^{n-2}}{dt^{n-2}} + \ldots + a_0 : C^n([a,b],\mathbb{C}) \longrightarrow C^0([a,b],\mathbb{C})$$

schreiben. Wendet man $P(\frac{d}{dt})$ auf die Funktion $\exp(\lambda\, t)$, für ein $\lambda \in \mathbb{C}$ an, so erhält man wegen $\frac{d}{dt}\exp(\lambda\, t) = \lambda \exp(\lambda\, t)$:

$$P(\frac{d}{dt})(\exp(\lambda\, t)) = P(\lambda)\exp(\lambda\, t) = (\lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots + a_0)\exp(\lambda\, t).$$

Hat also das Polynom P n verschiedene reelle Nullstellen $\lambda_1, \lambda_2, \ldots, \lambda_n$, so sind offenbar die n Funktionen $\exp(\lambda_i\, t)$ Lösungen von (3.10). Und in der Tat sind diese Funktionen auch linear unabhängig, bilden also ein Fundamentalsystem von (3.10), denn deren Wronski-Determinante W in $t = 0$ ist hier einfach die Determinante der Matrix

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & \vdots & \ldots & \vdots \\ \lambda_1^{n-2} & \lambda_2^{n-2} & \ldots & \lambda_n^{n-2} \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{pmatrix} \tag{3.15}$$

also $W(0) = \prod_{j>i}(\lambda_j - \lambda_i) \neq 0$, woraus die obige Behauptung mit Theorem 3.1.1 folgt. In diesem simplen Fall hat das Polynom P die Aufspaltung $P(\lambda) = (\lambda - \lambda_1)\cdot(\lambda - \lambda_2)\ldots\cdot(\lambda - \lambda_n)$ in (miteinander kommutierende) Linearfaktoren, und die analoge Aussage gilt für $P(\frac{d}{dt})$, also

$$P(\frac{d}{dt}) = (\frac{d}{dt} - \lambda_1)\circ(\frac{d}{dt} - \lambda_2)\circ(\frac{d}{dt} - \lambda_n)$$

woran man ebenfalls sofort erkennt, dass $P(\frac{d}{dt})(\exp(\lambda\, t)) = 0$ genau für $\lambda \in \{\lambda_1, \ldots, \lambda_n\}$ folgt. Ist nun beispielsweise λ_1 eine Nullstelle von P 2. Ordnung und sind $\lambda_2, \ldots, \lambda_{n-1}$ $n-2$ Nullstellen 1. Ordnung, so hat P die Zerlegung

$$P(\lambda) = (\lambda - \lambda_1)^2\cdot(\lambda - \lambda_2)\ldots\cdot(\lambda - \lambda_{n-1}).$$

Es stellt sich somit die Frage, welche Funktion ausser $\exp(\lambda_1\, t)$ vom Differentialoperator $(\frac{d}{dt} - \lambda_1)^2$ annuliert (,,zur konstanten Nullfunktion gemacht") wird. In der Tat rechnen wir sofort für $t\, e^{\lambda_1 t}$ nach:

$$(\frac{d}{dt} - \lambda_1)^2(t\, e^{\lambda_1 t}) = (\frac{d}{dt} - \lambda_1)(e^{\lambda_1 t}) = 0.$$

Somit wäre sicherlich $\{t \exp(\lambda_1 t), \exp(\lambda_1 t), \exp(\lambda_2 t), \ldots, \exp(\lambda_{n-1} t)\}$ ein Fundamentalsystem von (3.10) in dem hier diskutierten 2. Spezialfall, falls wir die lineare Unabhängigkeit von $\{t \exp(\lambda_1 t), \exp(\lambda_1 t), \exp(\lambda_2 t), \ldots, \exp(\lambda_{n-1} t)\}$ nachweisen können. Wir beweisen dies und behandeln den allgemeinen Fall nun in

Theorem 3.2.1. *Es habe das der homogenen, linearen Differentialgleichung n-ter Ordnung* (3.10) *zugeordnete Polynom $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \ldots + a_0$ die (allgemeine) Aufspaltung*

$$P(\lambda) = \prod_{i=1}^{k} (\lambda - \lambda_i)^{r_i} \prod_{j=1}^{l} (\lambda - \mu_j)^{s_j} \prod_{j=1}^{l} (\lambda - \bar{\mu_j})^{s_j} \tag{3.16}$$

wobei $\lambda_1, \ldots, \lambda_k$ die (paarweise verschiedenen) reellen und $\{\mu_1, \bar{\mu_1}\}, \ldots, \{\mu_l, \bar{\mu_l}\}$ die (paarweise verschiedenen) zueinander konjugierten, komplexen Nullstellen von P seien, also mit $\mu_j = \alpha_j + i\beta_j$ und $\beta_j > 0$. Für die Ordnungen der Nullstellen gilt $r_i, s_j \geq 1$ und ausserdem

$$\sum_{i=1}^{k} r_i + 2\sum_{j=1}^{l} s_j = n.$$

Dann bilden die Funktionen

$$\begin{aligned} t^r\, e^{\lambda_i t}, \quad & r = 0, \ldots, r_i - 1, \quad i = 1, \ldots, k, \\ t^s\, e^{\alpha_j t} \cos(\beta_j t),\ t^s\, e^{\alpha_j t} \sin(\beta_j t), \quad & s = 0, \ldots, s_j - 1, \quad j = 1, \ldots, l \end{aligned} \tag{3.17}$$

ein Fundamentalsystem der homogenen, linearen Differentialgleichung (3.10) *n-ter Ordnung, d.h. bildet eine Basis des n-dimensionalen Lösungsraums von* (3.10).

Proof. Wir betrachten für beliebige $r, m \in \mathbb{N}$ und beliebiges $\mu \in \mathbb{C}$ die Funktion $y(t) := t^r\, e^{\mu t}$ und das Polynom $q(\lambda) = (\lambda - \mu)^m$ und berechnen sukzessive:

$$\begin{aligned} q(\frac{d}{dt})(y) = (\frac{d}{dt} - \mu)^{m-1}(r\, t^{r-1}\, e^{\mu t}) = r\,(r-1)(\frac{d}{dt} - \mu)^{m-2}\,(t^{r-2}\, e^{\mu t}) = \ldots \\ = r\,(r-1)\ldots(r-m+1)\, t^{r-m}\, e^{\mu t}, \end{aligned}$$

sodass gerade für $r < m$ $q(\frac{d}{dt})(y) = 0$ folgt. Anhand der Aufspaltung (3.16) von P beweist dies, dass exakt die Funktionen

$$\begin{aligned} t^r\, e^{\lambda_i t}, \quad & r = 0, \ldots, r_i - 1, \quad i = 1, \ldots, k \\ t^s\, e^{\mu_j t},\ t^s\, e^{\bar{\mu}_j t}, \quad & s = 0, \ldots, s_j - 1, \quad j = 1, \ldots, l, \end{aligned}$$

vom Differentialoperator $P(\frac{d}{dt})$ annuliert werden, was äquivalent zu der Aussage ist, dass die in (3.17) angegebenen (reellwertigen) Funktionen die Differentialgleichung (3.10) n-ter Ordnung allesamt lösen. Dies sind offenbar exakt n Funktionen, sodass nun nur noch deren lineare Unabhängigkeit nachgewiesen werden muss. Hierzu betrachten wir eine triviale Relation zwischen diesen, also reelle Zahlen $\sigma_{i,r}, \zeta_{j,s}, \eta_{j,s}$, für welche

$$\sum_{i=1}^{k}\sum_{r=0}^{r_i-1} \sigma_{i,r}\, t^r\, e^{\lambda_i t} + \sum_{j=1}^{l}\sum_{s=0}^{s_j-1} \zeta_{j,s}\, t^s\, e^{\alpha_j t} \cos(\beta_j t) + \sum_{j=1}^{l}\sum_{s=0}^{s_j-1} \eta_{j,s}\, t^s\, e^{\alpha_j t} \sin(\beta_j t) \equiv 0$$

für alle $t \in \mathbb{R}$ simultan gelte. Setzen wir $\tilde{\zeta}_{j,s} := \frac{\zeta_{j,s} - i\eta_{j,s}}{2}$ und $\tilde{\eta}_{j,s} := \frac{\zeta_{j,s} + i\eta_{j,s}}{2}$, so ist diese Relation äquivalent zur Relation

$$\sum_{i=1}^{k} \sum_{r=0}^{r_i - 1} \sigma_{i,r}\, t^r\, e^{\lambda_i t} + \sum_{j=1}^{l} \sum_{s=0}^{s_j - 1} \tilde{\zeta}_{j,s}\, t^s\, e^{\mu_j t} + \sum_{j=1}^{l} \sum_{s=0}^{s_j - 1} \tilde{\eta}_{j,s}\, t^s\, e^{\bar{\mu}_j t} \equiv 0 \tag{3.18}$$

für alle $t \in \mathbb{R}$. Nun fixieren wir ein $i \in \{1, \ldots, k\}$ und versuchen $\sigma_{i,r_i-1} = 0$ zu zeigen, indem wir die Polynome $p_i(\lambda) := P(\lambda) \cdot (\lambda - \lambda_i)^{-1}$ und $q_i(\lambda) := P(\lambda) \cdot (\lambda - \lambda_i)^{-r_i}$ betrachten. Wir sehen wie oben anhand der Linearfaktor-Zerlegung von P, dass exakt die Funktionen

$$t^r\, e^{\lambda_{i'} t}, \qquad \text{für } r = 0, \ldots, r_{i'} - 1, \quad \text{für } i' \neq i, \quad \text{und} \quad \text{für } r < r_i - 1, \quad \text{falls} \quad i' = i,$$
$$t^s\, e^{\mu_j t},\ t^s\, e^{\bar{\mu}_j t}, \qquad \text{für } s = 0, \ldots, s_j - 1, \quad j = 1, \ldots, l$$

von $p_i(\frac{d}{dt})$ annuliert werden und dass wegen $p_i(\lambda) = q_i(\lambda) \cdot (\lambda - \lambda_i)^{r_i - 1}$ und $q_i(\lambda_i) \neq 0$

$$\begin{aligned} p_i(\frac{d}{dt})(t^{r_i-1}\, e^{\lambda_i t}) &= q_i(\frac{d}{dt}) \circ (\frac{d}{dt} - \lambda_i)^{r_i - 1}(t^{r_i-1}\, e^{\lambda_i t}) \\ &= (r_i - 1)!\, q_i(\frac{d}{dt})(e^{\lambda_i t}) = (r_i - 1)!\, q_i(\lambda_i)\, e^{\lambda_i t} \neq 0 \end{aligned}$$

für alle $t \in \mathbb{R}$ gilt. Wenden wir ausserdem $p_i(\frac{d}{dt})$ auf die triviale Relation (3.18) an, so erhalten wir insgesamt:

$$\begin{aligned} 0 = p_i(\frac{d}{dt})\Big(\sum_{i=1}^{k} \sum_{r=0}^{r_i - 1} \sigma_{i,r}\, t^r\, e^{\lambda_i t} + \sum_{j=1}^{l} \sum_{s=0}^{s_j - 1} \tilde{\zeta}_{j,s}\, t^s\, e^{\mu_j t} + \sum_{j=1}^{l} \sum_{s=0}^{s_j - 1} \tilde{\eta}_{j,s}\, t^s\, e^{\bar{\mu}_j t}\Big) \\ = 0 + \ldots 0 + \sigma_{i,r_i-1}(r_i - 1)!\, q_i(\lambda_i)\, e^{\lambda_i t} + 0 + \ldots + 0, \end{aligned}$$

also in der Tat $\sigma_{i,r_i-1} = 0$. Falls die Ordnung r_i der Nullstelle λ_i grösser als 1 sein sollte, so betrachten wir desweiteren $p_i(\lambda) := P(\lambda) \cdot (\lambda - \lambda_i)^{-2}$ und wieder $q_i(\lambda) := P(\lambda) \cdot (\lambda - \lambda_i)^{-r_i}$ und erhalten wie oben, indem man einfach $r_i - 1$ nun durch $r_i - 2$ ersetzt, dass wegen $p_i(\lambda) = q_i(\lambda) \cdot (\lambda - \lambda_i)^{r_i - 2}$ und $q_i(\lambda_i) \neq 0$

$$\begin{aligned} p_i(\frac{d}{dt})(t^{r_i-2}\, e^{\lambda_i t}) &= q_i(\frac{d}{dt}) \circ (\frac{d}{dt} - \lambda_i)^{r_i - 2}(t^{r_i-2}\, e^{\lambda_i t}) \\ &= (r_i - 2)!\, q_i(\frac{d}{dt})(e^{\lambda_i t}) = (r_i - 2)!\, q_i(\lambda_i)\, e^{\lambda_i t} \neq 0 \end{aligned}$$

für alle $t \in \mathbb{R}$ gilt. Somit folgt wieder durch Anwendung von $p_i(\frac{d}{dt})$ auf die triviale Relation (3.18), zusammen mit $\sigma_{i,r_i-1} = 0$:

$$0 = 0 + \ldots 0 + \sigma_{i,r_i-2}(r_i - 2)!\, q_i(\lambda_i)\, e^{\lambda_i t} + 0 + \ldots + 0,$$

und daher $\sigma_{i,r_i-2} = 0$. Durch sukzessive Anwendung dieses Verfahrens erhalten wir somit, dass die Koeffizienten $\sigma_{i,r}$ für $r = 0, \ldots, r_i - 1$ alle gleich Null sein müssen, und analog folgert man dies für die Koeffizienten $\tilde{\zeta}_{j,s}$ und $\tilde{\eta}_{j,s}$ und damit auch für alle $\zeta_{j,s}$ und $\eta_{j,s}$. Dies beweist die lineare Unabhängigkeit der in (3.17) angegebenen Lösungen von (3.10). $\square$

Beispiel: Eine klassische Anwendung dieses Satzes besteht in der vollständigen Lösung der Differentialgleichung 2. Ordnung des „harmonischen Oszillators", d.h. der homogenen Differentialgleichung

$$\ddot{y}(t) + 2\rho\,\dot{y}(t) + \omega^2 y(t) = 0 \tag{3.19}$$

wobei ρ und ω positive, reelle Zahlen seien. In der Tat ist dies gerade die Bewegungsgleichung einer punktförmigen Masse m, die sich am Ende einer bis zur Länge $\mid y(t) \mid$ ausgelenkten Feder befindet und deren Gesamtkraft $m\ddot{y}(t)$ zum Zeitpunkt t die Summe aus der „Spannkraft" $-ky(t)$ der bis zur Länge $\mid y(t) \mid$ gedehnten Feder und einer Reibungskraft $-r\dot{y}(t)$ ist, wenn wir $\omega := \sqrt{\frac{k}{m}}$ und $\rho := \frac{r}{2m}$ in (3.19) definieren. k und r sind in diesem Fall positive Konstanten, die durch das Material der Feder gegeben sind und experimentell bestimmt werden können. Das zu (3.19) korrespondierende Polynom lautet $P(\lambda) = \lambda^2 + 2\rho\lambda + \omega^2$ und besitzt die beiden (im allgemeinen komplexen) Nullstellen

$$\lambda_1 = -\rho + \sqrt{\rho^2 - \omega^2} \qquad \text{und} \qquad \lambda_2 = -\rho - \sqrt{\rho^2 - \omega^2}. \tag{3.20}$$

Nun sind die folgenden Fälle zu unterscheiden:

Fall 1: $\rho > \omega$: Genau in diesem Fall sind die beiden Nullstellen reell und verschieden, also $\lambda_2 < \lambda_1 \in \mathbb{R}$. Theorem 3.2.1 liefert somit in diesem Fall, dass die allgemeine Lösung von (3.19) durch

$$y(t) = c_1\, e^{\lambda_1 t} + c_2\, e^{\lambda_2 t}$$

auf ganz $\mathbb{R}$, für reelle Koeffizienten c_1, c_2, gegeben ist. Da anhand der Formel (3.20) und wegen $\omega > 0$ ausserdem beide Nullstellen negativ sein müssen, also $\lambda_2 < \lambda_1 < 0$, klingen die Lösungen in diesem Fall recht schnell ab, oszillieren also offenbar gar nicht.

Fall 2: Es gilt exakt $\rho = \omega$. In diesem Fall liefert Formel (3.20): $\lambda_1 = \lambda_2 = -\rho < 0$, d.h. $P(\lambda) = (\lambda + \rho)^2$. In diesem Fall liefert Theorem 3.2.1, dass die allgemeine Lösung von (3.19) durch

$$y(t) = (c_1 + c_2\, t)\, e^{-\rho t}$$

auf ganz $\mathbb{R}$, für reelle Koeffizienten c_1, c_2, gegeben ist. Da $te^{-\rho t}$ für grosse t-Werte beinahe so klein wie $e^{-\rho t}$ ist, verhalten sich diese Lösungen ähnlich wie im ersten Fall, klingen also einfach exponentiell ab.

Fall 3: $\rho < \omega$: Genau in diesem Fall sind die beiden Nullstellen echt komplex und notwendigerweise zueinander konjugiert und verschieden, also $\lambda_1 \neq \lambda_2 = \bar{\lambda}_1 \in \mathbb{C}$ mit Realteil $\alpha := -\rho < 0$ und Imaginärteil $\beta := \sqrt{\omega^2 - \rho^2} > 0$ bzw. $-\beta < 0$. In diesem Fall liefert Theorem 3.2.1, dass die allgemeine Lösung von (3.19) durch

$$y(t) = e^{-\rho t}\, (c_1 \cos(\sqrt{\omega^2 - \rho^2}\, t) + c_2 \sin(\sqrt{\omega^2 - \rho^2}\, t))$$

auf ganz $\mathbb{R}$, für reelle Koeffizienten c_1, c_2, gegeben ist. Genau in diesem Fall schwingen also die Lösungen von (3.19) in der Tat, und zwar mit der Frequenz $\frac{\sqrt{\omega^2 - \rho^2}}{2\pi}$, klingen jedoch während ihrer Oszillation anhand der positiven Reibungskonstanten r, also wegen $\rho = \frac{r}{2m} > 0$, ebenfalls exponentiell ab.

3.3 Homogene lineare Systeme 1. Ordnung mit konstanten Koeffizienten

In diesem Abschnitt geben wir die Lösung der Matrix-Differentialgleichung (3.5) für zeitunabhängige Koeffizienten-Matrizen A explizit an. Wir definieren hierfür zunächst den Begriff der Exponentialreihe einer quadratischen Matrix:

Definition 3.3.1. *Sei $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine beliebige quadratische Matrix, so definieren wir deren ,,Exponentialreihe" durch:*

$$\exp(A) := \lim_{n\to\infty} \sum_{j=0}^{n} \frac{1}{j!} A^j, \tag{3.21}$$

wobei hier A^j als j-fache Multiplikation der Matrix A mit sich selbst zu verstehen sei.

Proposition 3.3.1. *1) Die Reihe $\sum_{j=0}^{\infty} \frac{1}{j!} A^j$ in (3.21) ist absolut konvergent für jedes $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$, sodass insbesondere $\exp(A)$ für jedes $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ existiert und ein Element von $\mathrm{Mat}_{n,n}(\mathbb{R})$ ist.*

2) Für zwei kommutierende Matrizen $A, B \in \mathrm{Mat}_{n,n}(\mathbb{R})$ gilt die ,,Funktionalgleichung":

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

Insbesondere ist für beliebiges $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ deren Exponential $\exp(A)$ invertierbar mit $(\exp(A))^{-1} = \exp(-A)$.

