Analysis III
Immer gröber werdender Mitschrieb WS 10/11

Ingo Bürk

27. April 2011

Inhaltsverzeichnis

Vorwort

Dieses Skript entstand im Rahmen der Analysis III - Vorlesung bei Hr. Prof. Timo Weidl als Vorlesungsmitschrieb.
Es kann nicht garantiert werden, dass dieses Dokument fehlerfrei ist und der Autor übernimmt für möglicherweise entstandene Schäden jeglicher Art keine Haftung. Dieser Mitschrieb ist kein offizielles Dokument der Universität Stuttgart, Mitarbeiter eben dieser tragen daher ebenfalls keine Verantwortung.

Bei Fragen oder dem Auffinden von Fehlern können Sie mir gerne eine Nachricht zukommen lassen. Schreiben Sie mir dafür einfach eine eMail an die folgende Adresse:

admin@airblader.de

Mit freundlichen Grüßen,
Ingo Bürk.

8 Elemente der Integrationstheorie

8.1 Ring, Algebra und Maß

Sei X eine Menge (die Grundmenge). Außerdem bezeichne $\mathrm{Pot}(X) = 2^X$ die Potenzmenge von X und es sei $A, B \in 2^X$.
Weiterhin bezeichne im Folgenden

$$\begin{aligned} A_X^C = A^C &:= \{\tilde{x} \in X \mid \tilde{x} \notin A\} && \text{(Komplement)} \\ A \setminus B &:= \{x \in A \mid x \notin B\} && \text{(Differenzmenge)} \end{aligned}$$

Definition 8.1 Ring
Eine Mengensystem $\mathcal{R} \subset 2^X$ heißt Ring genau dann, wenn die folgenden Eigenschaften erfüllt sind:

$$\rhd\ \forall A, B \in \mathcal{R} \ \Rightarrow \ A \cup B \in \mathcal{R} \tag{8.1}$$

$$\rhd\ \forall A, B \in \mathcal{R} \ \Rightarrow \ A \setminus B \in \mathcal{R} \tag{8.2}$$

Folgerung.

$$\rhd\ \forall A \ : \ A \setminus A = \emptyset \qquad \Rightarrow \ \emptyset \in \mathcal{R} \qquad (\text{für } \mathcal{R} \neq \emptyset)$$

$$\rhd\ \forall A, B \ : \ A \cap B = A \setminus (A \cup B) \ \Rightarrow \ \forall A, B \in \mathcal{R} \ : \ A \cap B \in \mathcal{R}$$

Definition 8.2 σ-Ring
Ein Mengensystem $\mathcal{R} \subset 2^X$ heißt σ-Ring genau dann, wenn neben (8.1) und (8.2) noch gilt:

$\rhd$ Für jedes abzählbare System von Mengen $A_k \in \mathcal{R}$ mit $k \in \mathbb{N}$ (8.3)

gilt auch $\bigcup\limits_{k \in \mathbb{N}} A_k \in \mathcal{R}$

Definition 8.3 Algebra
Ein Mengensystem $\mathcal{R} \subset 2^X$ heißt Algebra genau dann, wenn neben (8.1) und (8.2) noch gilt:

$$\rhd\ X \in \mathcal{R} \tag{8.4}$$

Definition 8.4 σ-Algebra
Ein Mengensystem $\mathcal{R} \subset 2^X$ heißt σ-Algebra genau dann, wenn (8.1), (8.2), (8.3) und (8.4) gelten.

Anmerkung. Für ein abzählbares System von Mengen $A_k \in \mathcal{R}$ mit $k \in \mathbb{N}$ gilt

$$\bigcap_{k \in \mathbb{N}} A_k = A_1 \setminus \bigcup_{k \in \mathbb{N}} (A_1 \setminus A_k)$$

Daraus folgt für einen σ-Ring $\mathcal{R}$ dann $\bigcap_{k \in \mathbb{N}} A_k \in \mathcal{R}$.

Aufgabe. Die Menge $\mathcal{F}$ bildet eine Algebra, aber keine σ-Algebra. Begründen Sie!

$$\mathcal{F} = \{A \subset \mathbb{N} \mid A \vee A^C \text{ endlich}\}$$

Im Folgenden betrachten wir Funktionen der Form $\varphi : \mathcal{R} \to \mathbb{R}^* = \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$, wobei höchstens einer der Werte $\pm\infty$ angenommen werden darf. Wir werden diese Voraussetzung nicht jedes Mal explizit erwähnen.

Definition 8.5 **(σ-)Additivität**

Die Abbildung $\varphi : \mathcal{R} \to \mathbb{R}^*$ heißt additiv

$$:\Leftrightarrow \forall A, B \in \mathcal{R} \ : \ (A \cap B) = \emptyset$$
$$\Rightarrow \ \varphi(A \cup B) = \varphi(A) + \varphi(B)$$

Die Abbildung $\varphi : \mathcal{R} \to \mathbb{R}^*$ heißt σ-additiv

$$:\Leftrightarrow \{A_k\}_{k\in\mathbb{N}}, A_k \in \mathcal{R}, A_k \cap A_j = \emptyset \ \forall_{k \neq j}$$
$$\Rightarrow \ \varphi\left(\bigcup_{k\in\mathbb{N}} A_k\right) = \sum_{k=1}^{\infty} \varphi(A_k)$$

Anmerkung. Offensichtlich hängt $A = \bigcup_{k\in\mathbb{N}} A_k$ nicht von der Reihenfolge der A_k ab. Für eine σ-additive Abbildung φ muss also auch $\sum_{k=1}^{\infty} \varphi(A_k)$ unabhängig von der Reihenfolge der Summanden sein. Nach dem Umordnungssatz von Riemann konvergiert diese Reihe damit entweder absolut, oder aber sie divergiert bestimmt.

Eigenschaften additiver Funktionen

1. $\varphi(\emptyset) = \emptyset$, denn $\varphi(A) = \varphi(A \mathbin{\dot\cup} \emptyset) = \varphi(A) + \varphi(\emptyset)$
   (Anmerkung: Dies gilt nur, wenn $\varphi(A)$ für mindestens eine Menge endlich ist.)
2. $A_1, \ldots, A_n \in \mathcal{R}$, $A_j \cap A_k = \emptyset \ \forall_{k \neq j}$
   $\Rightarrow \ \varphi(A_1 \cup \ \cdots \ \cup A_n) = \sum_{k=1}^{n} \varphi(A_k)$
3. $\varphi(A_1 \cap A_2) + \varphi(A_1 \cup A_2) = \varphi(A_1) + \varphi(A_2)$

Eigenschaften nicht-negativer additiver Funktionen ($\varphi \geq 0$)

4. $\varphi(A_1 \cup A_2) \leq \varphi(A_1) + \varphi(A_2)$
5. $A_1 \subset A_2 \ \Rightarrow \ \varphi(A_1) \leq \varphi(A_2)$

Satz 8.6

Sei $\mathcal{R} \subset 2^X$ ein σ-Ring und $\varphi : \mathcal{R} \to \mathbb{R}^*$ eine σ-additive Funktion. Außerdem sei $A_k \in \mathcal{R}$ mit $k \in \mathbb{N}$ eine durch die Inklusion geordnete, abzählbare Familie von Mengen, d.h. so dass $A_1 \subset A_2 \subset \cdots$.
Definiere zudem $A := \bigcup_{k\in\mathbb{N}} A_k$.

$$\Rightarrow \ \lim_{k\to\infty} \varphi(A_k) = \varphi(A)$$

Beweis. Definiere eine Folge von Mengen mit $B_1 := A_1$, $B_k := A_k \setminus A_{k-1}$ für $k \geq 2$. Die Mengen B_k sind nach Konstruktion paarweise disjunkt und es gilt $\bigcup_{k\in\mathbb{N}} B_k = A$.
Damit ist

$$\begin{aligned}
\varphi(A) &= \sum_{k=1}^{\infty} \varphi(B_k) \\
&= \lim_{n\to\infty} \sum_{k=1}^{n} \varphi(B_k) \qquad \text{(da die Reihe konvergiert)} \\
&= \lim_{n\to\infty} \varphi \underbrace{\left(\bigcup_{k=1}^{n} B_k \right)}_{=A_n} \\
&= \lim_{n\to\infty} \varphi(A_n)
\end{aligned}$$

Damit ist der Satz bewiesen. □

Definition 8.7 Maß, Maßraum
Ein Maß ist eine auf einem σ-Ring $\mathcal{R}$ gegebene, nicht-negative und σ-additive Funktion $\varphi : \mathcal{R} \to \mathbb{R}^*$.
Das Tripel $(X, \mathcal{R}, \varphi)$ nennt man Maßraum.