3) Für beliebiges $A \in \mathrm{Mat}_{n,n}(\mathbb{R})$ ist die Funktion $t \mapsto \exp(tA)$ auf ganz $\mathbb{R}$ beliebig häufig differenzierbar und besitzt die Ableitungen

$$\frac{d^n}{dt^n} \exp(tA) = A^n \exp(tA) \qquad \text{für alle} \qquad t \in \mathbb{R}. \tag{3.22}$$

Proof. Zu 1): Aus der Dreiecksungleichung und aus $\mid A \cdot B \mid \leq \mid A \mid \mid B \mid$ erhalten wir zunächst

$$\mid \sum_{j=l}^{n} \frac{1}{j!} A^j \mid \leq \sum_{j=l}^{n} \frac{1}{j!} \mid A \mid^j$$

für beliebige Indizes $l \leq n \in \mathbb{N}$. Desweiteren gilt für ein beliebiges $R > 0$ mit $\mid A \mid \leq R$ und jedes $n \in \mathbb{N}$:

$$\sum_{j=0}^{n} \frac{1}{j!} \mid A \mid^j \leq \sum_{j=0}^{\infty} \frac{1}{j!} R^j = \exp(R).$$

Also bilden die Partialsummen $\{\sum_{j=0}^{n} \frac{1}{j!} \mid A \mid^j\}$ und $\{\sum_{j=0}^{n} \frac{1}{j!} A^j\}$ Cauchyfolgen in $\mathbb{R}$ bzw. in $\mathrm{Mat}_{n,n}(\mathbb{R})$ und sind somit konvergent, anhand der Vollständigkeit von $(\mathbb{R}, \mid \cdot \mid)$ bzw. $(\mathrm{Mat}_{n,n}(\mathbb{R}), \mid \cdot \mid)$ (ausgestattet mit der euklidischen Norm (3.2)).
Die Funktionalgleichung in (2) kann man anhand der vorausgesetzten Kommutativität $A \cdot B = B \cdot A$ wie die Funktionalgleichung $\exp(z+w) = \exp(z)\,\exp(w)$, für $w, z \in \mathbb{C}$,

mittels des Cauchy-Doppelreihensatzes und der Binomischen Formel beweisen.
Zu 3.) Zunächst folgt aus Punkt (1) für die Partialsummen $P_n(t) := \sum_{j=0}^n \frac{1}{j!}\,(tA)^j$:

$$| P_n(t) - \exp(tA) | \leq \sum_{j=n+1}^{\infty} \frac{1}{j!}\,(| tA |)^j \leq \sum_{j=n+1}^{\infty} \frac{1}{j!}\,R^j$$

für jedes $R >| tA |$. Somit folgt aus der Konvergenz von $\exp(R)$, dass die Partialsummen P_n auf jedem kompakten Intervall von $\mathbb{R}$ gegen die Grenzfunktion $t \mapsto \exp(tA)$ gleichmässig konvergieren, was insbesondere die Stetigkeit der Funktion $t \mapsto \exp(tA)$ auf ganz $\mathbb{R}$ garantiert. Desweiteren erhalten wir aus der Funktionalgleichung:

$$\frac{1}{h}(\exp((t+h)\,A) - \exp(tA)) = \exp(tA)\,\frac{1}{h}(\exp(h\,A) - \mathbf{1}), \tag{3.23}$$

für beliebig fixiertes $t \in \mathbb{R}$ und jedes $h \in \mathbb{R} \setminus \{0\}$. Desweiteren gilt:

$$\frac{1}{h}(\exp(h\,A) - \mathbf{1}) - A = A^2 h \sum_{j=0}^{\infty} b_j (hA)^j$$

mit $b_j = \frac{1}{(j+2)!}$. Somit erhalten wir die Abschätzung

$$| \frac{1}{h}(\exp(h\,A) - \mathbf{1}) - A | \leq | A |^2 | h | \sum_{j=0}^{\infty} b_j\,(| hA |^j). \tag{3.24}$$

Da die Potenzreihe $\sum_{j=0}^{\infty} b_j\, z^j$ wegen $\limsup_{j\to\infty} \sqrt[j]{b_j} = \limsup_{j\to\infty} \sqrt[j]{1/j!} = 0$ auf jedem Kompaktum von $\mathbb{C}$ gegen eine stetige Funktion $P(z)$ konvergiert, die insbesondere $\lim_{z\to 0} P(z) = b_0 = 1/2$ erfüllt, folgt somit aus (3.24):

$$\exists \lim_{h\to 0} \frac{1}{h}(\exp(h\,A) - \mathbf{1}) \quad \text{und} \quad = A.$$

Zusammen mit (3.23) beweist dies, dass die Ableitung von $t \mapsto \exp(tA)$ in jedem $t \in \mathbb{R}$ existiert und durch $\exp(tA)\,A = A\,\exp(tA)$ gegeben ist. Schliesslich folgt (3.22) durch Iteration und insbesondere die unendlich häufige Differenzierbarkeit von $t \mapsto \exp(tA)$. □

Korollar 3.3.1. *Seien A, Y_0 beliebige quadratische Matrizen. So hat das Matrizenwertige AWP*

$$Y'(t) = A \cdot Y(t), \qquad Y(0) = Y_0 \tag{3.25}$$

die eindeutige Lösung $Y(t) = \exp(tA)\,Y_0$ auf ganz $\mathbb{R}$.

Proof. Nach Formel (3.22) ist $Y(t) = \exp(tA) \cdot Y_0$ eine auf ganz $\mathbb{R}$ existente Lösung des AWP's (3.25). Da das „Vektorfeld" $f : B \mapsto A \cdot B$

$$| f(B_1) - f(B_2) | \leq | A | \, | B_1 - B_2 |$$

für alle $B \in \mathrm{Mat}_{n,n}(\mathbb{R})$ erfüllt, folgt aus Theorem 2.0.4, dass dies die einzige Lösung sein muss. □

Bemerkung 3.3.1. *Man kann die Eindeutigkeit der Lösung $Y(t) = \exp(tA) \cdot Y_0$ auch durch Differentiation der Hilfsfunktion $\phi(t) := \exp(-t\,A)\,Y(t)$ für eine beliebige Lösung Y von (3.25) einsehen, denn nach Formel (3.22) zusammen mit der Produktregel für Differentiation erhalten wir:*

$$\phi'(t) = (-A\exp(-t\,A)\,Y + \exp(-t\,A)\,(A\,Y) = (-A + A)\,\exp(-t\,A)\,Y \equiv 0.$$

Somit folgt $\phi(t) \equiv \phi(0) = Y(0) = Y_0$ und daher mittels Punkt (2) von Proposition 3.3.1: $Y(t) = \exp(t\,A)\,\phi(t) = \exp(t\,A)\,Y_0$, für alle $t \in \mathbb{R}$.

Beispiel: Wir versuchen nun die durch $(2 \times 2)-$Systeme

$$\begin{pmatrix} \dot{X}_1 \\ \dot{X}_2 \end{pmatrix} = A \cdot \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \tag{3.26}$$

für konstante, diagonalisierbare $(2 \times 2)-$Matrizen

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

bewirkten „Strömungsbilder", d.h. die Verläufe aller Lösungskurven im $\mathbb{R}^2$ von (3.26), zumindest qualitativ zu erfassen und zu klassifizieren. Hierfür verwenden wir, dass eine Basis des Lösungsraums des Systems (3.26) durch die Spalten einer Fundamentalmatrix $Y : \mathbb{R} \to \mathrm{Mat}_{2,2}(\mathbb{R})$ gegeben wird und ein solches Y nach Proposition 3.1.2 dadurch charakterisiert wird, dass Y die Matrix-Differentialgleichung

$$Y'(t) = A \cdot Y(t), \qquad \text{für alle} \qquad t \in \mathbb{R} \tag{3.27}$$

löst und $Y(0)$ invertierbar ist. Wählen wir nun eine Startmatrix $Y(0)$ beliebig, so erhalten wir aus Korollar 3.3.1, dass $Y(t) = \exp(tA)\,Y(0)$ die eindeutige Lösung des zu (3.27) gehörenden AWP's (3.25) zu dieser Startmatrix $Y(0)$ liefert, also dass die beiden Spalten von $\exp(tA)\,Y(0)$ eine Basis des Lösungsraums des $(2\times 2)-$Systems (3.26) bilden. Da die explizite Berechnung von $\exp(tA)$ für eine beliebige (diagonalisierbare) (2×2)-Matrix A extrem aufwendig sein kann, nutzen wir nun die vorausgesetzte Diagonalisierbarkeit von A, also dass eine invertierbare Matrix B existiert, für die $BAB^{-1} = \mathrm{diag}(\lambda_1, \lambda_2)$, mit zwei reellen Eigenwerten λ_1, λ_2 von A, gilt. Multiplikation von (3.27) mit B liefert nun

$$(B \cdot Y)'(t) = B \cdot A \cdot Y(t) = B \cdot A \cdot B^{-1} \cdot (B \cdot Y)(t) = \mathrm{diag}(\lambda_1, \lambda_2)(B \cdot Y)(t), \tag{3.28}$$

sodass also die Matrix $\tilde{Y} := B \cdot Y$ die eindeutige Fundamentalmatrix des vereinfachten $(2 \times 2)-$Systems

$$\dot{X} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \cdot X \tag{3.29}$$

liefert, die in $t = 0$ in der invertierbaren Matrix $\tilde{Y}(0) = B \cdot Y(0)$ startet, und somit nach Korollar 3.3.1 explizit durch

$$\tilde{Y}(t) = \exp(t\,\mathrm{diag}(\lambda_1, \lambda_2)) \cdot (B \cdot Y(0)) = \begin{pmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{pmatrix} \cdot (B \cdot Y(0)) \qquad \text{für alle} \qquad t \in \mathbb{R}$$

gegeben wird, falls $Y(t) = \exp(tA)\,Y(0)$ die obige Fundamentalmatrix des ursprünglichen $(2 \times 2)-$Systems mit Startmatrix $Y(0)$ ist. Wählen wir nun

$$Y(0) := B^{-1} \cdot \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

so sind die Spalten von $\tilde{Y}$ somit durch

$$\begin{pmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{pmatrix} \qquad \text{und} \qquad \begin{pmatrix} -e^{\lambda_1 t} \\ e^{\lambda_2 t} \end{pmatrix} \tag{3.30}$$

gegeben und bilden eine Basis des Lösungsraums $\text{Lös}(\text{diag}(\lambda_1, \lambda_2))$. Wegen $\tilde{Y} = B \cdot Y$ ist nun

$$Y(t) = B^{-1} \cdot \tilde{Y}(t) = B^{-1} \cdot \begin{pmatrix} e^{\lambda_1 t} & -e^{\lambda_1 t} \\ e^{\lambda_2 t} & e^{\lambda_2 t} \end{pmatrix} \tag{3.31}$$

die (eindeutige) Fundamentalmatrix des ursprünglichen $(2 \times 2)-$Systems (3.27), deren erste Spalte Y_1 die Lösungskurve dieses Systems zum Startvektor $Y_1(0) = B^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und deren zweite Spalte Y_2 die Lösungskurve dieses Systems zum Startvektor $Y_2(0) = B^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ ist. Hiermit können wir nun alle Lösungskurven von (3.27) bei konkreter Vorgabe von A explizit berechnen. Wir unterscheiden nun hierzu die folgenden Fälle:

<u>Fall 1</u>: $0 < \lambda_1 < \lambda_2$. In diesem Fall werden gemäss (3.30) beide Komponenten der beiden Startvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ exponentiell beschleunigt, die erste Komponente jedoch langsamer als die zweite Komponente, sodass die beiden entsprechenden Lösungskurven des vereinfachten Systems (3.29) für $t \to \infty$ parabelförmig in die Ferne entschwinden und für $t \to -\infty$ parabelförmig in den Ursprung fliessen. Nur falls ein Startvektor auf der „x-" oder „y-Achse" liegt, so bewegt sich die in diesem startende Lösungskurve von (3.29) für $t \to \infty$ entlang der jeweiligen Achse, also linear, in unendliche Ferne.

<u>Fall 2</u>: $0 < \lambda_2 < \lambda_1$. In diesem Fall werden ebenfalls gemäss (3.30) beide Komponenten der beiden Startvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ exponentiell beschleunigt, die erste Komponente nun jedoch schneller als die zweite Komponente, sodass die beiden entsprechenden Lösungskurven des vereinfachten Systems (3.29) für $t \to \infty$ (nur qualitativ) ähnlich wie Graphen von Wurzelfunktionen $x \mapsto x^r$, mit $r \in (0, 1)$, in die Ferne entschwinden und für $t \to -\infty$ in den Ursprung fliessen. Nur falls ein Startvektor exakt auf der „x-" oder „y-Achse" liegt, bewegt sich die in diesem startende Lösungskurve von (3.29) für $t \to \infty$ entlang der jeweiligen Achse, also linear, in unendliche Ferne.

<u>Fall 3</u>: $0 = \lambda_1,\ \lambda_2 > 0$. In diesem Fall wird gemäss (3.30) (wegen $e^0 = 1$) die erste Komponente der beiden Startvektoren $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ nicht bewegt, jedoch wird deren

zweite Komponente exponentiell beschleunigt. Somit strömen die beiden entsprechenden Lösungskurven von (3.29) für $t \to \infty$ parallel zur „y-Achse" nach oben und fallen senkrecht auf die „x-Achse" für $t \to -\infty$. Nur Startvektoren, die exakt auf der „x-Achse" liegen, werden durch die Wirkung des Beschleunigungsfeldes (3.30) nicht bewegt, bleiben also für alle $t \in (-\infty, \infty)$ einfach an ihrem Ort sitzen.

<u>Fall 4</u>: $0 = \lambda_2$, $\lambda_1 > 0$. In diesem Fall wird gemäss (3.30) (wegen $e^0 = 1$) nun die zweite Komponente der beiden Startvektoren $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}$ nicht bewegt, jedoch wird deren erste Komponente exponentiell beschleunigt. Somit strömen die beiden entsprechenden Lösungskurven von (3.29) für $t \to \infty$ parallel zur x-Achse nach rechts bzw. nach links und pirschen sich von rechts bzw. von links horizontal an die „y-Achse" heran, während $t \to -\infty$ strebt. Nur Startvektoren, die auf der „y-Achse" liegen, werden durch die Wirkung des Beschleunigungsfeldes (3.30) nicht bewegt, bleiben also für alle $t \in (-\infty, \infty)$ einfach an ihrem Ort sitzen.

<u>Fall 5</u>: $\lambda_1 < 0 < \lambda_2$. Dies ist der interessanteste Fall. Hier wird gemäss (3.30) der Betrag der ersten Komponente der beiden Startvektoren $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}$ exponentiell verringert, während deren zweite Komponente exponentiell beschleunigt wird, für wachsende $t-$Werte. Somit pirschen sich die beiden entsprechenden Lösungskurven von (3.29) für $t \to \infty$ von rechts bzw. von links an die y-Achse heran und gleiten für $t \to -\infty$ langsam von oben auf die x-Achse herab. Die in $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}$ startenden Lösungskurven entschwinden also für $t \to \infty$ entlang der y-Achse in unendlicher Ferne. Nur eine auf der x-Achse startende Lösungskurve wird von rechts oder von links in den Ursprung $(0, 0)$ hineingesogen.

<u>Fall 6</u>: $\lambda_1 = 0 = \lambda_2$. In diesem Fall werden beide Startvektoren $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}$ gemäss (3.30) nicht bewegt, sodass also überhaupt kein Strömungsbild entsteht.

Die Fälle (7): $[0 > \lambda_1 > \lambda_2]$, (8): $[0 > \lambda_2 > \lambda_1]$, (9): $[0 = \lambda_1,\ \lambda_2 < 0]$, (10) $[0 = \lambda_2,\ \lambda_1 < 0]$ und (11) $[\lambda_2 < 0 < \lambda_1]$ ergeben exakt dieselben Lösungskurven wie die Fälle (1)-(5), werden nur in entgegengesetzter Richtung durchlaufen. Kehrt man also in den Fällen (7)-(11) die Zeit um, ersetzt also t durch $-t$, so ergeben sich exakt dieselben Strömungsbilder wie in den Fällen (1)-(5).
Ist nun eine diagonalisierbare $(2 \times 2)-$Matrix A vorgelegt, so diagonalisiere man diese mittels Konjugation mit einer GL_2-Matrix B, entscheide anhand der Eigenwerte von A welcher obige Fall vorliegt und wende dann wegen $\tilde{Y} = B \cdot Y$ auf das im adäquaten Fall beschriebene Strömungsbild die Matrix B^{-1} an, siehe Formel (3.31), um das Strömungsbild des durch A gegebenen Systems (3.26) zu erhalten. Man gewinnt also in der Tat auf diese Weise alle qualitativ unterschiedlichen Strömungsbilder zu $(2 \times 2)-$Systemen (3.26) zu beliebig vorgelegten diagonalisierbaren $(2 \times 2)-$Matrizen A.

4 Stetige und stetig differenzierbare Abhängigkeit von Lösungen von ihren Anfangswerten

Definition 4.0.2. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen. Wir nennen eine Differentialgleichung $y' = f(\,\cdot\,, y)$,,eindeutig lösbar als Anfangswertproblem", wenn zwei Lösungen $y_1, y_2 : I_i \to \mathbb{R}^n$, $i = 1, 2$, von $y' = f(\,\cdot\,, y)$ die in einem Punkt $t_0 \in I_1 \cap I_2$ übereinstimmen, bereits auf $I_1 \cap I_2$ übereinstimmen, also falls aus der Existenz eines $t_0 \in I_1 \cap I_2$ mit $y_1(t_0) = y_2(t_0)$ bereits $y_1 \equiv y_2$ auf $I_1 \cap I_2$ folgt. In diesem Fall existiert also für ein fixiertes Paar $(\tau, x) \in \Omega$ eine eindeutige, maximale Lösung $y_{\tau,x} : I_{(\tau,x)} \longrightarrow \mathbb{R}^n$ des AWP's*

$$y' = f(\,\cdot\,, y), \qquad y(\tau) = x, \tag{4.1}$$

und wir definieren die sogenannte Lösungsfunktion $\phi \equiv \phi_f : \mathcal{D} := \{(t, \tau, x) \in I_{\tau,x} \times \Omega\} \to \mathbb{R}^n$ durch $\phi(\,\cdot\,, \tau, x) := y_{\tau,x}(\cdot)$. Wir setzen ausserdem $\phi'(t, \tau, x) := \frac{\partial}{\partial t}\phi(t, \tau, x)$.

Theorem 4.0.1. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, und die Differentialgleichung*

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

sei eindeutig als Anfangswertproblem lösbar. Dann ist die Lösungsfunktion ϕ und ihre partielle Zeit-Ableitung $\phi' \equiv \frac{\partial \phi}{\partial t}$ stetig und deren Definitionsbereich $\mathcal{D}$ offen.

Proof. Wir fixieren ein Tripel $(t, \tau, x) \in \mathcal{D}$, d.h. $(\tau, x) \in \Omega$ und $t \in I_{\tau,x}$ und betrachten zu (t, τ, x) die maximale Lösung $y_{\tau,x} : I_{\tau,x} \to \mathbb{R}^n$ von (4.1). Nach Theorem 2.0.6 ist das Definitionsintervall $I_{\tau,x}$ der maximalen Lösung $y_{\tau,x}$ offen. Somit können wir zwei Zeitpunkte $a < b \in I_{\tau,x}$ mit $a < \tau$ und $t < b$ wählen. Anhand der Stetigkeit von $y_{\tau,x}$ ist die Menge

$$K := \{(s, y_{\tau,x}(s)) \mid s \in [a, b]\}$$

kompakt. Es existiert somit ein $\epsilon > 0$, sodass die Umgebung

$$U_{5\epsilon}(K) := \{(s, y) \in \mathbb{R} \times \mathbb{R}^n \mid \mathrm{dist}((s, y), K) < 5\epsilon\}$$

von K noch in Ω enthalten ist, und wir definieren die Funktion

$$\varphi(s, y) := \max\{0, \min\{1, 2 - (2\epsilon)^{-1} d((s, y), K)\}\}$$

von $\mathbb{R} \times \mathbb{R}^n$ nach $\mathbb{R}_+$. φ ist stetig und erfüllt $\varphi \equiv 1$ auf $U_{2\epsilon}(K)$ und $\varphi \equiv 0$ auf $\mathbb{R} \times \mathbb{R}^n \setminus U_{4\epsilon}(K)$. Nun definieren wir das Vektorfeld $g : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ durch $g := f\,\varphi$ auf $U_{5\epsilon}(K)$ und durch $g \equiv 0$ auf $\mathbb{R} \times \mathbb{R}^n \setminus \overline{U_{4\epsilon}(K)}$ und sehen, dass g anhand der Stetigkeit von f

beschränkt ist, also dass $| g |\leq M$ auf $\mathbb{R} \times \mathbb{R}^n$ für ein $M > 0$ gilt, und auf $U_{2\epsilon}(K)$ mit f übereinstimmt. Wir betrachten nun eine Folge von Tripeln $\{(t_m, \tau_m, x_m)\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$, die gegen ein Tripel (t, τ, x) konvergiert. Nach Theorem 2.0.1 und Theorem 2.0.6 existieren maximale Lösungen z_m der AWP's

$$y' = g(\,\cdot\,, y), \qquad y(\tau_m) = x_m,$$

deren Existenzintervall nach Theorem 2.0.7 anhand der Beschränktheit von g ganz $\mathbb{R}$ sein muss. Wir wählen zwei beliebige Zahlen $c < d \in \mathbb{R}$ mit $t_m, t, \tau_m, \tau \in [c, d]$ und erhalten aus $| z_m' |=| g(\,\cdot\,, z_m) |\leq M$ zusammen mit dem Hauptsatz der Differential- und Integralrechnung:

$$| z_m(s) |\leq| z_m(\tau_m) | +M \mid s - \tau_m |\leq| x_m | + M(d-c)$$

und genauso

$$| z_m(s) - z_m(s') |\leq \int_{s'}^{s} | g(r, z_m(r)) | \; dr \leq M \mid s - s' | \qquad \text{für beliebige} \quad s, s' \in [c, d].$$

Wegen $x_m \to x$ folgt hieraus sowohl die gleichmässige Beschränktheit als auch die gleichgradige Stetigkeit der Folge $\{z_m\}$, sodass der Satz von Arzela und Ascoli die Existenz einer Teilfolge $\{z_{m_l}\}$ und einer auf $[c, d]$ stetigen Funktion z mit

$$\sup_{[c,d]} | z_{m_l} - z |\to 0, \tag{4.2}$$

für $l \to \infty$ garantiert. Anhand der gleichmässigen Stetigkeit von g auf ganz $\mathbb{R} \times \mathbb{R}^n$ folgt hieraus die ebenfalls gleichmässige Konvergenz

$$g(\cdot, z_{m_l}) \to g(\cdot, z) \qquad \text{auf} \qquad [c, d] \tag{4.3}$$

für $l \to \infty$. Da wir das Intervall $[c, d]$ beliebig gross wählen dürfen, gelten die Konvergenzen in (4.2) und (4.3) somit auf jedem kompakten Intervall aus $\mathbb{R}$, sodass uns der Hauptsatz der Differential- und Integralrechnung und der Konvergenzsatz von Lebesgue zusammen mit $(\tau_m, x_m) \to (\tau, x)$ liefern:

$$z(s) \longleftarrow z_{m_l}(s) = x_{m_l} + \int_{\tau_{m_l}}^{s} g(r, z_{m_l}(r))\, dr \longrightarrow x + \int_{\tau}^{s} g(r, z(r))\, dr$$

also $z(s) = x + \int_\tau^s g(r, z(r))\, dr$, für beliebige $s \in \mathbb{R}$, was mit Proposition 2.0.1 zu $[z'(s) = g(s, z(s))$ mit $z(\tau) = x]$ äquivalent ist. Nun zeigen wir, dass diese Funktion z in der Tat die eindeutige Lösung des AWP's (4.1) auf $[a, b]$ ist, also dass

$$z = y_{\tau,x} \qquad \text{auf} \quad [a, b] \tag{4.4}$$

gilt. Hierfür zeigen wir zuerst, dass

$$(s, z(s)) \in U_\epsilon(K) \tag{4.5}$$

für alle $s \in [a,b]$ gilt. Wegen $z(\tau) = x = y_{\tau,x}(\tau)$ erhalten wir zunächst, dass $(\tau, z(\tau)) = (\tau, x) \in K$ per Definition von K gilt, sodass (4.5) für $s = \tau$ gilt. Nun nehmen wir an, dass ein $s \in (\tau, b]$ existiert, für welches (4.5) nicht gilt. Wir setzen

$$s_0 := \inf\{s \in [\tau, b] \mid (s, z(s)) \notin U_\epsilon(K)\}$$

und sehen anhand der Stetigkeit von z und $(\tau, z(\tau)) \in K$:

$$\begin{aligned} \tau < s_0 \leq b, \qquad (s, z(s)) \in U_\epsilon(K) \qquad &\text{für} \qquad s \in [\tau, s_0) \\ \text{und} \qquad &\operatorname{dist}((s_0, z(s_0)), K) = \epsilon. \end{aligned}$$