Definition 8.8 Messbarkeit
$A \subset X$ heißt messbar $:\Longleftrightarrow \ A \in \mathcal{R}$

Beispiel 1. Sei $X \neq \emptyset$. Dann gibt es die zwei folgenden trivialen σ-Algebren auf X:

1. $\mathcal{R}_1 = \{\emptyset, X\}$
2. $\mathcal{R}_2 = 2^X$
   Für diese σ-Algebra ist z.B. $\varphi(X) = \#\{x \in X\}$ (d.h. die Anzahl der Elemente, falls X finit ist, andernfalls $+\infty$) ein Maß, das so genannte Zählmaß.

Aufgabe. $(X, \mathcal{R}, \varphi)$ sei ein Maßraum, wobei $\mathcal{R}$ eine σ-Algebra ist und $0 \leq \varphi(x) < \infty$. Außerdem gelte für alle $A_k \in \mathcal{R}$ die Inklusionskette $A_1 \supset A_2 \supset \cdots$ sowie $A := \bigcap_{k=1}^{\infty} A_k$. Zeigen Sie, dass dann

$$\varphi(A) = \lim_{k\to\infty} \varphi(A_k)$$

gilt.

8.2 Zur Konstruktion des Lebesgue-Maßes

Sei $X = \mathbb{R}^p$ mit $p \in \mathbb{N}$. Außerdem bezeichne $\mathbb{R}^p \ni x = (x_1, \ldots, x_p)$ mit $x_j \in \mathbb{R}$ einen Punkt in $\mathbb{R}^p$.
Ferner sei Q der offene, halboffene oder abgeschlossene Quader

$$Q = \{x \in \mathbb{R}^p \mid a_j \leq x_j \leq b_j \ , j = 1, \ldots, p\},$$

wobei neben „$\leq$“ auch „$<$“ und das in allen Kombinationen erlaubt ist.
Falls $Q \neq \emptyset$, d.h. $a_j \leq b_j$ für $j = 1, \dots, p$, so sei

$$m(Q) = \prod_{j=1}^{p} (b_j - a_j).$$

Mit $\mathcal{E}$ bezeichnen wir die Menge aller elementaren Mengen, wobei eine elementare Menge eine endliche, disjunkte Vereinigigung von Quadern der obigen Form ist. Es gelten folgende Eigenschaften:

1. $\mathcal{E}$ ist ein Ring.
2. m ist wohldefiniert auf $\mathcal{E}$, denn für $\mathcal{E} \ni E = \bigcup_{j=1}^{n} Q_j$ paarweise disjunkt $\Rightarrow\ m(E) = \sum_{j=1}^{n} m(Q_j)$. „Wohldefiniert“ bedeutet hier, dass der Wert unabhängig von der konkreten (endlichen) Darstellung ist.
3. m ist additiv auf $\mathcal{E}$.

Definition 8.9 reguläre Funktion

Eine (auf $\mathbb{R}^*$ erweiterte) nicht-negative, additive Funktion φ auf $\mathcal{E}$ heißt regulär, falls für alle $A \in \mathcal{E}$ und für alle $\delta > 0$ Mengen $F = F_{\delta,A}$ und $G = G_{\delta,A}$ existieren mit

- $F, G \in \mathcal{E}$ (8.5)
- F abgeschlossen, G offen (8.6)
- $F \subset A \subset G$ (8.7)
- $\varphi(G) - \delta \leq \varphi(A) \leq \varphi(F) + \delta$ (8.8)

Es gilt: m (nach obiger Definition) ist eine reguläre Funktion.

Anmerkung. Sei $\{\mathcal{R}_\tau\}_{\tau \in T}$ eine Folge von σ-Ringen $\mathcal{R}_\tau \subset 2^X$ und $\mathcal{R} := \bigcap_{\tau \in T} \mathcal{R}_\tau$ der Durchschnitt all dieser σ-Ringe. Dann gilt, dass $\mathcal{R}$ auch wieder ein σ-Ring ist, d.h. es gibt einen minimalen σ-Ring $\mathcal{R}$, welcher $\mathcal{E}$ enthält.

8.2.1 Das äußere Maß μ^*

Definition 8.10 Äußeres Maß

Sei $\mu : \mathcal{E} \to \mathbb{R}^*$, $\mu \geq 0$, μ additiv und regulär, sowie $E \subset \mathbb{R}^p$ eine beliebige Menge. Es seien abzählbar viele $A_n \in \mathcal{E}$ gegeben, die E überdecken, d.h. so dass

$$E \subset \bigcup_{n=1}^{\infty} A_n \text{ ist.}$$

Dann sei das äußere Maß μ^* wie folgt definiert:

$$\mu^*(E) := \inf_{E \subset \bigcup_{n=1}^{\infty} A_n} \sum_{n=1}^{\infty} \mu(A_n)$$

Satz 8.11
Es gilt:

$$E \in \mathcal{E} \;\Rightarrow\; \mu^*(E) = \mu(E)$$
$$E = \bigcup_{n=1}^{\infty} E_n,\ E_n \in 2^{\mathbb{R}^p} \;\Rightarrow\; \mu^*(E) \leq \sum_{k=1}^{\infty} \mu^*(E_k) \qquad \text{(Subadditivität)}$$

Man sagt, dass μ^* eine σ-subadditive Fortsetzung von $\mu : \mathcal{E} \to \mathbb{R}^*$ auf $2^{\mathbb{R}^p}$ ist.
Um diesen Satz zu beweisen benötigen wir einen Hilfssatz, den Satz von Heine-Borel. Diesen werden wir nun formulieren und beweisen.

Satz 8.12 **Satz von Heine-Borel**
Sei (M, d) ein metrischer Raum und $F \subset M$. Die Menge F ist kompakt genau dann, wenn aus jeder (abzählbaren) Überdeckung

$$F \subset \bigcup_{n=1}^{\infty} A_n$$

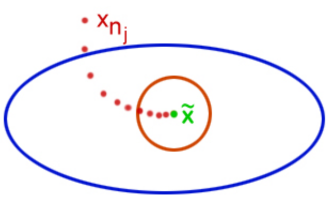
von F mit offenen Mengen A_n eine endliche Überdeckung ausgewählt werden kann, d.h. so dass gilt

$$F \subset \bigcup_{j=1}^{m} A_{n_j}.$$

Beweis. „⇒") Angenommen, F ist kompakt. Wir nehmen an, es existiere keine endliche Teilüberdeckung und leiten daraus einen Widerspruch ab.
Wenn keine solche endliche Teilüberdeckung existiert, so ist folgende Menge nicht leer, d.h. wir finden darin ein Element x_n:

$$\emptyset \neq F \setminus \left(\bigcup_{k=1}^{n} A_k \right) \ni x_n$$

Auf diese Weise erhalten wir eine Folge $(x_n)_{n \in \mathbb{N}}$ in F.



Da F kompakt ist, folgt, dass eine konvergente Teilfolge existiert: $x_{n_j} \to \tilde{x} \in F$. Dann ist

$$\tilde{x} \in F \subset \bigcup_{k=1}^{\infty} A_k \;\Rightarrow\; \exists_{k_0} \;:\; \tilde{x} \in A_{k_0}.$$

Da $x_{n_j} \to \tilde{x}$, folgt $x_{n_j} \in U_\varepsilon(\tilde{x}) \subset A_{k_0}$ für $j \to \infty$ ($n_j \to \infty$). Insbesondere ist dann also $x_{n_j} \in A_{k_0}$.