Wegen $f = g$ in $U_\epsilon(K)$ erhalten wir hieraus: $z'(s) = g(s, z(s)) = f(s, z(s))$ für $s \in [\tau, s_0)$, und z löst somit wegen $z(\tau) = x$ das AWP (4.1) zumindest auf dem Teilintervall $[\tau, s_0)$ von $[a,b]$. Da wir voraussetzen, dass die Differentialgleichung $y' = f(\,\cdot\,, y)$ eindeutig lösbar ist, so folgt hieraus bereits die Behauptung (4.4) auf $[\tau, s_0)$, also $z = y_{\tau,x}$ auf $[\tau, s_0)$. Somit erhalten wir aus der Stetigkeit von z und von $y_{\tau,x}$, dass $(s_0, z(s_0)) = (s_0, y_{\tau,x}(s_0))$ und somit $(s_0, z(s_0)) \in K$ per Definition von K gelten muss, was jedoch $\operatorname{dist}((s_0, z(s_0)), K) = \epsilon > 0$ widerspricht. Also gilt (4.5) auf $[\tau, b]$. Analog erhält man (4.5) auf $[a, \tau]$, indem man die Existenz eines $s \in [a, \tau)$ annimmt, für welches (4.5) nicht gilt, betrachtet

$$\tilde{s}_0 := \sup\{s \in [a, \tau] \mid (s, z(s)) \notin U_\epsilon(K)\}$$

und gelangt wieder zum Widerspruch zwischen $\operatorname{dist}((\tilde{s}_0, z(\tilde{s}_0)), K) = \epsilon > 0$ und $(\tilde{s}_0, z(\tilde{s}_0)) \in K$. Somit folgt (4.5) auf $[a,b]$, und wie oben erhalten wir aus Kombination hieraus mit $f = g$ in $U_\epsilon(K)$, dass $z'(s) = f(s, z(s))$ für $s \in [a,b]$ zusammen mit $z(\tau) = x$ gilt, was $z = y_{\tau,x}$ auf $[a,b]$, also Behauptung (4.4) zur Folge hat.
Desweiteren folgt aus der gleichmässigen Konvergenz der Funktionen z_{m_l} gegen z, dass

$$| \, z_{m_l}(s) - z(s) \, | < \epsilon \qquad \text{für} \qquad s \in [a,b]$$

und somit wegen (4.4) insbesondere $(s, z_{m_l}(s)) \in U_\epsilon(K)$ für alle $s \in [a,b]$, für hinreichend grosse l gilt. Somit erfüllen wegen $f = g$ auf $U_\epsilon(K)$ die Funktionen z_{m_l} die Differentialgleichungen $z'_{m_l}(t) = f(t, z_{m_l}(t))$, für $t \in [a,b]$. Wegen $t, \tau \in (a,b)$ und $(t_m, \tau_m) \to (t, \tau)$ folgt $t_{m_l}, \tau_{m_l} \in (a,b)$, und zusammen mit $z_m(\tau_m) = x_m$ erhalten wir anhand der eindeutigen Lösbarkeit der Gleichung $y' = f(\,\cdot\,, y)$, dass

$$z_{m_l}(s) = y_{\tau_{m_l}, x_{m_l}}(s) \qquad \forall s \in [a,b]$$

für grosse l gilt, was insbesondere zeigt, dass das Intervall $[a,b]$ in den Existenz-Intervallen $I_{\tau_{m_l}, x_{m_l}}$ der maximalen Lösungen $y_{\tau_{m_l}, x_{m_l}}$ der Gleichung $y' = f(\,\cdot\,, y)$ zu den Anfangswerten „$y(\tau_{m_l}) = x_{m_l}$" enthalten sein muss, falls l hinreichend gross ist. Insbesondere gilt also $t_{m_l} \in [a,b] \subset I_{\tau_{m_l}, x_{m_l}}$ und auch $(\tau_{m_l}, x_{m_l}) \in \Omega$, also

$$(t_{m_l}, \tau_{m_l}, x_{m_l}) \in \mathcal{D} \qquad \text{für hinreichend grosse } l. \tag{4.6}$$

Desweiteren folgt aus der gleichmässigen Konvergenz der z_{m_l} gegen die Funktion $z = y_{\tau,x}$ auf $[a,b]$ zusammen mit $t_m \to t$ anhand der Definition der Lösungsfunktion ϕ:

$$\phi(t_{m_l}, \tau_{m_l}, x_{m_l}) = y_{\tau_{m_l}, x_{m_l}}(t_{m_l}) = z_{m_l}(t_{m_l}) \to z(t) = y_{\tau,x}(t) = \phi(t, \tau, x).$$

Da der ermittelte Limes $y_{\tau,x}(t)$ unabhängig von der oben ausgewählten Teilfolge $\{m_l\}$ aller Indizes m ist, folgt $\phi(t_m, \tau_m, x_m) \to \phi(t, \tau, x)$, also die behauptete Stetigkeit von ϕ auf $\mathcal{D}$. Wegen

$$\frac{\partial \phi}{\partial t}(t, \tau, x) = (y_{\tau,x})'(t) = f(t, y_{\tau,x}(t)) = f(t, \phi(t, \tau, x))$$

erhalten wir hieraus auch die Stetigkeit von $\frac{\partial \phi}{\partial t}$ auf $\mathcal{D}$. Schliesslich beweisen wir die Offenheit von $\mathcal{D}$. Angenommen, $\mathcal{D}$ wäre nicht offen, so existierte ein Tripel $(t, \tau, x) \in \mathcal{D}$ und eine Folge $\{(t_m, \tau_m, x_m)\} \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n$ mit $(t_m, \tau_m, x_m) \to (t, \tau, x)$ und $(t_m, \tau_m, x_m) \notin \mathcal{D}$ für alle $m \in \mathbb{N}$. Oben bewiesen wir jedoch (siehe (4.6)), dass aus einer beliebigen, gegen ein fixiertes Tripel (t, τ, x) konvergenten Folge die Existenz einer Teilfolge $\{t_{m_l}, \tau_{m_l}, x_{m_l}\}$ mit $(t_{m_l}, \tau_{m_l}, x_{m_l}) \in \mathcal{D}$ für hinreichend grosse l folgt. Da dies ein Widerspruch ist, folgt die Behauptung. $\square$

Bemerkung 4.0.2. *Wir werden Differentialgleichungen betrachten, deren Vektorfeld f von einem (vektorwertigen) „Parameter" $\lambda \in \mathbb{R}^k$ abhängt. Es sei also $f : \Omega \to \mathbb{R}^n$ stetig mit $\Omega \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^k$ offen, und für jeden fixierten Parameter λ sei die Differentialgleichung $y' = f(\,\cdot\,, y, \lambda)$ eindeutig als Anfangswertproblem lösbar. Somit erhalten wir eine wohldefinierte Lösungsfunktion ϕ_f, die jedem Tupel (t, τ, x, λ) den Ortsvektor $y_{\tau,x,\lambda}(t)$ der eindeutigen Lösung $y_{\tau,x,\lambda}$ des AWP's*

$$y' = f(\cdot, y, \lambda), \qquad y(\tau) = x \tag{4.7}$$

zum Zeitpunkt t zuordnet. Um die Resultate des soeben bewiesenen Theorems auf ϕ_f anwenden zu können, führen wir nun das stetige Vektorfeld $F : \Omega \to \mathbb{R}^n \times \mathbb{R}^k$ durch $F(s, w) := (f(s, w), 0) \in \mathbb{R}^n \times \mathbb{R}^k$ mit $w := (y, \lambda)$ ein. Das AWP (4.7) mit Parameter λ ist somit zum eindeutig lösbaren AWP

$$w' = F(\cdot, w), \qquad w(\tau) = (x, \lambda) \tag{4.8}$$

ohne Parameter äquivalent. Das Vektorfeld F besitzt anhand der Voraussetzung an f ebenfalls eine wohldefinierte Lösungsfunktion ϕ_F, die jedem Tupel $(t, \tau, x, \lambda) \in \mathcal{D}_F \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^k$ das Paar $\begin{pmatrix} y_{\tau,x,\lambda}(t) \\ \lambda \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^k$ zuordnet, wobei $y_{\tau,x,\lambda}$ die eindeutige Lösung des AWP's (4.7) mit fixiertem Parameter λ ist. Somit sehen wir, dass die Lösungsfunktion $\phi \equiv \phi_f$ zu f gerade durch die ersten n Komponenten von ϕ_F gegeben wird, also dass $(\phi(t, \tau, x, \lambda))_i = (\phi_F(t, \tau, x, \lambda))_i$ für $i = 1, \ldots, n$ gilt, da in der Tat $\phi(t, \tau, x, \lambda) = y_{\tau,x,\lambda}(t)$ ist. Da nach Theorem 4.0.1 der Definitionsbereich $\mathcal{D}_F$ von ϕ_F offen und ϕ_F zusammen mit ϕ_F' stetig ist, erhalten wir somit automatisch die Offenheit des Definintionsbereichs $\mathcal{D} \equiv \mathcal{D}_f$ von ϕ_f in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^k$ und die Stetigkeit der ursprünglichen Lösungsfunktion ϕ_f und von ϕ_f' auf ihrem Definintionsbereich $\mathcal{D}$.

Theorem 4.0.2. *Es sei $f : \Omega \longrightarrow \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen, und die partielle Ableitung $\partial_x f : \Omega \to \mathrm{Mat}_{n,n}(\mathbb{R})$ existiere und sei stetig. Weiter sei $y : I \times [-\epsilon, \epsilon] \to \mathbb{R}^n$ eine Schar von Funktionen, sodass für jedes $h \in [-\epsilon, \epsilon]$ die Funktion $t \mapsto y(t, h)$ eine*

Lösung der Differentialgleichung $y' = f(\cdot, y)$ *ist, und für ein* $\tau \in I$ *existiere die partielle Ableitung* $\partial_h y(\tau, 0)$*. Dann existiert* $\partial_h y(t, 0)$ *für alle* $t \in I$ *und gehorcht dem linearen System von* n *Differentialgleichungen*

$$\partial_t(\partial_h y)(t, 0) = \partial_x f(t, y(t, 0)) \cdot \partial_h y(t, 0) \qquad \text{für alle } t \in I. \tag{4.9}$$

Proof. Wir fixieren zunächst ein Intervall $[a, b] \subset I$ beliebig und betrachten Zeitpunkte $t, \tau \in [a, b]$. Da wir voraussetzen, dass die partiellen Ableitungen von f nach x_i auf Ω existieren und stetig sind, ist f lokal Lipschitz-stetig in x, d.h. zu jedem Punkt $(t_0, x_0) \in \Omega$ existiert eine Umgebung $B_\rho(t_0, x_0) \subset \Omega$ und eine Konstante $L_\rho \in [0, \infty)$, sodass

$$\mid f(t, x_1) - f(t, x_2) \mid \leq L_\rho \mid x_1 - x_2 \mid \tag{4.10}$$

für beliebige Paare $(t, x_1), (t, x_2) \in B_\rho(t_0, x_0)$ gilt. Somit folgt aus Theorem 2.0.4 oder auch Theorem 2.0.5, dass die Differentialgleichung $y' = f(\cdot, y)$ eindeutig als Anfangswertproblem lösbar ist. Somit existiert die Lösungsfunktion ϕ aus Definition 4.0.2 und erfüllt

$$\phi(t, \tau, y(\tau, h)) = y(t, h) \qquad \forall\, t \in [a, b], \tag{4.11}$$

da $y(\cdot, h)$ die eindeutige Lösung der Differentialgleichung $y' = f(\cdot, y)$ ist, die zum Zeitpunkt τ den Wert $y(\tau, h)$ annimmt, was andererseits gerade die Definition von $\phi(\cdot, \tau, y(\tau, h))$ ist. Aus der Existenz der partiellen Ableitung $\partial_h y(\tau, 0) =: z_0$ folgt insbesondere $y(\tau, h) \to y(\tau, 0)$ für $h \to 0$. Ist nun $\{h_k\}$ eine beliebige Nullfolge, so folgt zunächst aus (4.11), dass das Produkt $[a, b] \times \{\tau\} \times \{y(\tau, h_k)\}$ im Definitionsgebiet $\mathcal{D}$ von ϕ enthalten ist und anhand der Konvergenz $y(\tau, h) \to y(\tau, 0)$ sogar in einem Kompaktum von $\mathcal{D}$ enthalten sein muss, also positiven Abstand von $\partial\mathcal{D}$ haben muss. Da die Lösungsfunktion ϕ nach Theorem 4.0.1 stetig ist, folgt deren gleichmässige Stetigkeit auf kompakten Teilmengen ihres Definitionsbereichs $\mathcal{D}$. Insgesamt folgt hieraus:

$$y(\cdot, h) = \phi(\cdot, \tau, y(\tau, h)) \to \phi(\cdot, \tau, y(\tau, 0)) = y(\cdot, 0)$$

gleichmässig auf $[a, b]$, für $h \to 0$. Wir definieren nun die Funktionen $z_h(t) := \frac{1}{h}(y(t, h) - y(t, 0))$ und sehen mittels Proposition 2.0.1 und dem verallgemeinerten Mittelwertsatz für jedes $h \neq 0$ mit hinreichend kleinem Betrag $\mid h \mid < \epsilon' < \epsilon$ und für jedes $t \in [a, b]$:

$$\begin{aligned} z_h(t) &= \frac{1}{h}(y(\tau, h) - y(\tau, 0)) + \int_\tau^t \frac{1}{h}\big(f(s, y(s, h)) - f(s, y(s, 0))\big)\, ds \\ &= \frac{1}{h}(y(\tau, h) - y(\tau, 0)) + \int_\tau^t \Big(\int_0^1 \partial_x f(s, \sigma y(s, h) + (1 - \sigma) y(s, 0))\, d\sigma\Big) z_h(s)\, ds \\ &\equiv z_h(\tau) + \int_\tau^t A_h(s) \cdot z_h(s)\, ds, \end{aligned} \tag{4.12}$$

wobei $A_h(s)$ das Integral $\int_0^1 \ldots d\sigma$ in der grossen Klammer bezeichne. Anhand der gleichmässigen Konvergenz $y(\cdot, h) \to y(\cdot, 0)$ auf $[a, b]$ und der gleichmässigen Stetigkeit von $\partial_x f$ auf kompakten Teilmengen von Ω erhalten wir, dass zu jedem $\delta > 0$

$$\mid A_h(s) - \partial_x f(s, y(s, 0)) \mid \leq \int_0^1 \mid \partial_x f(s, \sigma y(s, h) + (1 - \sigma) y(s, 0)) - \partial_x f(s, y(s, 0)) \mid \, d\sigma < \delta$$

für alle $s \in [a, b]$ gilt, falls $\mid h \mid$ hinreichend klein ist, d.h. dass A_h gegen $\partial_x f(\,\cdot\,, y(\,\cdot\,, 0))$ gleichmässig auf $[a, b]$ für $h \to 0$ konvergiert. Insbesondere existiert eine Konstante $M > 0$, für die $\mid A_h \mid \leq M$ auf $[a, b]$, für $\mid h \mid$ hinreichend klein, gilt, und somit folgt aus (4.12), also aus $z_h'(t) = A_h(t) \cdot z_h(t)$:

$$\mid z_h'(t) \mid \leq M \mid z_h(t) \mid \qquad \text{für alle } t \in [a, b].$$

Nach dem Lemma von Gronwall, Theorem 2.0.2, folgt hieraus:

$$\mid z_h(t) \mid \leq \mid z_h(\tau) \mid \exp(M \mid t - \tau \mid) \qquad \forall\, t \in [a, b].$$

Wegen $z_h(\tau) \to \partial_h y(\tau, 0) = z_0$, für $h \to 0$, folgt hieraus

$$\mid z_h(t) \mid \leq \Gamma \exp(M\,(b - a)) \qquad \forall\, t \in [a, b],$$

für ein $\Gamma > \mid z_0 \mid$, und zusammen mit $\mid z_h'(t) \mid \leq M \mid z_h(t) \mid$ schliesslich

$$\mid z_h'(t) \mid \leq M\,\Gamma \exp(M\,(b - a)) \qquad \forall\, t \in [a, b],$$

also mittels des Hauptsatzes der Diff.- und Integralrechnung:

$$\mid z_h(t) - z_h(s) \mid \leq M\,\Gamma \exp(M\,(b - a)) \mid t - s \mid \qquad \forall\, t, s \in [a, b].$$

Dies beweist, dass die Funktionen-Familie $\{z_h\}$ für kleine Werte von $\mid h \mid$ gleichgradig stetig und gleichmässig beschränkt auf $[a, b]$ ist, woraus mittels des Satzes von Arzela und Ascoli die Existenz einer Folge $\{z_{h_k}\}$, mit $h_k \to 0$, und einer auf $[a, b]$ stetigen Funktion z folgen, sodass $z_{h_k} \to z$ gleichmässig auf $[a, b]$ gilt. Aus (4.12) folgt hieraus zusammen mit $\frac{1}{h_k}(y(\tau, h_k) - y(\tau, 0)) \to \partial_h y(\tau, 0) = z_0$ und anhand der gleichmässigen Konvergenz der A_{h_k} gegen $\partial_x f(\,\cdot\,, y(\,\cdot\,, 0))$ auf $[a, b]$:

$$z(t) = z_0 + \int_\tau^t \partial_x f(s, y(s, 0)) \cdot z(s)\, ds, \qquad \text{für alle } t \in [a, b],$$

was nach Proposition 2.0.1 dazu äquivalent ist, dass z das lineare AWP

$$z'(t) = \partial_x f(t, y(t, 0)) \cdot z(t), \qquad z(\tau) = z_0 \tag{4.13}$$

auf $[a, b]$ löst. Da die Matrix $A(\,\cdot\,) := \partial_x f(\,\cdot\,, y(\cdot, 0))$ auf $[a, b]$ stetig ist, folgt aus Proposition 3.1.1, dass das AWP (4.13) auf $[a, b]$ keine weitere Lösung haben kann. Da die eindeutige Lösung dieses AWP's unabhängig von der obigen Auswahl der konvergenten Folge $\{z_{h_k}\}$ ist, folgt aus dem Teilfolgen-Prinzip die gleichmässige Konvergenz der gesamten Familie $z_h \to z$ auf $[a, b]$. Wegen $z_h(t) := \frac{1}{h}(y(t, h) - y(t, 0))$ impliziert dies, dass die partielle Ableitung $\partial_h y(\,\cdot\,, 0)$ auf ganz $[a, b]$ existiert und gerade die eindeutige Lösung z des AWP's (4.13) ist, also dass in der Tat $\partial_h y(\,\cdot\,, 0)$ dem linearen System (4.9) auf $[a, b]$ und somit auf ganz I gehorcht. $\square$

Theorem 4.0.3. *[Optimale Regularität der Lösungsfunktion, theoretische Version] Es sei $f : \Omega \to \mathbb{R}^n$ stetig, $\Omega \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ offen, bezeichne $(t, x, \mu) \mapsto f(t, x, \mu)$, und die partiellen Ableitungen $\partial_x f$ und $\partial_\mu f$ von f existieren und seien stetig. Dann besitzt die Lösungsfunktion ϕ der Differentialgleichung $y' = f(\,\cdot\,, y, \mu)$ mit Parameter μ zusammen mit ihrer Zeit-Ableitung ϕ' stetige partielle Ableitungen $\partial_x\phi$, $\partial_\mu\phi$, $\partial_\tau\phi$, $(\partial_x\phi)'$, $(\partial_\mu\phi)'$ und $(\partial_\tau\phi)'$ nach x, τ und μ, und diese erfüllen die Anfangswertprobleme*

$$\begin{aligned}
(\partial_x\phi)'(t, \tau, x, \mu) &= \partial_x f(t, \phi(t, \tau, x, \mu), \mu) \cdot \partial_x\phi(t, \tau, x, \mu), \\
\partial_x\phi(\tau, \tau, x, \mu) &= \mathbf{1}_n, \qquad (4.14) \\
(\partial_\mu\phi)'(t, \tau, x, \mu) &= \partial_x f(t, \phi(t, \tau, x, \mu), \mu) \cdot \partial_\mu\phi(t, \tau, x, \mu) \\
&\quad + \partial_\mu f(t, \phi(t, \tau, x, \mu), \mu), \\
\partial_\mu\phi(\tau, \tau, x, \mu) &= 0 \qquad (4.15) \\
&\text{und} \\
(\partial_\tau\phi)'(t, \tau, x, \mu) &= \partial_x f(t, \phi(t, \tau, x, \mu), \mu) \cdot \partial_\tau\phi(t, \tau, x, \mu) \\
\partial_\tau\phi(\tau, \tau, x, \mu) &= -f(\tau, x, \mu). \qquad (4.16)
\end{aligned}$$

Existieren die $k-$ten partiellen Ableitungen von f nach x und μ und sind stetig, so existieren die $(k-1)-$ten partiellen Ableitungen von $\partial_x\phi$, $\partial_\mu\phi$, $\partial_\tau\phi$, $(\partial_x\phi)'$, $(\partial_\mu\phi)'$ und $(\partial_\tau\phi)'$ nach x, τ und μ und sind stetig.

Proof. Wie in Bemerkung 4.0.2 erklärt wurde, führen wir zur korrekten Behandlung der Abhängigkeit von f vom Parameter μ das Vektorfeld $F(s, x, \mu) := (f(s, x, \mu), 0) \in \mathbb{R}^n \times \mathbb{R}^m$ ein, und betrachten anstatt des AWP's (4.7) mit Parameter μ das künstlich erweiterte AWP

$$\begin{pmatrix} y \\ \mu \end{pmatrix}' = F(\cdot, y, \mu), \qquad \begin{pmatrix} y \\ \mu \end{pmatrix}(\tau) = \begin{pmatrix} x \\ \mu \end{pmatrix} \qquad (4.17)$$

ohne Parameter. Da F nach Voraussetzung an f nach allen Komponenten des Paares (x, μ) stetig differenziert werden kann, erfüllt F eine lokale Lipschitz-Bedingung der Form (4.10) um jeden Punkt $(t, x, \mu) \in \Omega$ und besitzt somit eine wohldefinierte Lösungsfunktion ϕ_F. Wir fixieren nun ein Tripel $(\tau, x, \mu) \in \Omega$ beliebig, wählen ein Paar $a < b \in I_{\tau,x,\mu}$ mit $a < \tau < b$ und wählen ein $\epsilon > 0$, sodass

$$B_\epsilon([a, b] \times \{(\tau, x, \mu)\}) \subset \mathcal{D}$$

gilt, was anhand der Offenheit des Definitionsbereichs $\mathcal{D}$ der Lösungsfunktion $\phi = \phi_f$ von f bzw. des Definitionsbereichs $\mathcal{D}_F$ von ϕ_F und anhand der Kompaktheit von $[a, b]$ möglich ist. Ferner setzen wir für $i = 1, \ldots, n + m$ und $\mid h \mid < \epsilon$:

$$y_i(t, h) := \phi_F(t, \tau, (x, \mu) + h\, e_i) \qquad \text{für} \qquad t \in [a, b].$$

Es sind $y_i(t, h)$ Lösungen der AWP's

$$y' = F(\,\cdot\,, y), \qquad y(\tau) = \begin{pmatrix} x \\ \mu \end{pmatrix} + h\, e_i$$

$i = 1, \ldots, n+m$. Wir untersuchen zunächst die Fälle für $i = 1, \ldots, n$. Wegen $y_i(\tau, h) = \begin{pmatrix} x + h\, e_i \\ \mu \end{pmatrix}$ existieren die partiellen Ableitungen $\partial_h y_i(\tau, 0) = \begin{pmatrix} e_i \\ 0 \end{pmatrix}$. Somit folgt aus Theorem 4.0.2, dass die partiellen Ableitungen

$$z_i(t) := \partial_h y_i(t, 0) = \partial_{x_i} \phi_F(t, \tau, x, \mu)$$

in allen $t \in [a, b]$ existieren und die linearen AWP's

$$z_i'(t) = \partial_{(x,\mu)} F(t, y_i(t, 0)) \cdot z_i(t) = \partial_{(x,\mu)} F(t, \phi(t, \tau, x, \mu), \mu) \cdot z_i(t), \qquad z_i(\tau) = \begin{pmatrix} e_i \\ 0 \end{pmatrix} \tag{4.18}$$

auf $[a, b]$ lösen. Man beachte hierbei, dass wegen $F_i = f_i$, für $i = 1, \ldots, n$, und $F_j = 0$, für $j = n+1, \ldots, n+m$, $\partial_{(x,\mu)} F$ die Form

$$\partial_{(x,\mu)} F(t, x, \mu) = \begin{pmatrix} \partial_x f & \partial_\mu f \\ 0 & 0 \end{pmatrix} (t, x, \mu)$$

hat. Da die Lösungsfunktion ϕ zu f nach Theorem 4.0.1 stetig ist und die partiellen Ableitungen $\partial_x f$ und $\partial_\mu f$ von f nach Voraussetzung existieren und stetig sind, erweist sich das in z lineare Vektorfeld

$$g(s, z, \tau, x, \mu) := \partial_{(x,\mu)} F(s, \phi(s, \tau, x, \mu), \mu) \cdot z \tag{4.19}$$

$z \in \mathbb{R}^{n+m}$ nach $\mathbb{R}^{n+m}$ mit Parameter (τ, x, μ) als stetig in all seinen Variablen. Somit folgt aus Proposition 3.1.1, dass erstens die Differentialgleichung $z' = g(\,\cdot\,, z, \tau, x, \mu)$ eindeutig lösbar ist und zweitens zu beliebigen Anfangswerten $z(\sigma) = z_0 \in \mathbb{R}^{m+n}$, mit $\sigma \in [a, b]$ beliebig, eine eindeutige, maximale Lösung auf ganz $[a, b]$ besitzt. Somit besitzt das Vektorfeld g eine wohldefinierte und stetige Lösungsfunktion ϕ_g in den Variablen $(t, \sigma, z, \tilde{\tau}, \tilde{x}, \tilde{\mu}) \in [a, b] \times [a, b] \times \mathbb{R}^{m+n} \times B_\epsilon(\tau, x, \mu) \subset [a, b] \times [a, b] \times \mathbb{R}^{m+n} \times \mathbb{R}^{m+n+1}$ und mit Werten $\phi_g(t, \sigma, z, \tilde{\tau}, \tilde{x}, \tilde{\mu})$ im $\mathbb{R}^{n+m}$, wie in Bemerkung 4.0.2 erklärt wurde. Die eindeutigen, maximalen Lösungen z_i der AWP's

$$z_i' = g(\,\cdot\,, z_i, \tau, x, \mu), \qquad z_i(\tau) = \begin{pmatrix} e_i \\ 0 \end{pmatrix}$$

lassen sich mittels ϕ_g nun in der Form $z_i(t) = \phi_g(t, \tau, (e_i, 0), \tau, x, \mu)$ schreiben. Somit gilt per Definition der z_i's:

$$\partial_{x_i} \phi_F(t, \tau, x, \mu) = \phi_g(t, \tau, (e_i, 0), \tau, x, \mu)$$

für $i = 1, \ldots, n$ und für alle $t \in [a, b]$. Da die Lösungsfunktion ϕ_g nach Theorem 4.0.1 stetig ist, folgt hieraus die Stetigkeit der partiellen Ableitungen $\partial_{x_i} \phi_F$ auf $[a, b] \times B_\epsilon(\tau, x, \mu)$. Fassen wir ausserdem die AWP's (4.18), für $i = 1, \ldots, n$, spaltenweise zusammen, so erhalten wir zusammen mit $(\phi_F)_{n+j} = \mu_j$, $j = 1, \ldots, m$, dass die Jacobi-Matrix

$\partial_x\phi_F(t,\tau,x,\mu) = \begin{pmatrix}\partial_x\phi(t,\tau,x,\mu)\\ 0\end{pmatrix}$ der partiellen Ableitungen von ϕ_F nach den Komponenten x_i des Startvektors x der Matrix-wertigen Differentialgleichung

$$\begin{aligned}\begin{pmatrix}(\partial_x\phi)'(t,\tau,x,\mu)\\ 0\end{pmatrix} &= (\partial_x\phi_F)'(t,\tau,x,\mu)\\ &= \partial_{(x,\mu)}F(t,\phi(t,\tau,x,\mu),\mu)\cdot\partial_x\phi_F(t,\tau,x,\mu)\\ = \begin{pmatrix}\partial_x f & \partial_\mu f\\ 0 & 0\end{pmatrix}&(t,\phi(t,\tau,x,\mu),\mu)\cdot\begin{pmatrix}\partial_x\phi(t,\tau,x,\mu)\\ 0\end{pmatrix},\end{aligned}$$

genügt, also dass

$$(\partial_x\phi)'(t,\tau,x,\mu) = \partial_x f(t,\phi(t,\tau,x,\mu),\mu)\cdot\partial_x\phi(t,\tau,x,\mu)$$

auf $[a,b]$ mit $\partial_x\phi(\tau,\tau,x,\mu) = \mathbf{1}_n$ gilt, wie in (4.14) behauptet wurde. Insbesondere beweist dies die Stetigkeit der Zeit-Ableitung $(\partial_x\phi)'$ von $\partial_x\phi$ auf $[a,b]\times B_\epsilon(\tau,x,\mu)$. Nun untersuchen wir die Fälle für $i = n+1,\dots,n+m$. Wegen $y_{n+j}(\tau,h) = (x,\mu+h\,e_j)$, $j\in\{1,\dots,m\}$, existieren die partiellen Ableitungen $\partial_h y_{n+j}(\tau,0) = (0,e_j)$. Somit folgt aus Theorem 4.0.2, dass die partiellen Ableitungen

$$z_j(t) := \partial_h y_{n+j}(t,0) = \partial_{\mu_j}\phi_F(t,\tau,x,\mu)$$

in allen $t\in[a,b]$ existieren und die linearen AWP's

$$\begin{aligned}z_j'(t) &= \partial_{(x,\mu)}F(t,y_{n+j}(t,0))\cdot z_j(t) = \partial_{(x,\mu)}F(t,\phi(t,\tau,x,\mu),\mu)\cdot z_j(t)\\ &= \begin{pmatrix}\partial_x f & \partial_\mu f\\ 0 & 0\end{pmatrix}(t,\phi(t,\tau,x,\mu),\mu)\cdot z_j(t),\qquad z_j(\tau) = (0,e_j)\end{aligned}\tag{4.20}$$

auf $[a,b]$ lösen. Wie oben verwenden wir nun, dass das Vektorfeld $g(s,z,\tau,x,\mu) := \partial_{(x,\mu)}F(s,\phi(s,\tau,x,\mu),\mu)\cdot z$ aus (4.19) mit Parameter (τ,x,μ) linear in z und in allen Variablen stetig ist, sodass wegen Proposition 3.1.1 das Vektorfeld g eine wohldefinierte Lösungsfunktion ϕ_g auf $[a,b]\times[a,b]\times\mathbb{R}^{m+n}\times B_\epsilon(\tau,x,\mu)$ besitzt. Insbesondere die eindeutigen Lösungen z_j der AWP's

$$z_j' = g(\,\cdot\,,z_j,\tau,x,\mu),\qquad z_j(\tau) = (0,e_j)$$

lassen sich mittels ϕ_g in der Form $z_j(t) = \phi_g(t,\tau,(0,e_j),\tau,x,\mu)$ schreiben. Somit gilt per Definition der z_j's:

$$\partial_{\mu_j}\phi_F(t,\tau,x,\mu) = \phi_g(t,\tau,(0,e_j),\tau,x,\mu)$$

für $j = 1,\dots,m$ und für alle $t\in[a,b]$. Da die Lösungsfunktion ϕ_g nach Theorem 4.0.1 stetig ist, folgt hieraus die Stetigkeit der partiellen Ableitungen $\partial_{\mu_j}\phi_F$ auf $[a,b]\times B_\epsilon(\tau,x,\mu)$, also insbesondere die Stetigkeit der partiellen Ableitungen $\partial_{\mu_j}\phi$ auf $[a,b]\times B_\epsilon(\tau,x,\mu)$. Fassen wir nun wieder die AWP's (4.20), für $j = 1,\dots,m$, spaltenweise zusammen, so erhalten wir zusammen mit $(\phi_F)_{n+j} = \mu_j$, $j = 1, \ldots, m$, dass die Jacobi-Matrix $\partial_\mu \phi_F(t, \tau, x, \mu) = \begin{pmatrix} \partial_\mu \phi(t, \tau, x, \mu) \\ \mathbf{1}_m \end{pmatrix}$ der partiellen Ableitungen von ϕ_F nach den Komponenten μ_j des Parameters μ der Matrix-wertigen Differentialgleichung

$$\frac{d}{dt} \begin{pmatrix} \partial_\mu \phi(t, \tau, x, \mu) \\ \mathbf{1}_m \end{pmatrix} = (\partial_\mu \phi_F)'(t, \tau, x, \mu)$$
$$= \begin{pmatrix} \partial_x f & \partial_\mu f \\ 0 & 0 \end{pmatrix} (t, \phi(t, \tau, x, \mu), \mu) \cdot \begin{pmatrix} \partial_\mu \phi(t, \tau, x, \mu) \\ \mathbf{1}_m \end{pmatrix},$$

genügt, also dass

$$(\partial_\mu \phi)'(t, \tau, x, \mu) = \partial_x f(t, \phi(t, \tau, x, \mu), \mu) \cdot \partial_\mu \phi(t, \tau, x, \mu) + \partial_\mu f(t, \phi(t, \tau, x, \mu), \mu)$$

auf $[a, b]$ mit $\partial_\mu \phi(\tau, \tau, x, \mu) = 0$ gilt, wie in (4.14) behauptet wurde und woraus automatisch die Stetigkeit von $(\partial_\mu \phi)'$ auf $[a, b] \times B_\epsilon(\tau, x, \mu)$ folgt.
Zur Untersuchung der partiellen Ableitungen $\partial_\tau \phi$ und ihrer Zeit-Ableitung $(\partial_\tau \phi)'$ definieren wir die Funktionen

$$y(t, h) := \phi_F(t, \tau + h, x, \mu) = \begin{pmatrix} \phi(t, \tau + h, x, \mu) \\ \mu \end{pmatrix} \qquad \text{für } t \in [a, b],$$

für $\mid h \mid < \epsilon$, die anhand der Stetigkeit von ϕ_F in t und h stetig sind. Da ϕ_F die Lösungsfunktion zu F ist, lösen die Funktionen $y(t, h)$ die Differentialgleichung $y' = F(\,\cdot\,, y)$ und erfüllen ausserdem per Definition von ϕ_F:

$$y(\tau + h, h) = \phi_F(\tau + h, \tau + h, x, \mu) = \begin{pmatrix} x \\ \mu \end{pmatrix} = \phi_F(\tau, \tau, x, \mu) = y(\tau, 0).$$

Kombinieren wir dies mit dem Hauptsatz der Differential- und Integralrechnung und beachten wir die Stetigkeit von F (bzw. von f) und von $(t, h) \mapsto y(t, h)$, so erhalten wir

$$\frac{1}{h}(y(\tau, h) - y(\tau, 0) = \frac{1}{h}(y(\tau, h) - y(\tau + h, h)) = -\frac{1}{h} \int_\tau^{\tau + h} y'(s, h)\, ds$$
$$= -\frac{1}{h} \int_\tau^{\tau + h} F(s, y(s, h))\, ds \to -F(\tau, y(\tau, 0)) = -F(\tau, x, \mu).$$

Dies beweist die Existenz von $\partial_h y(\tau, 0) = \lim_{h \to 0} \frac{1}{h}(y(\tau, h) - y(\tau, 0))$ zusammen mit $\partial_h y(\tau, 0) = -F(\tau, x, \mu)$. Wenden wir also erneut Theorem 4.0.2 hier an, so erhalten wir, dass $\partial_h y(t, 0)$, also $\partial_\tau \phi_F(t, \tau, x, \mu)$, in allen $t \in [a, b]$ existiert und ausserdem das lineare AWP

$$z'(t) = \partial_{(x,\mu)} F(t, y(t, 0)) \cdot z(t) = \partial_{(x,\mu)} F(t, \phi(t, \tau, x, \mu), \mu) \cdot z(t), \qquad z(\tau) = -F(\tau, x, \mu) \tag{4.21}$$

auf $[a, b]$ löst, was wegen $\phi_F(t, \tau, x, \mu) = \begin{pmatrix} \phi(t, \tau, x, \mu) \\ \mu \end{pmatrix}$ und $F(t, x, \mu) = \begin{pmatrix} f(t, x, \mu) \\ 0 \end{pmatrix}$

$$(\partial_\tau \phi)'(t, \tau, x, \mu) = \partial_x f(t, \phi(t, \tau, x, \mu), \mu) \cdot \partial_\tau \phi(t, \tau, x, \mu) \quad \text{mit } \partial_\tau \phi(\tau, \tau, x, \mu) = -f(\tau, x, \mu) \tag{4.22}$$

impliziert, wie in (4.16) behauptet wurde. Wie oben können wir zum Vektorfeld $g(s, z, \tau, x, \mu) := \partial_{(x,\mu)} F(s, \phi(s, \tau, x, \mu), \mu) \cdot z$ mit Parameter (τ, x, μ) aus (4.19) die auf $[a, b] \times [a, b] \times \mathbb{R}^{m+n} \times B_\epsilon(\tau, x, \mu)$ wohldefinierte und stetige Lösungsfunktion ϕ_g verwenden, um mittels dieser die eindeutige Lösung $t \mapsto \partial_\tau \phi_F(t, \tau, x, \mu)$ des AWP's (4.21) in der Form

$$\partial_\tau \phi_F(t, \tau, x, \mu) = \phi_g(t, \tau, -F(\tau, x, \mu), \tau, x, \mu)$$

zu schreiben. Da ϕ_g nach Theorem 4.0.1 stetig ist und auch F auf $B_\epsilon(\tau, x, \mu)$ stetig ist, folgt die Stetigkeit der partiellen Ableitung $\partial_\tau \phi_F$, also die Stetigkeit seiner ersten n Komponenten $\partial_\tau \phi$ auf $[a, b] \times B_\epsilon(\tau, x, \mu)$. Da $\partial_\tau \phi$ die Differentialgleichung (4.22) löst und $\partial_x f$ stetig ist, haben wir hiermit automatisch auch die Stetigkeit der Zeit-Ableitung $(\partial_\tau \phi)'$ auf $[a, b] \times B_\epsilon(\tau, x, \mu)$ nachgewiesen.
Nun beweisen wir die zweite Behauptung des Theorems per Induktion über k. Die Induktionsbehauptung lautet also:
,,Ist f k-fach stetig differenzierbar nach x und μ, so existieren die $(k-1)$-ten partiellen Ableitungen von $\partial_x \phi$, $\partial_\tau \phi$, $\partial_\mu \phi$, $(\partial_x \phi)'$, $(\partial_\tau \phi)'$ und $(\partial_\mu \phi)'$ nach x, μ und τ in einem beliebig fixierten Punkt $(t, \tau, x, \mu) \in \mathcal{D}_f$ und sind in diesem Punkt stetig."
Aus dem soeben Bewiesenen folgt, dass aus der stetigen Differenzierbarkeit von f nach x und μ, also für $k = 1$, die Stetigkeit von $\partial_x \phi$, $\partial_\tau \phi$, $\partial_\mu \phi$, $(\partial_x \phi)'$, $(\partial_\tau \phi)'$ und $(\partial_\mu \phi)'$ folgt, was für $k = 1$ gerade die Induktionsbehauptung ist. Es sei nun $k \geq 1$ beliebig gegeben und f als k-fach stetig differenzierbar nach x und μ vorausgesetzt. Nach Induktionshypothese existieren somit die $(k-2)$-ten partiellen Ableitungen von $\partial_x \phi$, $\partial_\tau \phi$, $\partial_\mu \phi$, $(\partial_x \phi)'$, $(\partial_\tau \phi)'$, $(\partial_\mu \phi)'$ nach x, μ und τ und sind im beliebig fixierten Punkt $(t, \tau, x, \mu) \in \mathcal{D}_f$ stetig. Für das in z lineare und in allen Variablen stetige Vektorfeld

$$g(s, z, \tau, x, \mu) := \partial_{(x,\mu)} F(s, \phi(s, \tau, x, \mu), \mu) \cdot z, \tag{4.23}$$

aus (4.19) wissen wir bereits, dass es eine wohldefinierte Lösungsfunktion ϕ_g in den Variablen $(t, \sigma, z, \tilde{\tau}, \tilde{x}, \tilde{\mu}) \in \mathcal{D}_g \subset \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n+m} \times \mathbb{R}^{n+m+1}$ mit Werten $\phi_g(t, \sigma, z, \tilde{\tau}, \tilde{x}, \tilde{\mu})$ im $\mathbb{R}^{n+m}$ besitzt, dessen ersten n Komponenten $\tilde{\phi}_g := (\phi_g^1, \ldots, \phi_g^n)$

$$\partial_{x_i} \phi(t, \tau, x, \mu) = \tilde{\phi}_g(t, \tau, (e_i, 0), \tau, x, \mu), \qquad i = 1, \ldots, n \tag{4.24}$$
$$\partial_{\mu_j} \phi(t, \tau, x, \mu) = \tilde{\phi}_g(t, \tau, (0, e_j), \tau, x, \mu), \qquad j = 1, \ldots, m \tag{4.25}$$
$$\partial_\tau \phi(t, \tau, x, \mu) = \tilde{\phi}_g(t, \tau, -(f(\tau, x, \mu), 0), \tau, x, \mu) \tag{4.26}$$

erfüllen. Da wir soeben einsahen, dass ϕ nach x, μ und τ sogar $(k-1)$-fach stetig differenziert werden kann und $\partial_x f$ und $\partial_\mu f$ nach Induktionsvoraussetzung ebenfalls $(k-1)$-fach stetig nach x und μ differenziert werden können, erweist sich g anhand der Kettenregel als $(k-1)-$fach stetig differenzierbar nach z, τ, x und μ. Somit erfüllt g mit Parameter (τ, x, μ) (anstatt f mit Parameter μ) die Induktionsvoraussetzung für den Induktionsindex $k-1$, sodass die Induktionshypothese für die Lösungsfunktion ϕ_g von g, hier also mit z anstatt x, σ anstatt τ und $(\tilde{\tau}, \tilde{x}, \tilde{\mu})$ anstatt μ, ergibt, dass die $(k-2)$-ten partiellen Ableitungen von $\partial_z \phi_g$, $\partial_\sigma \phi_g$, $\partial_{\tilde{x}} \phi_g$, $\partial_{\tilde{\tau}} \phi_g$, $\partial_{\tilde{\mu}} \phi_g$, $(\partial_z \phi_g)'$, $(\partial_\sigma \phi_g)'$, $(\partial_{\tilde{x}} \phi_g)'$, $(\partial_{\tilde{\tau}} \phi_g)'$ und $(\partial_{\tilde{\mu}} \phi_g)'$ nach z, σ, $\tilde{x}$, $\tilde{\tau}$ und $\tilde{\mu}$ in einem beliebig fixierten Punkt $(t, \sigma, z, \tilde{\tau}, \tilde{x}, \tilde{\mu})$ aus dem Definitionsbereich $\mathcal{D}_g$ von ϕ_g existieren und stetig sind. In Kombination mit (4.24)-(4.26) zeigt dies anhand der Kettenregel, dass die $(k-1)$-ten partiellen Ableitungen von $\partial_x\phi$, $\partial_\mu\phi$ und $\partial_\tau\phi$ nach x, τ und μ auf $\mathcal{D}_f$ existieren und stetig sind. Da deren Zeit-Ableitungen $(\partial_x\phi)'$, $(\partial_\mu\phi)'$ und $(\partial_\tau\phi)'$ die AWP's (4.14), (4.15) und (4.16) lösen und $\partial_x f$ und $\partial_\mu f$ nach Induktionsvoraussetzung $(k-1)$-fach nach x und μ stetig partiell differenziert werden kann, folgt aus der soeben bewiesenen Existenz und Stetigkeit der $(k-1)$-ten partiellen Ableitungen von $\partial_x\phi$, $\partial_\mu\phi$ und $\partial_\tau\phi$ nach x, μ und τ ebenfalls die Existenz und Stetigkeit der $(k-1)$-ten partiellen Ableitungen von $(\partial_x\phi)'$, $(\partial_\mu\phi)'$ und $(\partial_\tau\phi)'$ nach x, τ und μ auf $\mathcal{D}_f$. □