Betrachte nun $n_j > k_0$, dann ist

$$F \setminus \underbrace{\bigcup_{k=1}^{n_j} A_k}_{\supset A_{k_0}} \ni x_{n_j} \Rightarrow x_{n_j} \notin A_{k_0}. ⚡$$

„$\Leftarrow$“) Angenommen es gibt eine endliche Teilüberdeckung. Weiterhin nehmen wir an, dass F nicht kompakt ist und führen auch das zum Widerspruch.

Wenn F nicht kompakt ist gibt es eine Folge $(x_k)_{k\in\mathbb{N}}$ mit $x_k \in F$, die keinen Häufungspunkt $y \in F$ besitzt. Es gilt also x_k ist kein Häufungspunkt $\forall_k$, d.h. $\forall_k \exists_{\varepsilon_k>0} \forall_{l\neq k} \; : \; x_l \notin U_{\varepsilon_k}(x_k)$. Wir konstruieren nun eine unendliche Überdeckung von F. Betrachte dazu folgende Menge G und einen beliebigen Punkt $y \in G$:

$$G := F \setminus \bigcup_{k=1}^{\infty} \{x_k\} \ni y$$

Aus dem Obigen folgt, dass ein $\varepsilon = \varepsilon(y) > 0$ existiert, so dass $x_k \notin U_{\varepsilon(y)}(y)$, d.h. $U_{\varepsilon(y)}(y) \subset G$. Das bedeutet also, dass G offen ist.

Wähle nun die überdeckende Folge $A_0 := G$ und $A_k = U_{\varepsilon_k}(x_k)$. G ist wie alle diese Umgebungen offen. Die Vereinigung dieser Folge überdeckt F und nach dem Obigen wissen wir, dass wir keine Menge entfernen können, ohne dass dabei die Überdeckungseigenschaft verloren geht. Es ist also nicht möglich, eine endliche Teilüberdeckung zu wählen. Dies steht im Widerspruch zu unserer Annahme. ⚡ □

Wir kommen nun zum Beweis von Satz 8.11.

Beweis. a) Sei $E \in \mathcal{E}$. Wir wissen, dass μ regulär ist. Daraus folgt, dass für alle $\delta > 0$ eine offene Menge $G^\delta \in \mathcal{E}$ existiert, so dass $E \subset G^\delta$ und $\mu(E) + \delta \geq \mu(G^\delta)$ ist.

Wähle nun $A_1 := G^\delta$. Diese Menge ist offen und es gilt $A_1 \in \mathcal{E}$ und $A_1 \supset E$. Daraus folgt, dass $\mu^*(E) \leq \mu(A_1) = \mu(G^\delta) \leq \mu(E) + \delta$ ist. Führe nun den Grenzübergang $\delta \to 0$ durch, so folgt daraus zunächst

$$\mu^*(E) \leq \mu(E). \tag{8.9}$$

Für die andere Richtung „$\geq$“ existiert nach Definition des äußeren Maßes für alle $\delta > 0$ eine abzählbare Familie von offenen und elementaren Mengen A_k^δ ($k \in \mathbb{N}$), so dass $E \subset \bigcup_{k=1}^{\infty} A_k^\delta$ und $\mu^*(E) + \delta \geq \sum_{k=1}^{\infty} \mu(A_k^\delta)$.

Da μ regulär ist existiert für alle $\delta > 0$ ein $F^\delta \in \mathcal{E}$ mit F^δ abgeschlossen, so dass $F^\delta \subset E$ und $\mu(F^\delta) + \delta \geq \mu(E)$.

$$\Rightarrow \underbrace{F^\delta}_{\Rightarrow \text{ kompakt}} \subset E \subset \bigcup_{k=1}^{\infty} \underbrace{A_k^\delta}_{\text{offen}}$$

Mit Hilfe des Satzes von Heine-Borel folgt $\exists_{n\in\mathbb{N}}$, so dass

$$F^\delta \subset \bigcup_{k=1}^{n} A_k^\delta.$$

Da μ eine additive Funktion ist folgt

$$\begin{aligned} \mu(F^\delta) &\leq \sum_{k=1}^{n} \mu(A_k^\delta) \\ \Leftrightarrow \mu(F^\delta) + \delta &\leq \sum_{k=1}^{n} \mu(A_k^\delta) + \delta \\ \Rightarrow \mu(E) &\leq \sum_{k=1}^{\infty} \mu(A_k^\delta) + \delta \\ &\leq \mu^*(E) + \delta + \delta \end{aligned}$$

Insgesamt folgt für $\delta \to 0$ also $\mu(E) \leq \mu^*(E)$. Zusammen mit (8.9) folgt dann **a)**.
b) $E = \bigcup_{n=1}^{\infty} E_n$. Wir nehmen an, es sei $\mu^*(E_n) < +\infty$. Betrachte ein $\delta > 0$ und eine Überdeckung mit offenen Mengen $A_{n,k}^\delta \in \mathcal{E}$, d.h. $E_n \subset \bigcup_{k=1}^{\infty} A_{n,k}^\delta$, so dass $\mu^*(E_n)+2^{-n}\delta \geq \sum_{k=1}^{\infty} \mu(A_{n,k}^\delta)$.

$$\Rightarrow E \subset \bigcup_{n,k} A_{n,k}^\delta \qquad \text{abzählbar viele Mengen.}$$

$$\begin{aligned} \Longrightarrow \quad \mu^*(E) &\leq \sum_{n,k} \mu(A_{n,k}^\delta) \\ &= \sum_n \sum_k \mu(A_{n,k}^\delta) \\ &\leq \sum_n \big(\mu^*(E_n) + 2^{-n}\delta\big) \\ &\leq \left(\sum_n \mu^*(E_n)\right) + \delta \end{aligned}$$

Für $\delta \to 0$ ist der Satz damit bewiesen. □

Seien A, B zwei Mengen. Die symmetrische Differenz ist definiert als

$$\begin{aligned} A \triangle B &:= (A \setminus B) \cup (B \setminus A) \\ &= (A \cup B) \setminus (A \cap B) \end{aligned}$$

Wenn $A, B \in \mathcal{R}$ Ring, so folgt, dass auch $A \triangle B \in \mathcal{R}$ ist.
Für zwei Teilmengen $A, B \subset \mathbb{R}^d$ definiere

$$d^*(A,B) := \mu^*(A \triangle B)$$

Außerdem sagen wir für $A_n, A \subset \mathbb{R}^d$, dass $A_n \xrightarrow{d^*} A \;:\Leftrightarrow\; d^*(A_n, A) \xrightarrow{n\to\infty} 0$.

Eigenschaften:

1. $A \triangle B = B \triangle A$. Daraus folgt $d^*(A,B) = d^*(B,A)$.
2. $A \triangle A = \emptyset$, d.h. $d^*(A,A) = 0$.
3. $A \triangle B \subset (A \triangle C) \cup (B \triangle C)$.

Es ist

$$\begin{aligned} d^*(A,B) &= \mu^*(A \triangle B) \\ &\leq \mu^*\left((A \triangle C) \cup (B \triangle C)\right) && \text{Monotonie} \\ &\leq \mu^*(A \triangle C) + \mu^*(B \triangle C) && \text{Subadditivität} \\ d^*(A,B) &\leq d^*(A,C) + d^*(C,B). \end{aligned}$$

Außerdem ist $d^*(A,B) = \mu^*(A \triangle B) \geq 0$.
d^* ist aber **keine** Metrik, denn $d^*(A,B) = 0 \not\Rightarrow A = B$. Wähle dafür zwei Mengen so, dass sie gleich sind, B aber einen Punkt mehr als A enthält. Die symmetrische Differenz ist dann eine einpunktige Menge mit $d^*(A,B) = 0$ aber $A \neq B$.