Theorem 4.0.4. *[Optimale Regularität der Lösungsfunktion, praktische Version] Es sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : \Omega \to \mathbb{R}^n$, $(t,x,\mu) \mapsto f(t,x,\mu)$, ein Vektorfeld mit Parameter μ, welches in allen Variablen k-fach stetig differenzierbar sei, für ein $k > 0$. Dann ist dessen Lösungsfunktion ϕ, $(t,\tau,x,\mu) \mapsto \phi(t,\tau,x,\mu)$, wohldefiniert, und ϕ ist zusammen mit ihrer partiellen Zeit-Ableitung $\phi' = \partial_t\phi$ in allen Variablen k-fach stetig differenzierbar. Insbesondere existieren die doppelten partiellen Ableitungen $\partial_t\partial_{x_i}\phi$ auf $\mathcal{D}_f$, sind stetig und stimmen mit $\partial_{x_i}\partial_t\phi$ auf ganz $\mathcal{D}_f$ überein.*

Proof. Wir führen Induktion über $k > 0$. Da die Differentialgleichung bereits für $k = 1$ eindeutig lösbar ist, erhalten wir aus Theorem 4.0.1, dass die Lösungsfunktion ϕ zu f für $k = 1$ wohldefiniert und zusammen mit ihrer partiellen Zeitableitung ϕ' stetig ist. Desweiteren folgt aus Theorem 4.0.3, dass die partiellen Ableitungen $\partial_x\phi$, $\partial_\mu\phi$ und $\partial_\tau\phi$ existieren und stetig sind. Damit ist also ϕ in all ihren Variablen stetig partiell differenzierbar. Wegen

$$\frac{\partial\phi}{\partial t}(t,\tau,x,\mu) = (y_{\tau,x,\mu})'(t) = f(t, y_{\tau,x,\mu}(t), \mu) = f(t, \phi(t,\tau,x,\mu), \mu) \tag{4.27}$$

erhalten wir hieraus auch die Stetigkeit aller partiellen Ableitungen von $\phi' \equiv \frac{\partial\phi}{\partial t}$.
Nun sei ein $k > 1$ fest gewählt. Da f nach Induktionsvoraussetzung insbesondere $(k-1)$-fach stetig differenzierbar ist, folgt per Induktionshypothese, dass ϕ $(k-1)$-fach stetig differenzierbar in all ihren Variablen ist. Somit folgt, dass das Vektorfeld g aus (4.23) mit Parameter (τ,x,μ) $(k-1)$-fach stetig differenzierbar in all seinen Variablen ist, also die Induktionsvoraussetzung für den Induktionsindex $k-1$ erfüllt. Wenden wir also die Induktionshypothese auf dieses g an, so erhalten wir, dass die Lösungsfunktion ϕ_g zu g wohldefiniert ist und zudem $(k-1)$-fach nach all ihren Variablen $t, \sigma, z, \tilde{\tau}, \tilde{x}$ und $\tilde{\mu}$ stetig differenziert werden kann. Anhand der Darstellungen

$$\begin{aligned}
\partial_{x_i}\phi(t,\tau,x,\mu) &= \tilde{\phi}_g(t,\tau,(e_i,0),\tau,x,\mu), \qquad i = 1,\dots,n \\
\partial_{\mu_j}\phi(t,\tau,x,\mu) &= \tilde{\phi}_g(t,\tau,(0,e_j),\tau,x,\mu), \qquad j = 1,\dots,m \\
\partial_\tau\phi(t,\tau,x,\mu) &= \tilde{\phi}_g(t,\tau,-(f(\tau,x,\mu),0),\tau,x,\mu)
\end{aligned}$$

aus (4.24), (4.25) und (4.26) der partiellen Ableitungen von ϕ nach x_i, μ_j und τ mittels der ersten n Komponenten $\tilde{\phi}_g$ von ϕ_g folgt, dass $\partial_x\phi$, $\partial_\mu\phi$ und $\partial_\tau\phi$ nach all ihren Variablen t, τ, x und μ $(k-1)$-fach stetig partiell differenziert werden können. Ausserdem folgt anhand von (4.27) anhand der $(k-1)$-fachen stetigen Differenzierbarkeit von ϕ und der Induktionsvoraussetzung an f, dass auch $\partial_t \phi$ nach allen Variablen t, τ, x und μ $(k-1)$-fach stetig partiell differenziert werden kann. Insgesamt beweist dies also, dass ϕ $k-$fach stetig differenzierbar in all ihren Variablen ist. Kombinieren wir dies erneut mit (4.27) und mit der Induktionsvoraussetzung an f, so folgt hieraus, dass zudem $\partial_t \phi$ k-fach stetig differenzierbar in all ihren Variablen t, τ, x und μ ist. Also funktionierte der Induktionsschritt von $k-1$ auf k, und das Theorem ist bewiesen, da die letzte Behauptung des Satzes nun direkt aus dem Satz von Schwarz folgt. $\square$

5 Phasenflüsse im $\mathbb{R}^n$ und auf Untermannigfaltigkeiten des $\mathbb{R}^n$

5.1 Phasenflüsse auf dem $\mathbb{R}^n$

Definition 5.1.1. *Es sei $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^n$, $(t, x) \mapsto f(t, x)$, ein k-fach stetig differenzierbares Vektorfeld, für ein $k > 0$. Bezeichne ϕ_f die nach Theorem 4.0.4 k-fach stetig differenzierbare Lösungsfunktion zu f, die jedem Tripel $(t, \tau, x) \in \mathcal{D}$ den Ortsvektor $y_{\tau,x}(t)$ der eindeutigen, maximalen Lösung $y_{\tau,x} : I_{(\tau,x)} \longrightarrow \mathbb{R}^n$ des AWP's*

$$y'(t) = f(t, y), \qquad y(\tau) = x \tag{5.1}$$

zuordnet, also durch $\phi_f(t, \tau, x) := y_{\tau,x}(t)$ definiert sei. Wir definieren nun den von f generierten ,,Phasenfluss" Ψ durch $\Psi(t, x) := \phi_f(t, 0, x)$, für jedes Paar $(0, x) \in \Omega$ und jedes $t \in I_{(0,x)}$.

Definition 5.1.2. *Wir nennen ein stetig differenzierbares Vektorfeld $f : \Omega \to \mathbb{R}^n$ vollständig, falls sein Phasenfluss $\Psi(t, x)$ für jedes fixierte Paar $(0, x) \in \Omega$ für alle $t \in \mathbb{R}$ existiert, d.h. falls das Existenzintervall $I_{0,x}$ der eindeutigen, maximalen Lösung $y_{0,x}$ von* (5.1) *(mit $\tau = 0$) ganz $\mathbb{R}$ ist.*

Ist beispielsweise $\Omega = \mathbb{R} \times \mathbb{R}^n$ und ist $f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ stetig differenzierbar und von höchstens linearem Wachstum in x (auf allen kompakten Zeit-Intervallen $[-R, R]$) im Sinne von (2.26), so folgt aus Theorem 2.0.7, dass f vollständig ist.

Theorem 5.1.1. *[Lie's Diffeomorphismen-Gruppe] Es sei $\Omega = \mathbb{R} \times \tilde{\Omega}$, $\tilde{\Omega} \subset \mathbb{R}^n$ offen und $f : \Omega \to \mathbb{R}^n$ ein vollständiges Vektorfeld der Klasse C^k, für ein $k > 0$, welches von t unabhängig sei, also $f(t, x) = f(s, x)$ für alle $s, t \in \mathbb{R}$ und jedes $x \in \tilde{\Omega}$ erfülle. Dann ist der von f erzeugte Phasenfluss Ψ nach t und x k-fach stetig differenzierbar und erfüllt bzgl. der Zeit t die ,,Funktionalgleichung"*

$$\Psi(t_1 + t_2, \cdot\,) = \Psi(t_2, \cdot\,) \circ \Psi(t_1, \cdot\,) \qquad \forall\, t_1, t_2 \in \mathbb{R}. \tag{5.2}$$

Hieraus folgt insbesondere $\Psi(t_1, \cdot\,) \circ \Psi(t_2, \cdot\,) = \Psi(t_2, \cdot\,) \circ \Psi(t_1, \cdot\,)$ $\forall\, t_1, t_2 \in \mathbb{R}$ und dass $\Psi(t, \cdot\,) : \tilde{\Omega} \to \tilde{\Omega}$ ein orientierungserhaltender Diffeomorphismus der Klasse C^k mit $\Psi(t, \cdot\,)^{-1} = \Psi(-t, \cdot\,)$, für jedes $t \in \mathbb{R}$ ist. Ausserdem kann die Jacobi-Determinante von Ψ für jeden festen Punkt $x \in \tilde{\Omega}$ durch die Formel

$$\det(\partial_x \Psi)(t, x) = \exp\Big(\int_0^t \operatorname{div}(f)(s, \Psi(s, x))\, ds\Big) > 0 \tag{5.3}$$

für alle Zeiten $t \in \mathbb{R}$ angegeben werden.

Proof. Wegen $\Psi(t,\,\cdot\,) = \phi(t,0,\,\cdot\,)$ ist Ψ anhand der Voraussetzung an f und nach Theorem 4.0.4 k-fach stetig differenzierbar in jedem $t \in \mathbb{R}$ und jedem $x \in \tilde{\Omega}$. Wir fixieren nun einen Punkt $x \in \tilde{\Omega}$ und ausserdem zwei Zeiten $t_1, t_1 \in \mathbb{R}$ beliebig. Da wir f als vollständig voraussetzen, ist die Funktion $y_{0,x}$ aus (5.1) in allen Zeiten definiert und $y_{0,x}(t)$ in $\tilde{\Omega}$ enthalten. Wir können somit zunächst anstatt $y_{0,x} \equiv \Psi(\,\cdot\,,x)$ die Funktion $Y(t) := y_{0,x}(t+t_1)$ betrachten. Wegen $Y'(t) = (y_{0,x})'(t+t_1) = f(t+t_1, y_{0,x}(t+t_1)) = f(t, Y(t))$ und $Y(0) = y_{0,x}(t_1) = \Psi(t_1,x)$ ist Y die eindeutige Lösung des AWP's

$$y'(t) = f(t,y(t)), \qquad y(0) = \Psi(t_1,x),$$

welche mittels der Lösungsfunktion ϕ zu f durch $Y(t) = \phi(t,0,\Psi(t_1,x))$, für jedes $t \in \mathbb{R}$, angegeben werden kann. Vergleichen wir dies mit der Definition von Ψ in Definition 5.1.1, so bedeutet diese Übereinstimmung gerade: $Y(t) = \Psi(t, \Psi(t_1,x))$, sodass wir zusammen mit $y_{0,x}(t) = \Psi(t,x)$ und der Definition von Y sehen:

$$\Psi(t_1+t_2,x) = y_{0,x}(t_1+t_2) = Y(t_2) = \Psi(t_2,\Psi(t_1,x)) = \Psi(t_2,\,\cdot\,)\circ\Psi(t_1,\,\cdot\,)(x).$$

Hieraus folgt insbesondere $\Psi(t_1,\cdot)\circ\Psi(t_2,\cdot) = \Psi(t_2,\cdot)\circ\Psi(t_1,\cdot)$, und beachten wir auch noch $\Psi(0,x) = y_{0,x}(0) = x$, so beweist dies, dass

$$\mathrm{Id}_{\tilde{\Omega}} = \Psi(t-t,\,\cdot\,) = \Psi(t,\,\cdot\,)\circ\Psi(-t,\,\cdot\,) = \Psi(-t,\,\cdot\,)\circ\Psi(t,\,\cdot\,)$$

gilt. Dies zeigt, dass $\Psi(t,\,\cdot\,)$ $\tilde{\Omega}$ auf $\tilde{\Omega}$ bijektiv und mit Inverser Abbildung $\Psi(-t,\,\cdot\,)$ abbildet. Da $\Psi(\pm t,\,\cdot\,)$ bereits als k-fach stetig differenzierbar auf $\tilde{\Omega}$ erkannt wurde, folgt dass $\Psi(t,\,\cdot\,)$ einen Diffeomorphismus der Klasse C^k von $\tilde{\Omega}$ auf $\tilde{\Omega}$ mit $\Psi(t,\,\cdot\,)^{-1} = \Psi(-t,\,\cdot\,)$, für jedes $t \in \mathbb{R}$, liefert. Schliesslich löst die Jacobi-Matrix $\partial_x\Psi \equiv \partial_x\phi(\cdot,0,\cdot)$ von Ψ nach Formel (4.14) das lineare AWP

$$\begin{aligned}(\partial_x\Psi)'(t,x) = \partial_x f(t,\Psi(t,x))\cdot\partial_x\Psi(t,x),&\\ \partial_x\Psi(0,x) = \mathbf{1}_n,&\end{aligned}$$

für alle $t \in \mathbb{R}$ und für beliebig fixiertes $x \in \tilde{\Omega}$, aus dem sich in Kombination mit Proposition 3.1.3 für die Jacobi-Determinante von Ψ die lineare Differentialgleichung

$$\begin{aligned}\frac{d}{dt}(\det(\partial_x\Psi))(t,x) &= \mathrm{Spur}(\partial_x f(t,\Psi(t,x)))\ \det(\partial_x\Psi)(t,x)\\ &= \mathrm{div}(f)(t,\Psi(t,x))\ \det(\partial_x\Psi)(t,x)\end{aligned} \tag{5.4}$$

mit Startwert $\det(\partial_x\Psi)(0,x) = 1$, für beliebig fixiertes $x \in \tilde{\Omega}$ ergibt. Somit erhalten wir in der Tat die in (5.3) angegebene Formel für $\det(\partial_x\Psi)(\,\cdot\,,x)$ anhand des Beispiels 5 aus Kapitel 1, welche insbesondere $\det(\partial_x\Psi)(t,\,\cdot\,) > 0$ auf $\tilde{\Omega}$ für alle Zeiten t und somit die Erhaltung der Orientierung der Diffeomorphismen $\Psi(t,\,\cdot\,)$ impliziert. □

Theorem 5.1.2. *[Satz von Liouville] Es sei $\Omega = \mathbb{R}\times\tilde{\Omega}$, $\tilde{\Omega} \subset \mathbb{R}^n$ offen und $f : \tilde{\Omega} \to \mathbb{R}^n$ ein vollständiges Vektorfeld, welches von t unabhängig sei. Desweiteren sei $B \subset \tilde{\Omega}$ eine beliebige kompakte Teilmenge, und es bezeichne $V_B(t) := \mathcal{L}^n(\Psi(t,B))$ das n-dimensionale*

Lebesgue-Mass von $\Psi(t, B)$*, also das ,,klassische Volumen" derjenigen Teilmenge von* $\tilde{\Omega}$*, in die sich die Menge* B *nach Verstreichen der Zeit* t *vermöge des Flusses* Ψ *deformiert hat. Dann gilt Liouville's Volumen-Formel:*

$$\frac{d}{dt} V_B(t) = \int_{\Psi(t,B)} \operatorname{div}(f)(y)\, d\mathcal{L}^n(y), \qquad \forall\, t \in \mathbb{R}. \tag{5.5}$$

Proof. Aus Theorem 5.1.1 und Formel (5.4) wissen wir, dass die Jacobi-Determinante $\det(\partial_x \Psi)(t, x)$ in jedem Paar $(t, x) \in \Omega$ stetig ist, stetig nach t differenziert werden kann, positiv ist und ausserdem der Differentialgleichung

$$\frac{d}{dt}(\det(\partial_x \Psi))(t, x) = \operatorname{div}(f)(\Psi(t, x))\, \det(\partial_x \Psi)(t, x)$$

genügt. Hieraus folgt erstens, dass auch das von t abhängige Integral $t \mapsto \int_B \det(\partial_x \Psi)(t, x)\, d\mathcal{L}^n(x)$ stetig nach t differenziert werden kann, und zwar vermöge der Formel:

$$\frac{d}{dt} \int_B \det(\partial_x \Psi)(t, x)\, d\mathcal{L}^n(x) = \int_B \operatorname{div}(f)(\Psi(t, x))\, \det(\partial_x \Psi)(t, x)\, d\mathcal{L}^n(x).$$

Da ausserdem $\Psi(t, \cdot\,)$ einen Diffeomorphismus von $\tilde{\Omega}$ auf $\tilde{\Omega}$ liefert, erhalten wir mittels doppelter Anwendung des Transformationssatzes der Lebesgue'schen Mass-Theorie:

$$\frac{d}{dt} V_B(t) = \frac{d}{dt} \mathcal{L}^n(\Psi(t, B)) = \frac{d}{dt} \int_B \det(\partial_x \Psi)(t, x)\, d\mathcal{L}^n(x)$$
$$= \int_B \operatorname{div}(f)(\Psi(t, x))\, \det(\partial_x \Psi)(t, x)\, d\mathcal{L}^n(x) = \int_{\Psi(t,B)} \operatorname{div}(f)(y)\, d\mathcal{L}^n(y),$$

$\forall\, t \in \mathbb{R}$, wie in (5.5) behauptet wurde. □

Korollar 5.1.1. *Es sei* $\Omega = \mathbb{R} \times \tilde{\Omega}$*,* $\tilde{\Omega} \subset \mathbb{R}^n$ *offen und* $f : \tilde{\Omega} \to \mathbb{R}^n$ *ein von der Zeit* t *unabhängiges und vollständiges Vektorfeld, welches ausserdem eine konstante Divergenz* $\operatorname{div}(f) \equiv c \in \mathbb{R}$ *habe. Ist nun* $B \subset \tilde{\Omega}$ *eine beliebige kompakte Teilmenge, so lässt sich das Volumen* $V_B(t) := \mathcal{L}^n(\Psi(t, B))$ *des nach Verstreichen der Zeit* t *durch den von* f *erzeugten Fluss* Ψ *deformierten Körpers* $\Psi(t, B)$ *durch die Formel*

$$V_B(t) = \mathcal{L}^n(B)\, e^{c\,t} \tag{5.6}$$

$\forall\, t \in \mathbb{R}$ *berechnen.*

Proof. Aus Formel (5.5) folgt anhand der vorausgesetzten Konstanz $\operatorname{div}(f) \equiv c$ sofort, dass die Volumen-Funktion V_B das AWP

$$\frac{d}{dt} V_B(t) = c\, \mathcal{L}^n(\Psi(t, B)) = c\, V_B(t), \qquad V_B(0) = \mathcal{L}^n(B)$$

erfüllt. In Kombination mit Beispiel 5 aus Kapitel 1 folgt hieraus Formel (5.6). □

Bemerkung 5.1.1. *1) Falls das Kompaktum $B \subset \tilde{\Omega}$ aus Theorem 5.1.2 der Abschluss einer offenen, beschränkten Teilmenge des $\mathbb{R}^n$ ist und dessen Rand ∂B eine kompakte $(n-1)-$dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^1 im Sinne von Definition 5.2.1 des folgenden Kapitels ist, so sind dessen diffeomorphe Bilder $\Psi(t, \partial B)$ $(n-1)-$dimensionale Untermannigfaltigkeiten des $\mathbb{R}^n$ der Klasse C^1, die gerade $\Psi(t, B)$ beranden, für jedes $t \in \mathbb{R}$. Es gilt also $\partial\Psi(t, B) = \Psi(t, \partial B)$, und man kann Liouville's Formel (5.5) aus Theorem 5.1.2 mit dem Gauss'schen Integralsatz kombinieren und erhält:*

$$\frac{d}{dt} V_B(t) = \int_{\Psi(t,\partial B)} \langle f, \nu_{\Psi(t,\partial B)} \rangle \, d\mathcal{A},$$

wenn $\nu_{\Psi(t,B)}$ das ins Äussere von $\Psi(t, \partial B)$ weisende Einheitsnormalenfeld an $\partial\Psi(t, B) = \Psi(t, \partial B)$ bezeichnet. Diese Formel bestätigt die Anschauung, dass $\langle f, \nu_{\Psi(t,\partial B)} \rangle(y)$ in einem Punkt $y \in \Psi(t, \partial B)$ denjenigen Anteil von f misst, der in die Menge $\Psi(t, B)$ hineinweist bzw. aus $\Psi(t, B)$ hinausweist, sodass das Integral von $\langle f, \nu_{\Psi(t,\partial B)} \rangle$ über die gesamte Oberfläche $\partial\Psi(t, B)$ von $\Psi(t, B)$ die Geschwindigkeit der Zu- bzw. Abnahme des Volumens von $\Psi(t, B)$, also gerade $\frac{d}{dt} V_B(t)$, im Verlaufe der Zeit t misst.