Es gelten folgende Teilmengenbeziehungen:

$$\begin{aligned} &\rhd\ (A_1 \cup B_1) \triangle (A_2 \cup B_2) \;\subset\; (A_1 \triangle A_2) \;\cup\; (B_1 \triangle B_2) \\ &\rhd\ (A_1 \cap B_1) \triangle (A_2 \cap B_2) \;\subset\; (A_1 \triangle A_2) \;\cup\; (B_1 \triangle B_2) \\ &\rhd\ (A_1 \setminus B_1) \triangle (A_2 \triangle B_2) \;\subset\; (A_1 \triangle A_2) \;\cup\; (B_1 \triangle B_2) \end{aligned}$$

Daraus folgen die folgenden drei Ungleichungen:

$$\begin{aligned} &\rhd\ d^*(A_1 \cup B_1, A_2 \cup B_2) \;\leq\; d^*(A_1, A_2) + d^*(B_1, B_2) \\ &\rhd\ d^*(A_1 \cap B_1, A_2 \cap B_2) \;\leq\; d^*(A_1, A_2) + d^*(B_1, B_2) \\ &\rhd\ d^*(A_1 \setminus B_1, A_2 \setminus B_2) \;\leq\; d^*(A_1, A_2) + d^*(B_1, B_2) \end{aligned}$$

Seien nun $A, B \subset \mathbb{R}^d$ und $\mu^*(A) < \infty$. Daraus folgt

$$|\mu^*(A) - \mu^*(B)| \leq d^*(A,B).$$

Beweis. Wir nehmen an, es gelte $\mu^*(A) \geq \mu^*(B) \geq 0$. Es ist

$$\begin{aligned} \mu^*(A) &= \mu^*(A \triangle \emptyset) \\ &= d^*(A, \emptyset) \\ &\leq d^*(A,B) + d^*(B,\emptyset) \\ &\leq d^*(A,B) + \mu^*(B) \\ \mu^*(A) - \mu^*(B) &\leq d^*(A,B). \end{aligned}$$

□

Definition 8.13 **Endlich μ-messbare Menge**
Eine Menge $A \subset \mathbb{R}^d$ heißt endlich μ-messbar ($A \in \mathcal{M}_F(\mu)$) genau dann, wenn es eine Folge $A_n \in \mathcal{E}$ mit $A_n \xrightarrow{d^*} A$ gibt.

Definition 8.14 **μ-messbare Menge**
Eine Menge $A \subset \mathbb{R}^d$ heißt μ-messbar ($A \in \mathcal{M}(\mu)$) genau dann, wenn A eine höchstens abzählbare Vereinigung von $A_n \in \mathcal{M}_F(\mu)$ ist.

Satz 8.15
$\mathcal{M}(\mu)$ ist eine σ-Algebra. μ^* ist σ-additiv auf $\mathcal{M}(\mu)$.

Folgerung. $(\mathbb{R}^d, \mathcal{M}(\mu), \mu^*)$ ist ein Maßraum.

Beweis. Schritt 1: Es seien $A, B \in \mathcal{M}_F(\mu)$, d.h. es gibt Folgen $A_n \xrightarrow{d^*} A$ und $B_n \xrightarrow{d^*} B$ mit $A_n, B_n \in \mathcal{E}$. Dann ist auch $A_n \cup B_n \in \mathcal{E}$. Es gilt

$$d^*(A \cup B, A_n \cup B_n) \leq d^*(A, A_n) + d^*(B, B_n) \quad \to 0$$

Also folgt $A_n \cup B_n \xrightarrow{d^*} A \cup B$. Mit den anderen weiter oben erwähnten Teilmengenbeziehungen folgt

$$\begin{aligned} A_n \setminus B_n &\longrightarrow A \setminus B \\ A_n \cap B_n &\longrightarrow A \cap B \end{aligned}$$

Dies bedeutet, dass $\mathcal{M}_F(\mu)$ ein Ring ist. Betrachte nun $A, B \in \mathcal{M}_F(\mu)$ mit $A \cap B = \emptyset$, $A_n \xrightarrow{d^*} A$, $B_n \xrightarrow{d^*} B$, wobei $A_n, B_n \in \mathcal{E}$.
Aus der oben bewiesenen Ungleichung $|\mu^*(A) - \mu^*(B)| \leq d^*(A, B)$ folgt dann

$$\begin{aligned} \mu^*(A_n \cup B_n) &\longrightarrow \mu^*(A \cup B) \\ \mu^*(A_n \cap B_n) &\longrightarrow \mu^*(A \cap B) \end{aligned}$$

Außerdem gilt für beliebige Elementarmengen die Identität

$$\begin{aligned} \mu^*(A_n \cup B_n) + \mu^*(A_n \cap B_n) &= \mu(A_n \cup B_n) + \mu(A_n \cap B_n) \\ &= \mu(A_n) + \mu(B_n) \\ &= \mu^*(A_n) + \mu^*(B_n). \end{aligned}$$

Betrachte nun den Grenzübergang $n \to \infty$, so folgt

$$\mu^*(A \cup B) + \mu^*(\underbrace{A \cap B}_{=\emptyset}) = \mu^*(A) + \mu^*(B).$$

μ^* ist also additiv auf $\mathcal{M}_F(\mu)$.

Schritt 2: Sei nun $A \in \mathcal{M}(\mu) \;\Leftrightarrow\; A = \cup_{n=1}^{\infty} \tilde{A}_n$ mit $\tilde{A}_n \in \mathcal{M}_F(\mu)$. Definiere die Folge

$$A'_n = \underbrace{\underbrace{\left(\bigcup_{k=1}^{n} \tilde{A}_k \right)}_{\in \mathcal{M}_F(\mu)} \setminus \underbrace{\left(\bigcup_{k=1}^{n-1} \tilde{A}_n \right)}_{\in \mathcal{M}_F(\mu)}}_{\in \mathcal{M}_F(\mu)}.$$

Dabei gilt, dass die A'_n paarweise disjunkt sind, was sich unmittelbar aus ihrer Konstruktion ergibt. Außerdem gilt $\bigcup_{n\in\mathbb{N}} A'_n = A$. Da μ^* eine σ-subadditive Funktion ist folgt, dass $\mu^*(A) \le \sum_{k=1}^{\infty} \mu^*(A'_k)$. Zudem ist μ^* monoton und da $A \supset A'_1 \cup \cdots \cup A'_n$ gilt, folgt

$$\mu^*(A) \ge \mu^*(A'_1 \cup \cdots \cup A'_n) \overset{\text{Schritt 1}}{=} \sum_{k=1}^{n} \mu^*(A'_k)$$

Falls nun $\mu^*(A) < \infty$ ist, folgt, dass $\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(A'_k)$ (*). Dies gilt für beliebige disjunkte $A'_k \in \mathcal{M}_F(\mu)$ mit $A = \bigcup_{k=1}^{\infty} A'_k$.

Schritt 3: Sei nun $A \in \mathcal{M}(\mu)$ mit endlichem Maß, d.h. $\mu^*(A) < \infty$. Wir zeigen, dass daraus $A \in \mathcal{M}_F(\mu)$ folgt.