2) Insbesondere zeigt Formel (5.6), dass jedes vollständige, ,,divergenz-freie" Vektorfeld $f : \tilde{\Omega} \to \mathbb{R}^n$ einen ,,volumentreuen" Phasenfluss Ψ generiert. D.h. falls

$$\operatorname{div}(f)(x) := (\partial_{x^1} f_1 + \ldots + \partial_{x^n} f_n)(x) = 0$$

in jedem Punkt $x \in \tilde{\Omega}$ erfüllt ist, so folgt aus Formel (5.6):

$$\mathcal{L}^n(\Psi(t, B)) \equiv \mathcal{L}^n(B),$$

$\forall\, t \in \mathbb{R}$, also dass das $n-$dimensionale, klassische Volumen einer beliebigen kompakten Teilmenge B von $\tilde{\Omega}$ unter dem Einfluss des von f erzeugten Flusses Ψ invariant bleibt. Dies entspricht der ,,physikalischen Deutung" der Divergenz eines Vektorfeldes f, nämlich die Quellen und Senken von f in einer Umgebung eines Punktes $x \in \tilde{\Omega}$ zu messen, welche bereits der Gauss'sche Integralsatz

$$\int_{B_r(x)} \operatorname{div}(f) \, d\mathcal{L}^n = \int_{\partial B_r(x)} \langle f, \nu_{\partial B_r(x)} \rangle \, d\mathcal{A}$$

anhand der geometrischen Bedeutung von $\langle f, \nu_{\partial B_r(x)} \rangle$ nahelegt.

Beispiele:

1.) Sei $f(x) := c \in \mathbb{R}^n$ ein konstantes Vektorfeld auf ganz $\mathbb{R}^n$. Der von f erzeugte Fluss ist in diesem Fall explizit durch $\Psi(t, x) = x + ct$ auf $\mathbb{R} \times \mathbb{R}^n$ gegeben. Ein beliebiges Kompaktum B des $\mathbb{R}^n$ wird durch Ψ also in Richtung von c verschoben, d.h. es gilt $\Psi(t, B) = B + tc$, und die Translationsinvarianz des Lebesgue-Masses $\mathcal{L}^n$ liefert hier sofort:

$$V_B(t) = V_B(0) = \mathcal{L}^n(B)$$

für jedes $t \in \mathbb{R}$. Dies steht in Einklang mit Formel (5.6), da hier $c = \operatorname{div}(f) \equiv 0$ gilt.

2.) Sei $f(x) := \lambda x$, für $x \in \mathbb{R}^n$ und ein $\lambda \in \mathbb{R}$. f erzeugt nach Korollar 3.3.1 den Fluss $\Psi(t,x) = e^{\lambda t}x$, also eine Streckung oder Stauchung des $\mathbb{R}^n$ um den Faktor $e^{\lambda t}$. Somit liefert die elementare Formel $\mathcal{L}^n(s\,B) = s^n\,\mathcal{L}^n(B)$, für jedes $s \in \mathbb{R}_+$, in diesem einfachen Fall sofort:

$$V_B(t) = \mathcal{L}^n(e^{\lambda t}\,B) = e^{n\lambda t}\mathcal{L}^n(B).$$

Da hier $\operatorname{div}(f) = n\lambda$ gilt, stimmt diese elementar gewonnene Formel mit Formel (5.6) für $c = n\lambda$ wieder überein.

3.) Ist $f(x) := A \cdot x$, für eine beliebige reelle $(n \times n)-$Matrix A und $x \in \mathbb{R}^n$, so können wir den Fluss Ψ mittels Korollar 3.3.1 im Allgemeinen nicht mehr explizit berechnen, sodass eine elementare Berechnung von $V_B(t)$ sicherlich nicht mehr möglich ist. Wegen $\operatorname{div}(f) \equiv \operatorname{Spur}(A)$ gibt uns jedoch (5.6) die exakte Auskunft:

$$V_B(t) = e^{\operatorname{Spur}(A)\,t}\mathcal{L}^n(B), \tag{5.7}$$

für alle $t \in \mathbb{R}$.

4.) Speziell für die Matrix $(2 \times 2)-$Matrix $A = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}$ wissen wir aus Aufgabe 20:

$$\exp\big(t\begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}\big) = \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$$

sodass der von $f(x) := A \cdot x$ erzeugte Fluss $\Psi(t,x)$ jeden Punkt $x \in \mathbb{R}^2$ um den Winkel $t\omega$ im Uhrzeigersinn dreht. Da $\mathcal{L}^2$ invariant gegenüber Drehungen ist, erhalten wir hier sofort:

$$V_B(t) \equiv \mathcal{L}^2(B)$$

für jedes Kompaktum $B \subset \mathbb{R}^2$, was wegen $\operatorname{Spur}(A) = 0$ mit Formel (5.7) bzw. Formel (5.6) wieder übereinstimmt.

5.2 Untermannigfaltigkeiten des $\mathbb{R}^n$, Tangentialvektorfelder und Phasenflüsse auf Mannigfaltigkeiten

Definition 5.2.1. *1) Eine nicht-leere Menge $M \subset \mathbb{R}^n$ heisse eine $m-$dimensionale Untermannigfaltigkeit der Klasse C^l, $l > 0$, falls es zu jedem Punkt $y_0 \in M$ eine Umgebung V im $\mathbb{R}^n$, eine offene Menge $U \subset \mathbb{R}^m$ und eine C^l-Abbildung $P : U \xrightarrow{\cong} M \cap V$ gibt, die U homöomorph auf $M \cap V$ abbildet und ausserdem*

$$\operatorname{Rang}(DP(x)) = m \qquad \text{für alle} \quad x \in U$$

erfüllt. Solch eine Abbildung P heisst eine ,,lokale Parametrisierung" oder ,,lokale Karte" von M um y_0, und $M \cap V$ heisst eine ,,Kartenumgebung" von y_0.

2) Wir definieren den Tangentialraum $T_{y_0}M$ von M an den Punkt y_0 durch

$$T_{y_0}M := \mathrm{Bild}(DP(x_0)) = \mathrm{Span}\{\partial_{x^1}P(x_0), \partial_{x^2}P(x_0), \dots, \partial_{x^m}P(x_0)\},$$

wobei $x_0 \in U$ der eindeutige Punkt mit $P(x_0) = y_0$ sei, und wir definieren den Normalraum $T_{y_0}^{\perp}M$ als dessen orthogonales Komplement im $\mathbb{R}^n$, also

$$T_{y_0}^{\perp}M := \{v \in \mathbb{R}^n \mid \langle v, \tau\rangle = 0 \quad \forall\, \tau \in T_{y_0}M\}.$$

In den meisten Anwendungen aus Physik und Geometrie tauchen Untermannigfaltigkeiten als Nullstellenmengen von C^l-Funktionen $f : \mathbb{R}^n \to \mathbb{R}^{n-m}$ auf, für die 0 ein ,,regulärer Wert" (im Bild von f) ist, also sodass

$$\mathrm{Rang}(Df(y)) = n - m \qquad \text{für alle} \quad y \in [f = 0] = f^{-1}(0) \tag{5.8}$$

gilt. In diesem Fall lässt sich mittels des Umkehrsatzes eine hinreichend kleine Umgebung $V = B^m_\epsilon(x_0) \times B^{n-m}_\epsilon(z_0)$ im $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$ eines jeden Punktes $y_0 = (x_0, z_0) \in M$ durch die Inverse der bijektiven Abbildung $y = (x, z) \mapsto (x, f(x, z)) =: F(x, z)$ lokal parametrisieren, die insbesondere $V \cap M$ praktischerweise als Graph über einer m-dimensionalen Ebene parametrisiert, wie der folgende ,,Satz über implizite Funktionen" zeigt:

Theorem 5.2.1. *Sei $\tilde{\Omega} \subset \mathbb{R}^n$ eine offene Teilmenge und $f : \tilde{\Omega} \to \mathbb{R}^{n-m}$ eine C^l-Abbildung, für die* 0 *ein ,,regulärer Wert" (im Bild von f) ist, also sodass* (5.8) *für f gilt. Für einen fixierten Punkt $y_0 \in [f = 0]$ existiert somit eine Permutation der Koordinaten des $\mathbb{R}^n$, sodass gerade*

$$\det\left(\partial_{(y^{m+1},\dots,y^n)} f(y_0)\right) \neq 0 \tag{5.9}$$

gilt. Nach Ausführung dieser Permutation schreiben wir $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$ und $y = (x, z)$ und zeigen: Es existiert eine Umgebung $V = B^m_\epsilon(x_0) \times B^{n-m}_\epsilon(z_0)$ um den Punkt $y_0 = (x_0, z_0)$, sodass die Abbildung $F(x, z) := (x, f(x, z))$ einen C^l-Diffeomorphismus von V auf eine Umgebung $B^m_\epsilon(x_0) \times W$ von $(x_0, 0)$ liefert, und deren inverse C^l-Abbildung $P : B^m_\epsilon(x_0) \times W \to V$ hat die Graphen-Form $P(x, c) = (x, P^{m+1}(x, c), \dots, P^n(x, c))$ und bildet für jedes $c \in W$ das m-dimensionale ,,Blatt" $B^m_\epsilon(x_0) \times \{c\}$ homöomorph auf die Untermannigfaltigkeit $V \cap [f = c]$ ab. Insbesondere liefert die Abbildung $x \mapsto P(x, 0) =: (x, g(x))$ eine lokale C^l-Parametrisierung von $V \cap [f = 0]$ in ,,Graphen-Form" über dem $m-$dimensionalen Ball $B^m_\epsilon(x_0)$, es gilt

$$T_{(x,g(x))}[f = 0] = \mathrm{Span}(\{\partial_{x^1}P(x, 0), \partial_{x^2}P(x, 0), \dots, \partial_{x^m}P(x, 0)\}) \tag{5.10}$$

und ganz $[f = 0]$ ist eine $m-$dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ im Sinne von Definition 5.2.1.

Proof. Wir fixieren einen Punkt $y_0 \in [f = 0]$ und wissen per Voraussetzung an f, dass es $n - m$ linear unabhängige partielle Ableitungen von f in y_0 gibt. Somit können wir eine Permutation der Koordinaten $y^1, y^2, \dots, y^n$ des $\mathbb{R}^n$ durchführen, sodass gerade (5.9) in y_0 gilt. Wir zerlegen demnach $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$, schreiben $y = (x, z)$ und wissen nach (5.9), dass $\partial_z f(x_0, z_0)$ invertierbar ist. Die Jacobi-Matrix der Abbildung $F(x,z) := (x, f(x,z))$ hat die günstige Block-Gestalt

$$\partial_{(x,z)} F(x,z) = \begin{pmatrix} \mathbf{1}_m & 0 \\ \partial_x f(x,z) & \partial_z f(x,z) \end{pmatrix}$$

und muss somit anhand der Invertierbarkeit von $\partial_z f(x_0, z_0)$, wiederum im Punkt $y_0 = (x_0, z_0)$ invertierbar sein. Der Umkehrsatz garantiert somit die Existenz einer Umgebung $V = B_\epsilon^m(x_0) \times B_\epsilon^{n-m}(z_0)$ von $y_0 = (x_0, z_0)$ im $\mathbb{R}^n$, sodass F diese Umgebung V bijektiv, und somit C^l-diffeomorph, auf eine offene Umgebung von $F(x_0, z_0) = (x_0, f(y_0)) = (x_0, 0)$ im $\mathbb{R}^n$ abbildet. Anhand der Graphen-Gestalt von F und der Wahl von V muss diese Umgebung die Form $B_\epsilon^m(x_0) \times W$, für eine offene Umgebung W von 0 im $\mathbb{R}^{n-m}$, haben. Desweiteren sehen wir per Definition von F:

$$(x,c) = F \circ P(x,c) = (P^1(x,c), \ldots, P^m(x,c), f(P(x,c))) \qquad \forall\, (x,c) \in B_\epsilon^m(x_0) \times W,$$

woraus insbesondere $P^1(x,c) = x^1, \ldots, P^m(x,c) = x^m$, also die Graphen-Form $P(x,c) = (x, P^{m+1}(x,c), \ldots, P^n(x,c))$ folgt. Es gilt für $y = (x,z) \in V$ genau dann $f(y) = c$ für ein $c \in W$, wenn $F(x,z) = (x, f(x,z)) = (x,c)$ gilt, was zu $P(x,c) = (x,z) = y$ äquivalent ist. Somit liefert für jedes fixierte $c \in W$ die Einschränkung $P(\,\cdot\,, c)$ der bijektiven Abbildung $P : B_\epsilon^m(x_0) \times W \to V$ auf das m-dimensionale „Blatt" $B_\epsilon^m(x_0) \times \{c\}$ einen Homöomorphismus von $B_\epsilon^m(x_0) \times \{c\}$ auf die Menge $V \cap [f = c]$. Und da die m partiellen Ableitungen $\partial_{x^1} P(\,\cdot\,, c), \partial_{x^2} P(\,\cdot\,, c), \ldots, \partial_{x^m} P(\,\cdot\,, c)$ von $P(\,\cdot\,, c)$ anhand der Graphen-Form von P linear unabhängig sind, gilt insbesondere

$$\mathrm{Rang}(\partial_x P(x,c)) = m \qquad \text{für alle} \quad (x,c) \in B_\epsilon^m(x_0) \times W,$$

sodass die Einschränkung $P(\,\cdot\,, c)$ eine Parametrisierung von $V \cap [f = c]$ im Sinne von Definition 5.2.1, für jedes fixierte $c \in W$, liefert. Insbesondere beweist dies, dass $V \cap [f = c]$ und insbesondere ganz $[f = 0]$ $m-$dimensionale Untermannigfaltigkeiten des $\mathbb{R}^n$ im Sinne von Definition 5.2.1 sein müssen, dass die Abbildung $x \mapsto P(x,0) =: (x, g(x))$, mit $g(x) := (P^{m+1}(x,0), \ldots, P^n(x,0))$, eine lokale Parametrisierung von $[f = 0]$ um den fixierten Punkt $y_0 \in [f = 0]$ liefert, die hier sogar $V \cap [f = 0]$ als Graphen der C^l-Abbildung $g : B_\epsilon^m(x_0) \to B_\epsilon^{n-m}(z_0)$ über dem $m-$dimensionalen Ball $B_\epsilon^m(x_0)$ darstellt und schliesslich dass die m partiellen Ableitungen $\partial_{x^1} P(x,0), \partial_{x^2} P(x,0), \ldots, \partial_{x^m} P(x,0)$ den $m-$dimensionalen Tangentialraum $T_{(x,g(x))}[f = 0]$ aufspannen, wie in (5.10) behauptet wurde. $\square$

Bemerkung 5.2.1. *Der obige Satz zeigt also, dass sich jede hinreichend kleine Umgebung V einer sogenannten „gleichungs-definierten" Untermannigfaltigkeit $M = [f = 0]$ des $\mathbb{R}^n$ der Dimension m mittels des C^l-Diffeomorphismus' F in eine durch $c \in W$ parametrisierte - also $(n-m)-$dimensionale - Familie $m-$dimensionaler Bälle $B_\epsilon^m(x_0) \times \{c\}$ homöomorph abbilden, also „abplätten", lässt, sodass umgekehrt jeder Durchschnitt $V \cap [f = c]$ eine „gleichungs-definierte" Untermannigfaltigkeit der Dimension m sein muss, die sich mittels der zu F inversen Abbildung $P(\,\cdot\,, c)$ als Graph über dem „platten" Ball $B_\epsilon^m(x_0)$ parametrisieren lässt.*

Beispiel:

Beispielsweise ist der Rand

$$M_{a_1,\ldots,a_n} := \{y \in \mathbb{R}^n \mid \frac{y_1^2}{a_1^2} + \frac{y_2^2}{a_2^2} + \ldots + \frac{y_n^2}{a_n^2} = 1\}$$

des Voll-Ellipsoids

$$E_{a_1,\ldots,a_n} := \{y \in \mathbb{R}^n \mid \frac{y_1^2}{a_1^2} + \frac{y_2^2}{a_2^2} + \ldots + \frac{y_n^2}{a_n^2} \leq 1\}$$

mit den Achsenlängen $a_1, a_2, \ldots, a_n > 0$ eine $(n-1)-$dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$. Offenbar ist diese von der Klasse C^∞ und ,,gleichungs-definierbar", denn sie ist die Nullstellenmenge der $C^\infty-$Funktion

$$f(y_1, y_2, \ldots, y_n) := \frac{y_1^2}{a_1^2} + \frac{y_2^2}{a_2^2} + \ldots + \frac{y_n^2}{a_n^2} - 1,$$

deren Gradient

$$\nabla f(y_1, y_2, \ldots, y_n) = 2\,(\frac{y_1}{a_1^2}, \frac{y_2}{a_2^2}, \ldots, \frac{y_n}{a_n^2})$$

auf $\mathbb{R}^n \setminus \{0\}$, insbesondere also auf $M_{a_1,\ldots,a_n} = [f = 0]$, ungleich Null ist.
Insbesondere für $a_1 = a_2 = \ldots = a_n = a > 0$ erhalten wir hieraus, dass der Rand $\mathbb{S}_a^{n-1}$ des offenen Balles $B_a^n(0)$ vom Radius $a > 0$ eine $(n-1)-$dimensionale, ,,gleichungs-definierte" Untermannigfaltigkeit des $\mathbb{R}^n$ von der Klasse C^∞ ist.

Definition 5.2.2. *Es sei M eine $m-$dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ der Klasse C^l. Wir nennen eine Abbildung $G : M \to \mathbb{R}^s$,,von der Klasse C^k", für ein $k \in \{0, 1, \ldots, l\}$, falls für eine beliebige lokale Parametrisierung $P : U \xrightarrow{\cong} M \cap V$ von M, mit $U \subset \mathbb{R}^m$ offen, die Verkettung $G \circ P : U \to \mathbb{R}^s$ von der Klasse C^k (im klassischen Sinn) ist.*

Definition 5.2.3. *Es sei M eine $m-$dimensionale Untermannigfaltigkeit des $\mathbb{R}^n$ von der Klasse C^l. Wir nennen eine Abbildung $\Upsilon : M \to \mathbb{R}^n$ ein ,,Tangentialvektorfeld" an M von der Klasse C^k, für ein $k \in \{0, 1, \ldots, l\}$, falls $\Upsilon(y) \in T_yM$ für jedes $y \in M$ gilt und Υ eine Abbildung der Klasse C^k im Sinne von Definition 5.2.2 ist.*

Wir verwenden nun Theorem 5.2.1 insbesondere zur Fortsetzung von Tangentialvektorfeldern von einer beliebig fixierten, kompakten ,,gleichungs-definierten" Untermannigfaltigkeit $M = [f = 0]$ auf eine hinreichend dünne offene Umgebung $B_\delta(M)$ von M im $\mathbb{R}^n$.

Lemma 5.2.1. *Es sei $M = [f = 0]$ eine $m-$dimensionale, kompakte, gleichungs-definierte Untermannigfaltigkeit des $\mathbb{R}^n$ von der Klasse C^l und $\Upsilon : M \to \mathbb{R}^n$ ein Tangentialvektorfeld an M von der Klasse C^k, $0 \leq k \leq l$. Dann existiert ein $\delta > 0$, sodass Υ eine Fortsetzung $\tilde{\Upsilon} : B_\delta(M) \to \mathbb{R}^n$ der Klasse C^k auf die $\delta-$Umgebung*

$$B_\delta(M) := \{y \in \mathbb{R}^n \mid \mathrm{dist}(y, M) < \delta\}$$

von M besitzt. Es ist also $\tilde{\Upsilon}$ ein C^k-Vektorfeld auf $B_\delta(M)$ im klassischen Sinne mit $\tilde{\Upsilon} = \Upsilon$ auf M. Ausserdem existiert eine Konstante $C(\Upsilon) \in \mathbb{R}$ mit

$$\sup_{y \in B_\delta(M)} \mid \tilde{\Upsilon}(y) \mid \leq C(\Upsilon).$$

Proof. Nach Theorem 5.2.1 existiert zu jedem Punkt $y_0 \in M$ eine Umgebung V im $\mathbb{R}^n$, ein offener ,,Zylinder" U in $\mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$ und ein C^l-Diffeomorphismus $P : U \to V$, dessen Einschränkung auf $U \cap \mathbb{R}^m$ eine lokale Parametrisierung von $M \cap V$ liefert. Da M kompakt ist, können wir eine endliche Familie dieser ,,Karten" auswählen, die ganz M überdecken. Es existieren also endlich viele Karten $P_i : U_i \to V_i$, $i = 1, \ldots, N$, mit $M \subset \bigcup_{i=1}^N V_i$. Da M kompakt ist, das Komplement der Vereinigung der offenen Mengen V_i abgeschlossen im $\mathbb{R}^n$ ist und M durch $\bigcup_{i=1}^N V_i$ überdeckt wird, existiert ein $\delta > 0$, sodass auch noch der kompakte Abschluss der $\delta-$Umgebung $B_\delta(M)$ in der Vereinigung $\bigcup_{i=1}^N V_i$ enthalten ist. Zu dieser Familie $\{V_i\}$ und $B_\delta(M)$ existieren eine Familie von C^k-Funktionen $\eta_i : \mathbb{R}^n \to [0, 1]$ mit kompaktem Träger $\overline{\{y \in \mathbb{R}^n \mid \eta_i(y) \neq 0\}}$ in V_i und mit der ,,Partitions-Eigenschaft"

$$\sum_{i=1}^N \eta_i(y) = 1 \qquad \forall\, y \in \overline{B_\delta(M)}. \tag{5.11}$$

Wir fixieren nun eine dieser Umgebungen V_i und führen eine (eventuell notwendige) Permutation der Koordinaten des $\mathbb{R}^n$ aus, sodass nach Theorem 5.2.1 V_i die Form $V_i = B_\epsilon^m(x_0) \times B_\epsilon^{n-m}(z_0)$ um einen Punkt $y_0 = (x_0, z_0) \in M$ hat und sodass die lokale Parametrisierung P_i ein C^l-Diffeomorphismus der offenen Teilmenge $B_\epsilon^m(x_0) \times W_i$ des $\mathbb{R}^n$ auf die gesamte Umgebung V_i von y_0 ist, welche in Graphen-Form $P_i(x, c) = (x, P_i^{m+1}(x, c), \ldots, P_i^n(x, c))$ gegeben ist. Insbesondere liefert P_i also eine lokale Parametrisierung von $M \cap V_i$ in Graphenform $P_i(x, 0) = (x, g_i(x))$, für $x \in B_\epsilon^m(x_0)$. Zu jedem $y \in B_\delta(M) \cap V_i$ existiert nun genau ein Paar $(x, c) \in B_\epsilon^m(x_0) \times W_i$ mit $P_i(x, c) = y$, und wir definieren zunächst die lokale Fortsetzung Υ_i von Υ in einen Punkt $y \in B_\delta(M) \cap V_i$ durch

$$\Upsilon_i(y) := \Upsilon(P_i(x, 0)) = \Upsilon(x, g_i(x)).$$

Per dieser Definition ist Υ_i von der Klasse C^k im klassischen Sinne auf $B_\delta(M) \cap V_i$, da Υ als C^k-Vektorfeld an M im Sinne von Definition 5.2.3 vorausgesetzt wurde. Wir definieren nun die gesamte Fortsetzung $\tilde{\Upsilon}$ von Υ in einen beliebigen Punkt $y \in B_\delta(M)$ durch die ,,Konvex-Kombination"

$$\tilde{\Upsilon}(y) := \sum_{i=1}^N \eta_i(y)\, \Upsilon_i(y).$$

Anhand der C^k-Regularität der η_i und Υ_i ist $\tilde{\Upsilon} : B_\delta(M) \to \mathbb{R}^n$ ebenfalls von der Klasse C^k. Ist ausserdem $y \in M$ beliebig gewählt, so ergibt die Definition von Υ_j: $\Upsilon_j(y) = \Upsilon(y)$ für jedes j mit $y \in V_j$. Somit folgt aus (5.11) und aus $\eta_i \equiv 0$ auf $\mathbb{R}^n \setminus V_i$, für jedes i:

$$\tilde{\Upsilon}(y) = \sum_{i=1}^N \eta_i(y)\, \Upsilon_i(y) = (\sum_{i=1}^N \eta_i(y))\, \Upsilon(y) = \Upsilon(y),$$

für jedes $y \in M$, sodass $\tilde{\Upsilon}$ in der Tat eine C^k-Fortsetzung von Υ auf $B_\delta(M)$ ist. Desweiteren existiert anhand der Kompaktheit von M eine Konstante $C(\Upsilon) \in \mathbb{R}$ mit $\sup_{y\in M} \mid \Upsilon(y) \mid \leq C(\Upsilon)$. Somit erhalten wir wiederum aus der Definition von $\tilde{\Upsilon}$ und der Υ_i anhand der Eigenschaften der Partitions-Familie $\{\eta_i\}$:

$$\mid \tilde{\Upsilon}(y) \mid \leq \sum_{i=1}^{N} \eta_i(y) \mid \Upsilon_i(y) \mid \leq (\sum_{i=1}^{N} \eta_i(y))\, C(\Upsilon) = C(\Upsilon)$$

für jedes $y \in B_\delta(M)$, was die letzte Behauptung des Lemmas beweist. $\square$

Nun sei $U \subset \mathbb{R}^m$ eine offene Menge und $P : U \xrightarrow{\cong} M \cap V$ eine lokale C^l-Parametrisierung von M um einen festen Punkt $y_0 \in M$. Per Definition 5.2.1 gilt

$$\text{Rang}(DP(x)) = m \qquad \text{für alle} \quad x \in U,$$

sodass $DP(x)$ den $\mathbb{R}^m$ isomorph auf sein Bild im $\mathbb{R}^n$, also auf den Tangentialraum T_yM an M im Punkt $y = P(x)$ abbildet. Ist also $\upsilon : U \to \mathbb{R}^m$ ein Vektorfeld auf U der Klasse C^k, für $1 \leq k \leq l$, so ist $\Upsilon(y) := DP(x) \cdot \upsilon(x)$ ein Tangentialvektorfeld an $M \cap V$ der Klasse C^{k-1}. Ist P von der speziellen Graphenform $P(x) := (x, g(x))$, für eine C^l-Funktion $g : B^m_\epsilon(x_0) \to \mathbb{R}^{n-m}$, wie sie in Theorem 5.2.1 konstruiert wird, so gilt genauer

$$\Upsilon(x, g(x)) = (\upsilon(x), Dg(x) \cdot \upsilon(x)), \qquad \forall\, x \in B^m_\epsilon(x_0). \tag{5.12}$$

Ist umgekehrt $\Upsilon : M \cap V \to \mathbb{R}^n$ ein C^k-Tangentialvektorfeld an $M \cap V$, so ist dessen Projektion $\upsilon(x) := (\Upsilon^1(x, g(x)), \dots, \Upsilon^m(x, g(x)))$ auf den $\mathbb{R}^m$ das eindeutig bestimmte Vektorfeld auf $B^m_\epsilon(x_0)$, welches (5.12) erfüllt, da für $P(x) := (x, g(x))$

$$DP(x) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 0 & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ \partial_{x^1} g(x) & \partial_{x^2} g(x) & \dots & \partial_{x^{m-1}} g(x) & \partial_{x^m} g(x) \end{pmatrix}$$

gilt und diese Matrix in jedem $x \in B^m_\epsilon(x_0)$ einen Isomorphismus des $\mathbb{R}^m$ auf $T_{(x,g(x))}M$ liefert, sodass die ersten m Komponenten von Υ im Punkt $(x, g(x))$, also die Projektion $\upsilon(x)$ von $\Upsilon(x, g(x))$ auf den $\mathbb{R}^m$, bereits $\Upsilon(x, g(x))$ vermöge (5.12) festlegen. Insbesondere ist in diesem Fall auch υ von der Klasse C^k.

Lemma 5.2.2. *Es sei $M = [f = 0]$ eine $m-$dimensionale, kompakte, gleichungs-definierte Untermannigfaltigkeit des $\mathbb{R}^n$ von der Klasse C^l, $y_0 \in M$ ein beliebiger Punkt mit Karten-Umgebung $M \cap V$, $\Upsilon : M \cap V \to \mathbb{R}^n$ ein Tangentialvektorfeld an $M \cap V$ von der Klasse C^k, für $1 \leq k \leq l$, und $\tilde{\Upsilon} : B_\delta(M) \cap V \to \mathbb{R}^n$ die in Lemma 5.2.1 konstruierte Fortsetzung von Υ zu einem C^k-Vektorfeld auf $B_\delta(M) \cap V$. Dann gilt für die eindeutige Kurzzeit-Lösung y des AWP's*

$$y' = \tilde{\Upsilon}(y), \qquad y(0) = y_0, \tag{5.13}$$

dass $y(t) \in M$, $\forall\, t \in (-\rho, \rho)$.

Proof. Mittels einer lokalen Parametrisierung $P(x) := (x, g(x))$ von einem Ball $B^m_\epsilon(x_0)$ auf die Umgebung $M \cap V$ von $y_0 = (x_0, z_0)$ liefert das vorgegebene Tangentialvektorfeld $\Upsilon : M \cap V \to \mathbb{R}^n$ das eindeutig bestimmte C^k-Vektorfeld $\upsilon(x) := (\Upsilon^1(x, g(x)), \ldots, \Upsilon^m(x, g(x)))$ auf $B^m_\epsilon(x_0)$, welches (5.12), also

$$\Upsilon(P(x)) = DP(x) \cdot \upsilon(x)$$

$\forall\, x \in B^m_\epsilon(x_0)$ erfüllt. Das AWP

$$x' = \upsilon(x), \qquad x(0) = x_0$$

besitzt eine eindeutige Kurzzeit-Lösung $x(t)$ auf einem Intervall $(-\rho, \rho)$. Dessen Komposition $y(t) := P(x(t))$ mit P erfüllt nun anhand der Kettenregel

$$\begin{aligned} y'(t) = DP(x(t)) \cdot x'(t) = DP(x(t)) \cdot \upsilon(x(t)) &= \Upsilon(P(x(t))) \\ &= \tilde{\Upsilon}(P(x(t))) = \tilde{\Upsilon}(y(t)) \end{aligned}$$

auf $(-\rho, \rho)$ und ausserdem $y(0) = P(x(0)) = (x_0, g(x_0)) = y_0$. Da das AWP (5.13) eine eindeutige Kurzzeit-Lösung hat, muss diese durch die soeben ermittelte Lösung $y(t) = P(x(t))$ auf $(-\rho, \rho)$ gegeben sein, und da diese Funktion insbesondere nach $M \cap V$ abbildet, folgt die Behauptung des Lemmas. $\square$

Theorem 5.2.2. *Es sei M eine $m-$dimensionale, kompakte, gleichungs-definierte Untermannigfaltigkeit des $\mathbb{R}^n$ von der Klasse C^l und $\Upsilon : M \to \mathbb{R}^n$ ein Tangentialvektorfeld an M von der Klasse C^k, für $1 \leq k \leq l$. Dann erzeugt Υ einen Fluss $\Psi : \mathbb{R} \times M \to M$, der die ,,Funktionalgleichung"*

$$\Psi(t_1 + t_2, \cdot\,) = \Psi(t_2, \cdot\,) \circ \Psi(t_1, \cdot\,) \qquad \forall\, t_1, t_2 \in \mathbb{R}$$

erfüllt, aus der insbesondere folgt, dass $\Psi(t, \cdot) : M \to M$ ein Diffeomorphismus der Klasse C^k für jedes $t \in \mathbb{R}$ ist.

Proof. Nach Lemma 5.2.1 kann Υ zu einem C^k-Vektorfeld $\tilde{\Upsilon} : B_\delta(M) \to \mathbb{R}^n$ fortgesetzt werden. Nun wählen wir eine C^∞-Funktion $\eta : \mathbb{R}^n \to [0, 1]$ mit $\eta = 1$ auf $\overline{B_{\delta/2}(M)}$ und $\eta = 0$ auf $\mathbb{R}^n \setminus B_{3\delta/4}(M)$ und schneiden mittels dieser $\tilde{\Upsilon}$ ab, d.h. definieren das Vektorfeld $\Lambda : \mathbb{R}^n \to \mathbb{R}^n$ durch

$$\Lambda(y) := \begin{cases} \eta(y)\, \tilde{\Upsilon}(y) & : \quad y \in B_\delta(M) \\ \qquad 0 & : \quad y \in \mathbb{R}^n \setminus B_\delta(M). \end{cases}$$

Λ stimmt mit $\tilde{\Upsilon}$ auf $B_{\delta/2}(M)$ überein, ist wieder von der Klasse C^k und zudem beschränkt auf dem gesamten $\mathbb{R}^n$. Aus Theorem 2.0.7 folgt somit, dass Λ vollständig ist, also dass Λ einen globalen Fluss $\Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ mit allen in Theorem 5.1.1 angegebenen Eigenschaften erzeugt. Nun fixieren wir einen Punkt $y \in M$ beliebig und zeigen, dass $\Psi(t, y) \in M$ für jedes $t \in \mathbb{R}$ gilt. Hierzu definieren wir

$$\mathcal{T} := \{t \in \mathbb{R} \mid \Psi(t, y) \in M\}.$$

$\mathcal{T}$ ist nicht-leer, da $\Psi(0,y) = y \in M$ gilt. Desweiteren ist $\mathcal{T}$ eine abgeschlossene Teilmenge von $\mathbb{R}$. Denn ist $\{t_j\}$ eine beliebige Folge aus $\mathcal{T}$ mit $t_j \to t^* \in \mathbb{R}$, so folgt aus $\Psi(t_j, y) \in M$ zusammen mit der Stetigkeit von $\Psi(\,\cdot\,, y) \equiv y_{0,y}$ auf $\mathbb{R}$ und der Abgeschlossenheit von $M \subset \mathbb{R}^n$, dass auch $\Psi(t^*, y) \in M$ und somit $t^* \in \mathcal{T}$ gelten muss. Sei schliesslich $t^* \in \mathcal{T}$ beliebig fixiert. Aus Theorem 5.1.1 wissen wir, dass $\Psi(t + t^*, y) = \Psi(t, \Psi(t^*, y))$, also $\Psi(t+t^*, y) = y_{0,y_0}(t)$ mit $y_0 := \Psi(t^*, y)$, für alle $t \in \mathbb{R}$ gilt. Wenden wir Lemma 5.2.2 auf $y_0 = \Psi(t^*, y) \in M$ an und beachten wir die Übereinstimmung von Λ mit $\tilde{\Upsilon}$ auf $B_{\delta/2}(M)$, so erhalten wir: $\Psi(t + t^*, y) = y_{0,y_0}(t) \in M$ für alle $t \in (-\rho, \rho)$, für ein hinreichend kleines $\rho > 0$. Per Definition von $\mathcal{T}$ bedeutet dies: $(t^* - \rho, t^* + \rho) \subset \mathcal{T}$. Insgesamt haben wir also bewiesen, dass $\mathcal{T}$ eine nicht-leere, abgeschlossene und offene Teilmenge von $\mathbb{R}$ ist und somit $\mathcal{T} = \mathbb{R}$ gelten muss. Also bildet $\Psi(t,\,\cdot\,)$ für jedes $t \in \mathbb{R}$ die gesamte Mannigfaltigkeit M nach M ab, und da Λ mit $\tilde{\Upsilon}$ auf $B_{\delta/2}(M)$ übereinstimmt, wurde die Einschränkung $\Psi : \mathbb{R} \times M \to M$ des gesamten, von Λ generierten Flusses $\Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ de facto von $\tilde{\Upsilon}$ und damit von Υ erzeugt. Der gesamte von Λ erzeugte Fluss $\Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ erfüllt insbesondere die Funktionalgleichung (5.2), sodass wir diese nun automatisch für die Einschränkung $\Psi : \mathbb{R} \times M \to M$ des gesamten Flusses auf M erhalten. Zusammen mit $\Psi(0, \cdot) = \mathrm{id}_M$ zeigt dies wie im Beweis von Theorem 5.1.1, dass für jedes $t \in \mathbb{R}$ $\Psi(t,\,\cdot\,)$ eine bijektive C^k-Abbildung von M auf M mit C^k-Inversem $\Psi(-t,\,\cdot\,)$, also insbesondere ein C^k-Diffeomorphismus von M auf M ist. □

5.3 Anwendungen auf Variationsprobleme bei holonomen Nebenbedingungen

In diesem letzten Abschnitt sei $M := [f = 0]$ eine fixierte kompakte, $m-$dimensionale, gleichungs-definierte Untermannigfaltigkeit eines $\mathbb{R}^n$ der Klasse C^2 und $a < b \in \mathbb{R}$. Desweiteren definieren wir die Klasse

$$\mathcal{C}_{P,Q} := \{u \in C^1([a,b], \mathbb{R}^n) \mid f(u(x)) = 0,\ u'(x) \neq 0\ \forall x \in [a,b], u\,\text{injektiv},\ u(a) = P,\ u(b) = Q\}$$

aller regulärer C^1-Jordan-Kurven auf M, die zwei fixierte, verschiedene Punkte $P, Q \in M$ miteinander verbinden. Ausserdem sei $F \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R})$ eine beliebige C^1-Funktion in den Variablen $(x, y, p) \in [a,b] \times \mathbb{R}^n \times \mathbb{R}^n$ und

$$\mathcal{F}(u) := \int_a^b F(x, u(x), u'(x))\, dx$$

das zur ,,Lagrange-Funktion" F korrespondierende ,,Variations-Funktional" auf $C^1([a,b], \mathbb{R}^n)$. In diesem Abschnitt bezeichne ausnahmsweise das Symbol $'$ die Ableitung nach $x \in [a,b]$. Wir wollen in diesem Abschnitt eine Antwort auf die folgende klassische Fragestellung geben: Welche Integral- oder Differential-Gleichung muss eine Kurve $u \in \mathcal{C}_{P,Q}$ lösen, falls sie ein lokaler Minimierer von $\mathcal{F}$ innerhalb der ,,Nebenbedingungs-Klasse" $\mathcal{C}_{P,Q}$ ist, also falls

$$\mathcal{F}(u) \leq \mathcal{F}(v) \qquad \forall\, v \in \mathcal{C}_{P,Q} \quad \text{mit} \quad \| u - v \|_{C^1([a,b])} < \epsilon_0 \tag{5.14}$$

für ein beliebig kleines $\epsilon_0 > 0$ gilt ?

Wir definieren zunächst:

Definition 5.3.1. *Wir definieren die erste Variation von $\mathcal{F}$ in $u \in C^1([a,b],\mathbb{R}^n)$ in Richtung einer Störung $\varphi \in C^1([a,b],\mathbb{R}^n)$ durch*

$$\delta\mathcal{F}(u,\varphi) := \frac{d}{dt}\mathcal{F}(u+t\varphi)\mid_{t=0}= \int_a^b \partial_y F(x,u(x),u'(x))\cdot\varphi(x)+\partial_p F(x,u(x),u'(x))\cdot\varphi'(x)\,dx.$$

Definition 5.3.2. *Sei $u \in \mathcal{C}_{P,Q}$, so nennen wir eine Störung $\varphi \in C^1([a,b],\mathbb{R}^n)$ ein ,,tangentielles Vektorfeld entlang u", falls sie $\varphi(x) \in T_{u(x)}M$ in jedem $x \in [a,b]$ erfüllt.*

Bemerkung 5.3.1. *1) Im folgenden Theorem verwenden wir den Funktionenraum*

$$C^1_c((a,b),\mathbb{R}^n) := \{u \in C^1([a,b],\mathbb{R}^n) \mid \overline{\{x \in [a,b] \mid u(x) \neq 0\}} \subset (a,b)\}$$

aller C^1-Funktionen auf $[a,b]$, deren Träger $\overline{\{x \in [a,b] \mid u(x) \neq 0\}}$ im offenen Intervall (a,b) enthalten ist, d.h. die bereits vor den Intervallgrenzen a und b identisch zu verschwinden beginnen.

2) Desweiteren definieren wir hier zur Vorbereitung des Beweises von Theorem 5.3.2, d.h. zur Formulierung von Whitney's Fortsetzungssatz, für eine beliebige abgeschlossene Teilmenge $A \subset \mathbb{R}^m$ und für zwei beliebige Funktionen $g : A \to \mathbb{R}$ und $d : A \to \mathbb{R}^m$ den Quotienten

$$Q(y,x) := \frac{g(y) - g(x) - d(x)\cdot(y-x)}{\mid x-y \mid},$$

und für alle Kompakta $K \subset A$:

$$\rho_{K,g,d}(\delta) := \sup\{\mid Q(y,x) \mid\mid x,y \in K \text{ mit } 0 <\mid x-y \mid\leq \delta\}.$$

Nun können wir Whitney's berühmten ,,C^1-Fortsetzungssatz" formulieren:

Theorem 5.3.1. *[Whitney's C^1-Fortsetzungssatz] Es seien $A \subset \mathbb{R}^m$ eine beliebige abgeschlossene Teilmenge und $g : A \to \mathbb{R}$ und $d : A \to \mathbb{R}^m$ stetige Funktionen, für die*

$$\rho_{K,g,d}(\delta) \to 0 \quad \text{für } \delta \to 0$$

für jedes Kompaktum $K \subset A$ gelte. Dann lässt sich eine Funktion $\bar{g} \in C^1(\mathbb{R}^m,\mathbb{R})$ mit

$$[\bar{g} = g \quad \text{und} \quad \nabla\bar{g} = d] \text{ auf } A,$$

also eine C^1-Fortsetzung $\bar{g}$ von g (zu vorgegebenem Gradienten d auf A) auf den gesamten $\mathbb{R}^m$ explizit konstruieren.

Theorem 5.3.2. *[Notwendige Bedingung zur lokalen Minimierung von $\mathcal{F}$ bei holonomen Nebenbedingungen] Sei $u \in \mathcal{C}_{P,Q}$ ein lokaler Minimierer von $\mathcal{F}$ innerhalb der ,,Nebenbedingungs-Klasse" $\mathcal{C}_{P,Q}$, u erfülle also die lokale Minimierungs-Bedingung* (5.14) *für ein beliebig kleines $\epsilon_0 > 0$. Dann gilt*

$$\delta\mathcal{F}(u,\varphi) = 0$$

für jedes entlang u tangentielle Vektorfeld $\varphi \in C_c^1((a,b),\mathbb{R}^n)$.