A lässt sich schreiben als $A = \bigcup_{n=1}^{\infty} \tilde{A}_n = \bigcup_{n=1}^{\infty} A'_n$ mit $\tilde{A}_n \in \mathcal{M}_F(\mu)$ und paarweise disjunkten $A'_n \in \mathcal{M}_F(\mu)$, wobei wir die A'_n wie oben konstruieren. Setze nun $B_n = \bigcup_{k=1}^{n} A'_k = \bigcup_{k=1}^{n} \tilde{A}_n \subset A$. Wir wissen dann, dass

$$\begin{aligned}
d^*(B_n, A) &= \mu^*(A \bigtriangleup B_n) \\
&= \mu^*(A \setminus B_n) \\
&= \mu^*\left(\bigcup_{k=n+1}^{\infty} A'_k \right) \\
&\le \sum_{k=n+1}^{\infty} \mu^*(A'_k) \qquad \sigma\text{-Subadditivität} \\
&\overset{n\to\infty}{\longrightarrow} 0. \qquad \text{wg. Konvergenz in } (*)
\end{aligned}$$

Daraus folgt dann, dass

$$B_n = \bigcup_{k=1}^{n} \tilde{A}_k = \bigcup_{k=1}^{n} A'_k \xrightarrow{d^*} A.$$

Da $B_n \in \mathcal{M}_F(\mu)$ eine endliche Vereinigung finit messbarer Mengen ist folgt

$$\forall_{\delta>0}\ \exists\ \tilde{B}_n^{\delta} \in \mathcal{E} \qquad d^*(\tilde{B}_n^{\delta}, B) \le 2^{-n}\delta.$$

$$\begin{aligned}
\Rightarrow\ d^*(\tilde{B}_n^{\delta}, A) &\le d^*(\tilde{B}_n^{\delta}, B_n) + d^*(B_n, A) \\
&\le 2^{-n}\delta + d^*(B_n, A) \\
&\overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Dies ist gleichbedeutend mit $A \in \mathcal{M}_F(\mu)$.

Schritt 4: Wir wollen nun zeigen, dass μ^* auf $\mathcal{M}(\mu)$ eine σ-additive Funktion ist.
Sei $A \in \mathcal{M}(\mu)$ und es gelte $A = \bigcup_{n=1}^{\infty} \tilde{A}_n$ mit $\tilde{A}_n \in \mathcal{M}(\mu)$. Falls $\exists_{n=n_0} \mu^*(\tilde{A}_{n_0}) = +\infty$ gilt folgt wegen der Monotonie, dass $\mu^*(A) \geq \mu^*(\tilde{A}_{n_0}) = +\infty$ und daraus wiederum folgt

$$\underbrace{\mu^*(A)}_{=+\infty} = \underbrace{\mu^*(\tilde{A}_{n_0})}_{=+\infty} = \underbrace{\sum_{n\in\mathbb{N}} \mu^*(A_n)}_{=+\infty}.$$

Andernfalls gilt $\forall_n \mu^*(\tilde{A}_n) < \infty$ und daraus folgt $\forall_n \tilde{A}_n \in \mathcal{M}_F(\mu)$.
Nehmen wir nun an, dass $\tilde{A}_n$ paarweise disjunkt sind, so folgt mit Schritt 2), dass $A'_n = \tilde{A}_n$ gilt und daraus folgt

$$\mu^*(A) = \sum_{k=1}^{\infty} \mu^*(\tilde{A}_n).$$

Schritt 5: Wir müssen nun zeigen, dass $\mathcal{M}(\mu)$ wirklich eine σ-Algebra ist.
Seien $\tilde{A}_n \in \mathcal{M}(\mu)$ und $A = \bigcup_{n=1}^{\infty} \tilde{A}_n$. Zu zeigen ist, dass A messbar ist.
Da $\tilde{A}_n \in \mathcal{M}(\mu)$ ist folgt $\exists A'_{n,k} \in \mathcal{M}_F(\mu)$ mit $\tilde{A}_n = \bigcup_{k=1}^{\infty} A'_{n,k}$. Daraus folgt dann, dass $A = \bigcup_{n=1}^{\infty} \tilde{A}_n = \bigcup_{n,k=1}^{\infty} A'_{n,k}$ gilt, also gilt $A \in \mathcal{M}(\mu)$.
Sei nun $B \in \mathcal{M}(\mu) \;\Leftrightarrow\; B = \bigcup_{n=1}^{\infty} B'_n$. Daraus folgt

$$\tilde{A}_n \cap B = \bigcup_{k=1}^{\infty} (\underbrace{\tilde{A}_n \cap B'_k}_{\in \mathcal{M}_F(\mu)}), \text{ da } \mu^*(\tilde{A}_n \cap B'_k) \leq \mu^*(B'_k) < \infty \text{ ist.}$$

Daraus folgt, dass $\tilde{A}_n \cap B'_k \in \mathcal{M}_F(\mu)$ gilt.

$$A \cap B = \bigcup_{n=1}^{\infty} \tilde{A}_n \cap B = \bigcup_{k,n=1}^{\infty} (\underbrace{\tilde{A}_n \cap B'_k}_{\in \mathcal{M}_F(\mu)})$$

Daraus folgt $A \cap B \in \mathcal{M}(\mu)$.
Zu zeigen bleibt noch $A, B \in \mathcal{M}(\mu) \Rightarrow A \setminus B \in \mathcal{M}(\mu)$, $\mathbb{R}^d \in \mathcal{M}(\mu)$. Dies führen wir hier nicht mehr aus. $\square$

Im Zuge des Beweises haben wir gesehen, dass

$$\begin{aligned} A \in \mathcal{M}_F(\mu) \;&\Leftrightarrow\; A \in \mathcal{M}(\mu) \;\wedge\; \mu^*(A) = \mu(A) < \infty \\ &\Leftrightarrow\; \exists_{A_n \in \mathcal{E}} \;:\; \underbrace{d^*(A_n, A)}_{\mu^*(A_n \triangle A)} \overset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

Bemerkung: Die obige Aussage $\mu^*(A) = \mu(A)$ ist so zu lesen, dass für beliebige Mengen A in diesem Fall $\mu(A) := \mu^*(A)$ gesetzt wird, d.h. μ wird durch μ^* gewissermaßen fortgesetzt.

Gegeben sei nun eine offene Menge $A \subset \mathbb{R}^d$ mit $A \neq \emptyset$. Es existiere ein $\varepsilon = \varepsilon(x)$ mit $U_{\varepsilon(x)}(x) \subset A$. In diese Umgebung lässt sich ein Quader mit Seitenlänge $\varepsilon_1(x)$ finden, der ebenfalls vollständig in der Umgebung liegt. Wiederum innerhalb dieses Quaders lässt sich ein $r \in \mathbb{Q}^d$ finden mit $|x - r| < \frac{\varepsilon_1}{8}$. Wählt man nun einen Quader mit der Seitenlänge $\frac{\varepsilon_1(x)}{2}$, welcher r als Mittelpunkt hat, so liegt auch x in diesem Quader, der selbst immer noch komplett in der Umgebung liegt. Daraus folgt, dass eine offene Menge A immer als abzählbare Vereinigung von Würfeln darstellbar ist. Es folgt, dass $A \in \mathcal{M}(\mu)$. Ebenso folgt, dass abgeschlossene Mengen A in $\mathcal{M}(\mu)$ liegen.
Dabei ist wichtig, dass dies noch nicht alle Mengen in $\mathcal{M}(\mu)$ sind. Wähle z.B. G_n offen, F_n abgeschlossen und bilde dann $G^\sigma = \bigcap_{n=1}^\infty G_n$ und $F^\sigma = \bigcup_{n=1}^\infty F_n$. Die Mengen G^σ und F^σ heißen Borel-Mengen vom ersten Typ. Es gilt $G^\sigma, F^\sigma \in \mathcal{M}(\mu)$.

Aus diesem Zusammenhang heraus führt man die Borel-Algebra $\mathcal{B}$ ein, welche die kleinste σ-Algebra ist, die alle offenen Mengen enthält. Da $\mathcal{M}(\mu)$ bezüglich abzählbarer Operationen abgeschlossen ist folgt aus der Konstruktion, dass $\mathcal{B} \in \mathcal{M}(\mu)$ gelten muss. Man beachte, dass die Inklusion echt ist, es gilt also $\mathcal{B} \neq \mathcal{M}(\mu)$.
Betrachte $A \in \mathcal{M}(\mu)$ mit $\mu(A) < \infty$ und $A_n \in \mathcal{E}$ mit $d^*(A_n, A) \to 0$. μ sei regulär, G_n offen, F_n abgeschlossen und es gelte $F_n \subset A \subset G_n$. Dann gilt

$$\begin{aligned} d^*(G_n, A_n) &= \mu^*(G_n \setminus A_n) < 2^{-n} \\ d^*(F_n, A_n) &= \mu^*(A_n \setminus F_n) < 2^{-n} \end{aligned}$$

Es ist $A \subset G^\sigma = \bigcap_{n=1}^\infty G_n$ und es folgt $\mu^*(G^\sigma \setminus A) = \mu(G^\sigma \setminus A) = 0$. Entsprechend ist $\bigcup_{n=1}^\infty F_n = F^\sigma \subset A$ und man kann zeigen, dass auch hier $\mu(A \setminus F^\sigma) = 0$ ist. Insgesamt erhalten wir also Borel-Mengen mit $F^\sigma \subset A \subset G^\sigma$, so dass das Maß der Differenz Null ist.
In $\mathcal{M}(\mu)$ ist also jede Teilmenge einer messbaren Menge mit Maß Null selbst wieder messbar. In der Borel-Algebra $\mathcal{B}$ gilt dies im Allgemeinen allerdings **nicht**.

8.3 Messbare Funktionen

Gegeben sei ein Maßraum $(X, \mathcal{R}, \mu)$ mit einer σ-Algebra $\mathcal{R}$ sowie eine Funktion $f : X \to \{-\infty\} \cup \mathbb{R} \cup \{+\infty\}$. Wir hatten bereits definiert, dass eine Menge $A \subset X$ messbar heißt, wenn $A \in \mathcal{R}$ gilt. Wir wollen diesen Begriff nun auf Funktionen ausweiten:

Definition 8.16 **Messbare Funktion**
Die Funktion f heißt messbar $:\Leftrightarrow$

$$\forall_{a \in \mathbb{R}}\ E_a(f) = \{x \in X \mid f(x) > a\} \text{ ist messbar.}$$

Satz 8.17

All die folgenden Aussagen sind äquivalent:

- $\forall_{a \in \mathbb{R}}\ \{x \mid f(x) > a\}$ messbar
- $\forall_{a \in \mathbb{R}}\ \{x \mid f(x) \geq a\}$ messbar
- $\forall_{a \in \mathbb{R}}\ \{x \mid f(x) < a\}$ messbar
- $\forall_{a \in \mathbb{R}}\ \{x \mid f(x) \leq a\}$ messbar

Beweis. **(1) $\Rightarrow$ (2):** $\{x \mid f(x) \geq a\} = \bigcap_{n \in \mathbb{N}} \{x \mid f(x) > a - \frac{1}{n}\}$, man schneidet also $E_{a-\frac{1}{n}}(f)$.
(2) $\Rightarrow$ (3): $\{x \mid f(x) < a\} = X \setminus \{x \mid f(x) \geq a\}$, man bildet also das Komplement von X mit einer messbaren Menge.
Die anderen Fälle gehen analog und werden hier nicht bewiesen. $\square$

Folgerung. $I \subset \mathbb{R}$ sei ein Intervall. Dann gilt f messbar $\Rightarrow$ $f^{-1}(I)$ messbar.

Jede offene Menge in $\mathbb{R}$ ist eine disjunkte Vereinigung von höchstens abzählbar vielen offenen Intervallen I_k. Daraus folgt

$$f^{-1}(A) = \bigcup_{k \in \mathbb{N}} f^{-1}(I_k) \text{ für } A = \bigcup_{k \in \mathbb{N}} I_k \text{ offen.}$$

Daraus folgt, dass $f^{-1}(A)$ für offene Mengen A messbar ist. Daraus kann man folgern, dass sogar für alle Mengen A, die sich aus abzählbar vielen Vereinigungen bilden lassen, das Urbild messbar ist. Dies sind gerade die Borel-Mengen, d.h. es gilt $f^{-1}(A)$ messbar für Borel-Mengen $A \subset \mathbb{R}$.

Satz 8.18

f messbar $\Rightarrow$ $|f|$ messbar.

Beweis.

$$\{x \mid |f(x)| > a\} = \underbrace{\{x \mid f(x) > a\}}_{\text{messbar}} \cup \underbrace{\{x \mid f(x) < -a\}}_{\text{messbar}}$$

Damit ist der Satz bereits bewiesen. $\square$

Satz 8.19

Sei $\{f_n\}$ eine Folge messbarer Funktionen und definiere

$$g(x) := \sup_{n \in \mathbb{N}} f_n(x) \qquad \left(\inf_{n \in \mathbb{N}} f_n(x)\right)$$

$$h(x) := \limsup_{n \to \infty} f_n(x) \qquad \left(\liminf_{n \to \infty} f_n(x)\right)$$

Dann folgt, dass g, h messbare Funktionen sind.

Beweis. Beweis für den Fall „sup“:

$$\{x \mid g(x) > a\} = \underbrace{\bigcup_{n=1}^{\infty} \{x \mid f_n(x) > a\}}_{\text{messbar}}$$

Daraus folgt, dass $\{x \mid g(x) > a\}$ messbar ist. Der Fall „inf“ wird zur Übung überlassen.
Beweis für den Fall „lim sup“:

$$h(x) = \underbrace{\underbrace{\lim_{n\to\infty}}_{\inf_{n\in\mathbb{N}}} \underbrace{\sup_{k\geq n} f_k(x)}_{h_n(x) \text{ messbar}}}_{\text{messbar}}$$

Auch hier wird der andere Fall zur Übung überlassen. Damit ist der Satz bewiesen. □

Folgerung. Ist $\{f_n\}$ eine Folge messbarer Funktionen und gilt $\forall_{x\in X}$, dass $\exists\ f(x) = \lim_{n\to\infty} f_n(x)$, so folgt, dass f messbar ist.

Folgerung. Ist f messbar, so sind auch $f^+(x) = \max\{0, f(x)\}$ und $f^-(x) = -\min\{0, f(x)\}$ messbare Funktionen und es gilt $f = f^+ - f^-$ und $f^\pm \geq 0$. Außerdem ist auch $-f$ eine messbare Funktion.

Satz 8.20
Seien f, g messbare Funktionen. Dann sind auch $f+g, f-g, f\cdot g$ und $c\cdot f$ für $c \in \mathbb{R}$ messbar. Ebenso ist die Menge $\{x \mid f(x) < g(x)\}$ messbar, sowie alle Varianten mit $\leq, >, \geq$ statt dem „$<$“.

Folgerung. Sei p ein Polynom mit reellen Koeffizienten. Ist f messbar, so ist auch $p(f(x))$ messbar.

Treppenfunktionen Sei $t : X \to \{c_1, c_2, \ldots, c_N\} \subset \mathbb{R}$ mit $c_j \neq c_k$ für $j \neq k$, d.h., dass t nur einen endlichen Wertevorrat besitzt. Für eine Teilmenge $E \subset X$ führen wir die charakteristische Funktion ein:

$$\chi_E(x) = \begin{cases} 1 & x \in E \\ 0 & x \notin E \end{cases}$$

Dann ist E messbar genau dann, wenn auch χ_E messbar ist. Es ist

$$t(x) = \sum_{k=1}^{N} c_k \cdot \chi_{E_k}(x),$$

wobei man $E_i := \{x \in X \mid t(x) = c_i\}$ $(i = 1, \ldots, n)$ setzt. Die Funktion t ist dann genau dann messbar, wenn alle E_k messbar sind.

Satz 8.21
$f : X \to \mathbb{R}^*$ sei messbar. Dann existiert eine Folge von Treppenfunktionen t_n mit

$$\forall_{x \in X}\ t_n(x) \overset{n\to\infty}{\longrightarrow} f(x)$$

Falls $f \geq 0$ ist kann man t_n sogar monoton wählen, d.h. für jedes x ist $\{t_n(x)\}_{n\in\mathbb{N}}$ in n monoton wachsend.