Proof. Wir wählen ein entlang u tangentielles Vektorfeld $\varphi \in C_c^1((a,b),\mathbb{R}^n)$ beliebig. Schritt I) Wir konstruieren im ersten Schritt mittels Whitney's Fortsetzungssatzes und mittels der Technik der Partition der Eins ein tangentielles Vektorfeld Υ an M der Klasse C^1, welches φ von Spur(u) auf ganz M fortsetzt, d.h. welches

$$\Upsilon(u(x)) = \varphi(x) \quad \forall\, x \in [a,b] \tag{5.15}$$

erfüllt und ein an ganz M tangentielles C^1-Vektorfeld im Sinne von Definition 5.2.3 ist:
Da u eine reguläre Jordan-Kurve, also eine Parametrisierung für die gesamte 1-dimensionale Untermannigfaltigkeit Spur$(u) \subset M$ ist, existiert genau ein an M tangentielles C^1-Vektorfeld $\tilde{\varphi}$: Spur$(u) \to \mathbb{R}^n$ mit

$$\tilde{\varphi} \circ u = \varphi \quad \text{auf} \quad [a,b]. \tag{5.16}$$

Wie im Beweis von Lemma 5.2.1 können wir anhand der Kompaktheit von Spur(u) eine endliche Familie von lokalen ,,Karten" $P_i : U_i \to V_i$, $i = 1,\dots,N$, von M auswählen, die Spur(u) und ausserdem sogar eine (im $\mathbb{R}^n$) offene $\delta-$Umgebung $B_\delta(\text{Spur}(u))$, für ein hinreichend kleines $\delta > 0$, überdecken. Zu dieser Familie $\{V_i\}$ und zu $B_\delta(\text{Spur}(u))$ wählen wir eine Familie von C^1-Funktionen $\eta_i : \mathbb{R}^n \to [0,1]$ mit kompaktem Träger $\overline{\{y \in \mathbb{R}^n \mid \eta_i(y) \neq 0\}}$ in V_i und mit der ,,Partitions-Eigenschaft"

$$\sum_{i=1}^{N} \eta_i(y) = 1 \qquad \forall\, y \in \overline{B_\delta(\text{Spur}(u))}. \tag{5.17}$$

Wir fixieren nun eine dieser Umgebungen V_i und führen eine (eventuell notwendige) Permutation der Koordinaten des $\mathbb{R}^n$ aus, sodass nach Theorem 5.2.1 V_i die Form $V_i = B_\epsilon^m(\tilde{x}_0) \times B_\epsilon^{n-m}(z_0)$ um einen Punkt $y_0 = (\tilde{x}_0, z_0) \in \text{Spur}(u)$ hat und sodass die lokale Parametrisierung P_i ein C^2-Diffeomorphismus der offenen Teilmenge $B_\epsilon^m(\tilde{x}_0) \times W_i$ des $\mathbb{R}^n$ auf die gesamte Umgebung V_i von y_0 ist, welche in Graphen-Form $P_i(\tilde{x},c) = (\tilde{x}, P_i^{m+1}(\tilde{x},c),\dots,P_i^n(\tilde{x},c))$ gegeben ist. Insbesondere liefert P_i also eine lokale Parametrisierung von $M \cap V_i$ in Graphenform $P_i(\tilde{x},0) = (\tilde{x}, g_i(\tilde{x}))$, für $\tilde{x} \in B_\epsilon^m(\tilde{x}_0)$. Wir erhalten somit zunächst eine reguläre C^1-Jordan-Kurve u_i in $B_\epsilon^m(\tilde{x}_0)$, die von der Karte $P_i(\,\cdot\,,0)$ auf Spur$(u) \cap V_i$ abgebildet wird. Da $\partial_{\tilde{x}} P_i(\tilde{x},0) : \mathbb{R}^m \to T_{(\tilde{x},g_i(\tilde{x}))}M$ ein Isomorphismus ist, existiert zu jedem Punkt $\tilde{x} \in \text{Spur}(u_i)$, also zu jedem $y = (\tilde{x}, g_i(\tilde{x})) \in \text{Spur}(u) \cap V_i$, genau ein Vektor $v_i(\tilde{x}) \in \mathbb{R}^m$ mit

$$\partial_{\tilde{x}} P_i(\tilde{x},0) \cdot v_i(\tilde{x}) = \tilde{\varphi}(\tilde{x}, g_i(\tilde{x})). \tag{5.18}$$

Weil $\partial_{\tilde{x}} P_i(\,\cdot\,,0)$ und φ stetig differenzierbar sind, erfüllen die m Komponenten von υ_i : $\mathrm{Spur}(u_i) \to \mathbb{R}^m$ die Voraussetzungen des obigen C^1-Fortsetzungssatzes von Whitney (hier mit $A := \overline{\mathrm{Spur}(u_i)}$), sodass sich ein C^1-Vektorfeld $\bar{\upsilon}_i : B^m_\epsilon(\tilde{x}_0) \to \mathbb{R}^m$ mit

$$\bar{\upsilon}_i = \upsilon_i \quad \text{auf } \mathrm{Spur}(u_i) \tag{5.19}$$

konstruieren lässt. Wir definieren nun zunächst in einem beliebigen Punkt $y = (\tilde{x}, g_i(\tilde{x})) \in B_\delta(\mathrm{Spur}(u)) \cap M \cap V_i$ das an M tangentielle C^1-Vektorfeld Υ_i durch

$$\Upsilon_i(y) := \partial_{\tilde{x}} P_i(\tilde{x}, 0) \cdot \bar{\upsilon}_i(\tilde{x}) \tag{5.20}$$

und anschliessend die gesamte C^1-Fortsetzung Υ von $\tilde{\varphi}$ in einen beliebigen Punkt $y \in B_\delta(\mathrm{Spur}(u)) \cap M$ durch die ,,Konvex-Kombination"

$$\Upsilon(y) := \sum_{i=1}^N \eta_i(y)\, \Upsilon_i(y).$$

Anhand der C^1-Regularität der η_i und Υ_i ist $\Upsilon : B_\delta(\mathrm{Spur}(u)) \cap M \to \mathbb{R}^n$ ebenfalls von der Klasse C^1, und ausserdem erfüllt es $\Upsilon(y) \in T_yM$, da $\Upsilon_i(y) \in T_yM$ für jedes $i \in \{1, \ldots, N\}$ gilt und T_yM ein linearer Raum, also insbesondere konvex ist. Ist $y \in \mathrm{Spur}(u)$ ein beliebig fixierter Punkt, so gilt ausserdem $\Upsilon_j(y) = \tilde{\varphi}(y)$ für jedes j mit $y \in V_j$ anhand von (5.18), (5.19) und (5.20). Somit folgt aus (5.16), (5.17) und aus $\eta_i \equiv 0$ auf $\mathbb{R}^n \setminus V_i$, für jedes i:

$$\Upsilon(u(x)) = \sum_{i=1}^N \eta_i(u(x))\, \Upsilon_i(u(x)) = \big(\sum_{i=1}^N \eta_i(u(x))\big)\, \tilde{\varphi}(u(x)) = \tilde{\varphi}(u(x)) = \varphi(x),$$

für jedes $x \in [a,b]$, sodass Υ, wie in (5.15) versprochen, eine an M tangentielle C^1-Fortsetzung des entlang u tangentiellen Vektorfeldes φ auf die Umgebung $B_\delta(\mathrm{Spur}(u)) \cap M$ ist. Mittels einer Abschneide-Funktion kann schliesslich Υ, ähnlich wie zu Beginn des Beweises von Theorem 5.2.2, leicht zu einem auf ganz M definierten C^1-Tangentialvektorfeld fortgesetzt werden.

Schritt II) Υ kann wiederum wie im Beweis von Theorem 5.2.2 mittels Lemma 5.2.1 zu einem auf dem gesamten $\mathbb{R}^n$ definierten, beschränkten C^1-Vektorfeld Λ fortgesetzt werden. Theorem 5.1.1 garantiert zunächst, dass Λ einen globalen C^1-Fluss $\Psi : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, also insbesondere eine einparametrige Gruppe von C^1-Diffeomorphismen $\Psi(t,\,\cdot\,)$ von $\mathbb{R}^n$ auf $\mathbb{R}^n$ erzeugt, und Theorem 5.2.2 garantiert, dass die Einschränkung von Ψ auf $\mathbb{R} \times M$ eine einparametrige Gruppe von C^1-Diffeomorphismen von M auf M ist. Definieren wir nun

$$w(x,t) := \Psi(t, u(x)) \qquad \text{für } (x,t) \in [a,b] \times \mathbb{R},$$

so stellt sich w als ein günstiges ,,Variationsfeld" von u innerhalb der Klasse $\mathcal{C}_{P,Q}$ heraus, da es die folgenden Eigenschaften besitzt:

1) $w \in C^1([a,b] \times \mathbb{R}, \mathbb{R}^n)$, $w(\,\cdot\,,t)$ ist injektiv und regulär $\forall\, t \in \mathbb{R}$ und $\mathrm{Bild}(w) \subset M$.

2) $w(x,0) = u(x)$, $\forall\, x \in [a,b]$.

3) $\partial_t w(x,0) = \varphi(x)$, $\forall\, x \in [a,b]$, und es existieren die zweiten partiellen Ableitungen $\partial_t \partial_x w(x,t)$, sind stetig und erfüllen somit $\partial_t \partial_x w(x,t) = \partial_x \partial_t w(x,t)$ in allen $(x,t) \in [a,b] \times \mathbb{R}$. Insbesondere folgt hieraus:
$$\partial_t \partial_x w(x,0) = \partial_x \partial_t w(x,0) = \varphi'(x) \qquad \forall\, x \in [a,b]. \tag{5.21}$$

4) $w(a,t) = u(a) = P$ und $w(b,t) = u(b) = Q$ in allen $t \in \mathbb{R}$.

Eigenschaft (1) von w folgt sofort aus der Definition von w, da u und Ψ C^1-Funktionen sind und u injektiv und regulär ist. Man muss hierbei nur beachten, dass $x \mapsto \Psi(t, u(x))$ für $t \neq 0$ injektiv bleibt, da $\Psi(t, \cdot\,) : M \to M$ für jedes $t \in \mathbb{R}$ ein Diffeomorphismus ist und dass z.B. wegen (5.3)
$$\frac{d}{dx}\Psi(t,u(x)) = \partial_y \Psi(t,u(x)) \cdot u'(x) \neq 0$$
in allen $(x,t) \in [a,b] \times \mathbb{R}$ bleibt.
Eigenschaft (2) folgt ebenfalls direkt per Definition von w, wenn man noch $\Psi(0,y) = y$ $\forall\, y \in M$ beachtet.
Da Ψ von Λ erzeugt wird und $\Lambda(u(x)) = \varphi(x)$ in jedem $x \in [a,b]$ gilt, folgt per Definition 5.1.1 von Ψ:
$$\partial_t w(x,0) = \partial_t \Psi(0,u(x)) = \Lambda(\Psi(0,u(x))) = \Lambda(u(x)) = \varphi(x).$$
Desweiteren wissen wir aus Theorem 4.0.4, dass die doppelten partiellen Ableitungen $\partial_t \partial_{y_i} \Psi$ auf ganz $\mathbb{R} \times \mathbb{R}^n$ existieren, stetig sind und mit $\partial_{y_i} \partial_t \Psi$ auf ganz $\mathbb{R} \times \mathbb{R}^n$ übereinstimmen. Somit liefert die Kettenregel, dass auch die zweiten partiellen Ableitungen $\partial_t \partial_x w(x,t)$ existieren und stetig sind und dass
$$\partial_t \partial_x w(x,t) = \partial_t \partial_y \Psi(t,u(x)) \cdot u'(x) = \partial_y \partial_t \Psi(t,u(x)) \cdot u'(x) = \partial_x \partial_t w(x,t)$$
in allen $(x,t) \in [a,b] \times \mathbb{R}$ gilt, was Behauptung (3) und insbesondere Formel (5.21) beweist.
Anhand von $\Lambda(u(a)) = \varphi(a) = 0$ und $\Lambda(u(b)) = \varphi(b) = 0$ folgt sofort $\Psi(t,u(a)) = u(a)$ und $\Psi(t,u(b)) = u(b)$ in allen $t \in \mathbb{R}$ per Definition 5.1.1, was die vierte behauptete Eigenschaft von w zeigt.
Schritt III) Die Eigenschaften (1) und (4) zeigen, dass alle Funktionen $w(\,\cdot\,,t)$ in $\mathcal{C}_{P,Q}$ liegen. In Kombination mit Eigenschaft (2) sehen wir ausserdem, dass ein $\delta(\epsilon_0) > 0$ mit
$$\| \, u - w(\,\cdot\,,t) \, \|_{C^1([a,b])} < \epsilon_0$$
für $|\, t \,| < \delta$ existiert. Somit können wir mit $v := w(\,\cdot\,,t)$ für jedes $t \in (-\delta,\delta)$ die vorausgesetzte Ungleichung (5.14) testen und erhalten für die Komposition $f(t) := \mathcal{F}(w(\,\cdot\,,t))$: $f(0) \leq f(t)$ für $t \in (-\delta,\delta)$. Da f aus $C^1((-\delta,\delta))$ ist, liefert dies insbesondere $f'(0) =$

0 und somit anhand der Kettenregel, des „Differenzierbarkeitssatzes über parameterabhängige Integrale" und den Eigenschaften (2) und (3) von w:

$$0 = f'(0) = \frac{d}{dt} \int_a^b F(x, w(x,t), w'(x,t))\, dx \,|_{t=0}$$

$$= \int_a^b \partial_y F(x, w(x,0), w'(x,0)) \cdot \partial_t w(x,0) + \partial_p F(x, w(x,0), w'(x,0)) \cdot \partial_t \partial_x w(x,0)\, dx$$

$$= \int_a^b \partial_y F(x, u(x), u'(x)) \cdot \varphi(x) + \partial_p F(x, u(x), u'(x)) \cdot \varphi'(x)\, dx \equiv \delta\mathcal{F}(u, \varphi),$$

nach Definition 5.3.1. □

Wir zitieren nun zwei fundamentale Hilfsmittel der Variationsrechnung, um aus Theorem 5.3.2 ein sehr praktisches Resultat ableiten zu können. Für ihre Beweise sei hier auf das Buch [3], S. 165-166 und S. 317, verwiesen.

Lemma 5.3.1. *[Fundamentallemma der Variationsrechnung] Sei $\Psi \in C^0([a,b], \mathbb{R}^n)$ eine stetige Funktion, die*

$$\int_a^b \langle \Psi(x), \varphi(x) \rangle\, dx = 0$$

für alle $\varphi \in C_c^1((a,b), \mathbb{R}^n)$ erfülle. Dann gilt $\Psi(x) = 0$ für alle $x \in [a,b]$.

Lemma 5.3.2. *[Fundamentallemma der Variationsrechnung bei holonomen Nebenbedingungen] Sei $u \in \mathcal{C}_{P,Q}$ eine Kurve, die zwei Punkte $P \neq Q$ auf M verbindet, und eine stetige Funktion $\Psi \in C^0([a,b], \mathbb{R}^n)$ gehorche der Bedingung*

$$\int_a^b \langle \Psi(x), \varphi(x) \rangle\, dx = 0$$

für jedes entlang u tangentielle Vektorfeld $\varphi \in C_c^1((a,b), \mathbb{R}^n)$. Dann gilt:

$$\Psi(x) \in T_{u(x)}^{\perp} M \qquad \forall\, x \in [a,b],$$

d.h. dann muss $\Psi(x)$ in jedem $x \in [a,b]$ senkrecht auf den Tangentialraum von M im Kurvenpunkt $u(x)$ stehen.

Bemerkung 5.3.2. *Für den Beweis des folgenden Korollars sei darauf hingewiesen, dass der Normalraum $T_{y_0}^{\perp} M$ an eine durch $f : \mathbb{R}^n \to \mathbb{R}^{n-m}$ gleichungs-definierte Mannigfaltigkeit M der Dimension m von den $n-m$ Zeilen von $\partial_y f(y_0)$ aufgespannt wird, also dass*

$$T_{y_0}^{\perp} M = \mathrm{Span}\{\partial_y f_1(y_0), \ldots, \partial_y f_{n-m}(y_0)\} \tag{5.22}$$

in jedem $y_0 \in M$ gilt. Zum Beweis dieser Behauptung wähle man einen beliebigen Tangentialvektor $v \in T_{y_0} M$ und zu diesem eine Kurve $c \in C^1((-\rho, \rho), M)$ mit $c(0) = y_0$, $\dot{c}(0) = v$ und leite die Gleichung $f(c(t)) \equiv 0$ nach t ab:

$$0 = \frac{d}{dt} f(c(t)) \,|_{t=0} = \partial_y f(c(0)) \cdot \dot{c}(0) = \partial_y f(y_0) \cdot v.$$

Dies zeigt zunächst, dass die $n-m$ Zeilen von $\partial_y f(y_0)$ in $T_{y_0}^{\perp}M$ liegen. Da $\partial_y f_1(y_0), \ldots,$ $\partial_y f_{n-m}(y_0)$ nach Voraussetzung (5.8) *an f linear unabhängig sind und $T_{y_0}^{\perp}M$ $(n-m)-$dimensional ist, folgt die Behauptung in* (5.22).

Korollar 5.3.1. *[Euler-Lagrange-Gleichungen zu $\mathcal{F}$ bei holonomen Nebenbedingungen] Sei $u \in \mathcal{C}_{P,Q}$ ein lokaler Minimierer von $\mathcal{F}$ innerhalb der „Nebenbedingungs-Klasse" $\mathcal{C}_{P,Q}$ und ausserdem $u \in C^2([a,b],\mathbb{R}^n)$ und $\partial_p F \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^n)$, so existieren „Lagrange-Multiplikator-Funktionen" $\lambda_j \in C^0([a,b],\mathbb{R})$, $j = 1, \ldots, n-m$, sodass u eine Lösung der „Euler-Lagrange-Gleichungen" zur Funktion*

$$F^*(x,y,p) := F(x,y,p) + \sum_{j=1}^{n-m} \lambda_j(x)\, f_j(y)$$

ist, d.h. u erfüllt das folgende System aus n Differentialgleichungen:

$$\frac{d}{dx}(\partial_p F^*(x,u(x),u'(x))) = \partial_y F^*(x,u(x),u'(x)) \qquad \forall\, x \in [a,b]. \tag{5.23}$$

Proof. Nach Theorem 5.3.2 wissen wir:

$$0 = \delta\mathcal{F}(u,\varphi) = \int_a^b \partial_y F(x,u(x),u'(x)) \cdot \varphi(x) + \partial_p F(x,u(x),u'(x)) \cdot \varphi'(x)\, dx$$

für jedes entlang u tangentielle Vektorfeld $\varphi \in C_c^1((a,b),\mathbb{R}^n)$. Mittels partieller Integration (man beachte hierbei die Voraussetzungen „$u \in C^2$" und „$\partial_p F \in C^1$") erhalten wir hieraus

$$\int_a^b [\partial_y F(x,u(x),u'(x)) - \frac{d}{dx}(\partial_p F(x,u(x),u'(x)))] \cdot \varphi(x)\, dx = 0$$

für jedes entlang u tangentielle Vektorfeld $\varphi \in C_c^1((a,b),\mathbb{R}^n)$, sodass Lemma 5.3.2 garantiert, dass der Vektor $[\partial_y F(x,u(x),u'(x)) - \frac{d}{dx}(\partial_p F(x,u(x),u'(x)))]$ im Normalraum $T_{u(x)}^{\perp}M$ an M im Kurvenpunkt $u(x)$, für jedes $x \in [a,b]$, liegen muss. Da die Ableitungen $\partial_y f_1(u(x)), \ldots, \partial_y f_{n-m}(u(x))$ nach (5.22) den Normalraum $T_{u(x)}^{\perp}M$ in $u(x)$ aufspannen, existieren somit eindeutig bestimmte, reelle Zahlen $\tilde{\lambda}_1(x), \tilde{\lambda}_2(x), \ldots, \tilde{\lambda}_{n-m}(x)$, für die

$$\begin{aligned} \partial_y F(x,u(x),u'(x)) - \frac{d}{dx}(\partial_p F(x,u(x),u'(x))) &= \sum_{j=1}^{n-m} \tilde{\lambda}_j(x)\, \partial_y f_j(u(x)) \\ &\equiv (\partial_y f(u(x)))^T \cdot \begin{pmatrix} \tilde{\lambda}_1(x) \\ \vdots \\ \tilde{\lambda}_{n-m}(x) \end{pmatrix} \end{aligned} \tag{5.24}$$

in jedem $x \in [a,b]$ erfüllt ist. Da $\partial_y f(u(x))$ den maximalen Rang $n-m$ hat, folgt aus der allgemeinen Determinanten-Multiplikationsformel für $\det(A \cdot B^T)$, für zwei $((n-m) \times n)-$Matrizen A, B, hier angewandt auf $A = B = \partial_y f(u(x))$, dass

$$\det\left(\partial_y f(u(x)) \cdot (\partial_y f(u(x)))^T\right) > 0$$

gilt, also dass das Produkt $\partial_y f(u(x)) \cdot (\partial_y f(u(x)))^T$ in jedem $x \in [a,b]$ eine invertierbare $((n-m) \times (n-m))-$Matrix ist. Multiplizieren wir nun die Gleichung (5.24) von links zuerst mit $\partial_y f(u(x))$ und daraufhin von links mit $(\partial_y f(u(x)) \cdot (\partial_y f(u(x)))^T)^{-1}$, so isolieren wir $(\tilde{\lambda}_1(x), \ldots, \tilde{\lambda}_{n-m}(x))$ auf der rechten Seite in (5.24), sodass wir für jede Funktion $\tilde{\lambda}_j$ eine Gleichung erhalten, auf deren linker Seite wegen $u \in C^2([a,b], \mathbb{R}^n)$ und $\partial_p F \in C^1([a,b] \times \mathbb{R}^n \times \mathbb{R}^n, \mathbb{R}^n)$ ein Produkt stetiger Funktionen steht. Dies beweist die Stetigkeit der Funktionen $\tilde{\lambda}_j$ auf $[a,b]$. Setzen wir nun $\lambda_j := -\tilde{\lambda}_j$, so folgt für $F^*(x,y,p) := F(x,y,p) + \sum_{j=1}^{n-m} \lambda_j(x)\, f_j(y)$ in der Tat aus der ersten Gleichung in (5.24):

$$\begin{aligned}
&\frac{d}{dx}(\partial_p F^*(x,u(x),u'(x))) - \partial_y F^*(x,u(x),u'(x)) \\
&= \frac{d}{dx}(\partial_p F(x,u(x),u'(x))) - \partial_y F(x,u(x),u'(x)) - \sum_{j=1}^{n-m} \lambda_j(x)\, \partial_y f_j(u(x)) \\
&= \frac{d}{dx}(\partial_p F(x,u(x),u'(x))) - \partial_y F(x,u(x),u'(x)) + \sum_{j=1}^{n-m} \tilde{\lambda}_j(x)\, \partial_y f_j(u(x)) = 0
\end{aligned}$$

in jedem $x \in [a,b]$, was gerade Formel (5.23) beweist. $\square$

Literaturverzeichnis

[1] Amann, H.: Gewöhnliche Differentialgleichungen, 2. Auflage, de Gruyter, 1995.

[2] Hildebrandt, S.: Analysis I, Springer-Verlag, 2002.

[3] Hildebrandt, S.: Analysis II, Springer-Verlag, 2002.

[4] Reid, W.T.: Ordinary differential equations, Wiley, New York, 1971.

[5] Walter, W.: Gewöhnliche Differentialgleichungen, 7. Auflage, Springer-Verlag, 2000.