Beweis. Sei $f \geq 0$. Für festes $n \in \mathbb{N}$ definieren wir

$$E_{ni} := \left\{ x \ : \ \frac{i-1}{2^n} \leq f(x) \leq \frac{i}{2^n} \right\} \qquad , i = 1, \dots, n2^n$$
$$F_n := \{x \ : \ f(x) \geq n\}$$

Diese Mengen sind paarweise disjunkt und messbar. Dann ist

$$t_n(x) = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{E_{ni}}(x) + n\chi_{F_n}(x)$$

ebenfalls messbar. Für die Konvergenz fixiere ein $x \in X$ mit $f(x) < \infty$ und wähle $n \in \mathbb{N}$ mit $n \geq f(x)$. Dann folgt $|t_n(x) - f(x)| \leq 2^{-n} \overset{n\to\infty}{\longrightarrow} 0$. Falls $f(x) = \infty$ ist, so folgt $t_n(x) = n \overset{n\to\infty}{\longrightarrow} \infty$.
Für die Monotonie sei nun $x \in E_{ni}$. Dann folgt $t_n(x) = \frac{i-1}{2^n}$ und $f(x) \geq \frac{i-1}{2^n} = \frac{2(i-1)}{2^{n+1}}$, also ist $t_{n+1}(x) \geq t_n(x)$. Für $x \in F_n$ ist $t_n(x) = n$ und $f(x) \geq n = \frac{2^{n+1}n}{2^{n+1}}$, also ist $t_{n+1}(x) \geq t_n(x)$.
Wir kommen nun noch zum allgemeinen Fall, d.h. dass f nicht notwendigerweise nicht-negativ ist. Dann zerlege $f(x) = f^+(x) - f^-(x)$ mit $f^\pm(x) = \frac{|f| \pm f}{2} \geq 0$. Dann erhält man Treppenfunktionen $t_n^\pm(x)$ für $f^\pm$ wie oben und kann dementsprechend $t_n(x) = t_n^+(x) - t_n^-(x)$ schreiben, wobei $t_n^+(x) \uparrow f^+(x)$ und $t_n^-(x) \downarrow f^-(x)$ ist, also konvergiert $t_n(x) \to f^+(x) - f^-(x) = f(x)$. □

8.4 Das Lebesgue-Integral

Sei $(X, \mathcal{R}, \mu)$ ein Maßraum und $t(x) = \sum_{i=1}^n c_i \chi_{E_i}(x)$ eine messbare Treppenfunktion. Es sei

$$I_E(t) = \sum_{i=1}^n c_i \mu(E_i \cap E) \qquad \text{mit } E \in \mathcal{R}.$$

Definition 8.22 Lebesgue-Integral nicht-negativer Funktionen
Sei $f : X \to \mathbb{R}^*$, $f \geq 0$ messbar. Wir definieren

$$\int_E f \, \mathrm{d}\mu := \sup_{\substack{t \text{ Treppenfkt.} \\ 0 \leq t \leq f}} I_E(t).$$

Dies ist das Lebesgue-Integral von f über E mit Maß μ.

Bemerkung: Falls $f = t$ selbst eine Treppenfunktion ist gilt $\int_E f \, \mathrm{d}\mu = I_E(f)$.

Definition 8.23 Lebesgue-Integral
Sei $f : X \to \mathbb{R}^*$ messbar. Wir definieren nun

$$\int_E f \, \mathrm{d}\mu := \int_E f^+ \, \mathrm{d}\mu - \int_E f^- \, \mathrm{d}\mu,$$

falls eines der Integrale $\int_E f^\pm \, \mathrm{d}\mu$ endlich ist.

Definition 8.24 Integrierbarkeit
Eine messbare Funktion f heißt integrierbar, falls beide Integrale $\int_E f^\pm \, \mathrm{d}\mu$ endlich sind. Man schreibt dann, dass $f \in \mathcal{L}(E, \mu)$.

Eigenschaften:

1. f messbar und beschränkt, $\mu(E) < \infty$, so ist $f \in \mathcal{L}(E, \mu)$.
2. f messbar, $\mu(E) < \infty$, $a \leq f(x) \leq b \ \forall x \in X$, so ist $a\mu(E) \leq \int_E f \, \mathrm{d}\mu \leq b\mu(E)$.
3. $f, g \in \mathcal{L}(E, \mu)$, $f(x) \leq g(x)$ für alle $x \in X$, dann ist $\int_E f \, \mathrm{d}\mu \leq \int_E g \, \mathrm{d}\mu$.
4. $f \in \mathcal{L}(E, \mu)$, $c \in \mathbb{R} \ \Rightarrow \ (cf) \in \mathcal{L}(E, \mu)$ und $\int_E (cf) \, \mathrm{d}\mu = c \int_E f \, \mathrm{d}\mu$.
5. $\mu(E) = 0$, f messbar, dann ist $\int_E f \, \mathrm{d}\mu = 0$.
6. $f \in \mathcal{L}(E, \mu)$, $A \in \mathcal{R}$ und $A \subset E$, dann ist $f \in L(A, \mu)$.

Satz 8.25 Sigma-Additivität des Integrals
Sei $f : X \to \mathbb{R}^*$ messbar, $f \geq 0$ und für $A \subset \mathcal{R}$ sei $\varphi(A) = \int_A f \, \mathrm{d}\mu$. Dann ist

1. φ ist σ-additiv.
2. Falls f nicht definit, dafür aber $f \in \mathcal{L}(X, \mu)$ ist, so gilt 1.) ebenfalls.

Beweis. 1.) $\Rightarrow$ 2.): Sei $f \in \mathcal{L}(X, \mu)$ und $f(x) = f^+(x) - f^-(x)$. Dann folgt

$$\varphi(A) = \int_A f \, \mathrm{d}\mu = \int_A f^+ \, \mathrm{d}\mu - \int_A f^- \, \mathrm{d}\mu = \varphi^+(A) - \varphi^-(A) \qquad , \sigma\text{-additiv.}$$

Wir zeigen nun also 1.). **Erster Schritt:** Sei $f = \chi_E$ eine charakteristische Funktion. Dann ist

$$\varphi(A) = \int_A \chi_E \, \mathrm{d}\mu = \mu(E \cap A).$$

Sei $A = \bigcup_{j=1}^\infty A_j$ eine paarweise disjunkte Vereinigung. Dann folgt

$$\begin{aligned} \varphi(A) &= \mu\left(E \cap \bigcup_{j=1}^\infty A_j\right) = \sum_{j=1}^\infty \mu(E \cap A_j) \\ &= \sum_{j=1}^\infty \int_{A_j} \chi_E \, \mathrm{d}\mu = \sum_{j=1}^\infty \varphi(A_j). \end{aligned}$$

Zweiter Schritt: Sei nun $f = t = \sum_{k=1}^{n} c_k \chi_{E_k}$ eine Treppenfunktion. Dann folgt die Aussage unmittelbar aus dem ersten Schritt, da dies nur eine endliche Summe ist.
Dritter Schritt: f sei messbar und $f \geq 0$. Sei t Treppenfunktion mit $0 \leq t \leq f$.

$$I_A(t) = \int_A t \,\mathrm{d}\mu = \sum_{j=1}^{\infty} \int_{A_j} t \,\mathrm{d}\mu \leq \sum_{j=1}^{\infty} \int_{A_j} f \,\mathrm{d}\mu = \sum_{j=1}^{\infty} \varphi(A_j)$$

Damit folgt dann

$$\int_A f \,\mathrm{d}\mu = \varphi(A) = \sup_{0 \leq t \leq f} I_A(t) \leq \sum_{j=1}^{\infty} \varphi(A_j).$$

Umgekehrt gilt: Nach Definition von $\int_A f \mathrm{d}\mu$ existiert für jedes $\varepsilon > 0$ und jedes $A = A_1 \cup A_2$ mit $A_1 \cap A_2 = \emptyset$ eine Treppenfunktion $0 \leq t \leq f$, so dass

$$\int_{A_1} t \,\mathrm{d}\mu \geq \int_{A_1} f \,\mathrm{d}\mu - \varepsilon$$
$$\int_{A_2} t \,\mathrm{d}\mu \geq \int_{A_2} f \,\mathrm{d}\mu - \varepsilon$$

gilt. Also folgt

$$\begin{aligned}\varphi(A) = \int_A f \,\mathrm{d}\mu \geq \int_A t \,\mathrm{d}\mu &= \int_{A_1} t \,\mathrm{d}\mu + \int_{A_2} t \,\mathrm{d}\mu \\ &\geq \int_{A_1} f \,\mathrm{d}\mu + \int_{A_2} f \,\mathrm{d}\mu - 2\varepsilon.\end{aligned}$$

Für $\varepsilon \to 0$ folgt dann $\varphi(A) \geq \varphi(A_1) + \varphi(A_2)$. Dieser Fall lässt sich unmittelbar auf den Fall

$$\varphi\left(\bigcup_{j=1}^{n} A_j\right) \geq \sum_{j=1}^{n} \varphi(A_j) \qquad \forall_{n \in \mathbb{N}}$$

übertragen. Also folgt

$$\varphi\left(\bigcup_{j=1}^{\infty} A_j\right) \geq \varphi\left(\bigcup_{j=1}^{n} A_j\right) \geq \sum_{j=1}^{n} \varphi(A_j) \qquad \forall_{n \in \mathbb{N}}$$

Und für $n \to \infty$ damit

$$\varphi(A) = \varphi\left(\bigcup_{j=1}^{\infty} A_j\right) \geq \sum_{j=1}^{\infty} \varphi(A_j).$$

Damit ist der Satz bewiesen. $\square$

Folgerungen:

- f messbar, $f \geq 0$, $A, B \in \mathcal{R}$, $B \subset A$, so folgt $\int_B f \, \mathrm{d}\mu \leq \int_A f \, \mathrm{d}\mu$.
- f messbar, $A, B \in \mathcal{R}$, $B \subset A$, $\mu(A \setminus B) = 0$, so folgt $\int_A f \, \mathrm{d}\mu = \int_B f \, \mathrm{d}\mu$.

Beispiel: Es sei zum Beispiel

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

Es ist $\mu(\mathbb{Q}) = \sum_{n=1}^{\infty} \mu(\{q\}) = 0$, $(q_n)_{n\in\mathbb{N}}$ eine Abzählung in $\mathbb{Q}$, wobei μ hier das Lebesgue-Maß bezeichnet. Damit folgt dann

$$\int_{\mathbb{R}} f \, \mathrm{d}\mu = \int_{\mathbb{R}\setminus\mathbb{Q}} f \, \mathrm{d}\mu = 0.$$

Definition 8.26 **Fast überall**

Sei $H(\cdot)$ eine Aussageform. Man sagt, H gilt μ-fast überall (μ-f.ü.), genau dann wenn es eine Menge $E \in \mathcal{R}$ mit $\mu(E) = 0$ gibt, so dass $E \supset \{x \; : \; \neg H(x)\}$.

Satz 8.27

Sei $f \in \mathcal{L}(X, \mu)$, dann folgt $|f| \in \mathcal{L}(X, \mu)$ und $\left|\int_X f \, \mathrm{d}\mu\right| \leq \int_X |f| \, \mathrm{d}\mu$.

Beweis. Sei $f(x) = f^+(x) - f^-(x)$, $|f| = f^+ + f^-$. Da $f \in \mathcal{L}(X, \mu)$ ist folgt, dass auch $f^{\pm} \in \mathcal{L}(X, \mu)$ gilt. Sei nun $A = \{x \; : \; f(x) \geq 0\}$ und $B = \{x \; : \; f(x) < 0\}$. Es ist $A \cap B = \emptyset$ und $X = A \cup B$. Ferner ist $f^+(x) = 0$ auf B und $f^-(x) = 0$ auf A.

$$\begin{aligned} \int_X |f| \, \mathrm{d}\mu &= \int_A |f| \, \mathrm{d}\mu + \int_B |f| \, \mathrm{d}\mu = \int_A f^+ \, \mathrm{d}\mu + \int_B f^- \, \mathrm{d}\mu \\ &= \int_X f^+ \, \mathrm{d}\mu + \int_X f^- \, \mathrm{d}\mu < \infty \end{aligned}$$

Außerdem gilt $|f| \geq f$, $|f| \geq -f$, also folgt

$$\int_X |f| \, \mathrm{d}\mu \geq \int_X f \, \mathrm{d}\mu \text{ und } \int_X |f| \, \mathrm{d}\mu \geq \int_X -f \, \mathrm{d}\mu \qquad \text{, also folgt}$$
$$\int_X |f| \, \mathrm{d}\mu \geq \left|\int_X f \, \mathrm{d}\mu\right|$$

Damit ist der Satz bewiesen. □

Folgerungen:

- $g \in \mathcal{L}(E, \mu)$, f messbar, $|f(x)| \leq g(x)$ μ-fast überall, dann folgt $f \in \mathcal{L}(E, \mu)$ und $|f| \in \mathcal{L}(E, \mu)$.
- Sei $f \in \mathcal{L}(X, \mu)$, dann folgt, dass f μ-fast überall endlich ist.

8.5 Das Lebesgue-Integral und Konvergenz

$(X, \mathcal{R}, \mu)$ sei ein Maßraum.

Satz 8.28 Satz zur monotonen Konvergenz (Lebesgue)
Sei $f_n : X \to \mathbb{R}^*$ eine Folge messbarer Funktionen. Weiter gelte $0 \leq f_1(x) \leq f_2(x) \leq \cdots$, d.h. für festes x sei die Folge monoton in n. Es sei $f(x) := \lim_{n\to\infty} f_n(x) \geq 0$.

$$\Rightarrow \lim_{n\to\infty} \int_X f_n \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\mu.$$

Beweis. Sei $\tilde{X} = \{x \in X \mid f(x) > 0\}$. Daraus folgt, dass $\int_X f(x) \, \mathrm{d}\mu = \int_{\tilde{X}} f(x) \, \mathrm{d}\mu$, denn $\tilde{X}$ ist messbar, da f messbar, ihr Komplement ist aber eine Nullmenge. Weiterhin gilt $0 \leq f_n(x) \leq f(x)$ (für alle $n \in \mathbb{N}$), d.h. $f_n(x) = 0$ für alle $x \in X \setminus \tilde{X}$ und $\int_X f_n \, \mathrm{d}\mu = \int_{\tilde{X}} f_n \, \mathrm{d}\mu$. Wir können all diese Eigenschaften also nutzen, um von Anfang an nur auf der Menge $\tilde{X}$ zu operieren.
Schritt 1: Aus $0 \leq f_n(x) \leq f(x)$ folgt $\int_{\tilde{X}} f_n \, \mathrm{d}\mu \leq \int_{\tilde{X}} f \, \mathrm{d}\mu$ (Monotonie des Lebesgie-Integrals in der Funktion). Sei $\alpha := \sup_{n\in\mathbb{N}} \int_{\tilde{X}} f_n \, \mathrm{d}\mu \leq \int_{\tilde{X}} f \, \mathrm{d}\mu$. Nun ist aber gerade $\alpha = \lim_{n\to\infty} \int_{\tilde{X}} f \, \mathrm{d}\mu$. Damit haben wir den Fall „$\leq$" bewiesen.
Schritt 2: Sei $c \in]0,1[$. Wir betrachten messbare Treppenfunktionen t, für die $0 \leq t(x) \leq f(x)$ gilt. Wir bilden nun die Menge $E_n = \{x \in X \mid f_n(x) \geq ct(x)\}$ $(n \in \mathbb{N})$. Da $f_n \uparrow$ (in n) ist, gilt $E_1 \subset E_2 \subset \cdots$. Betrachte $\tilde{X} \cap E_n = \tilde{E}_n$, diese Menge ist immer noch monoton (d.h. $\tilde{E}_1 \subset \tilde{E}_2 \subset \cdots$). Da $\lim_{n\to\infty} f_n(x) = f(x)$ gilt, folgt $\bigcup_{n\in\mathbb{N}} \tilde{E}_n = \tilde{X}$. Nehmen wir nämlich an, dies wäre nicht so, so muss ein $p \in \tilde{X}$ existieren mit $p \notin \bigcup_{n\in\mathbb{N}} \tilde{E}_n$. Daraus folgt aber, dass für alle $n \in \mathbb{N}$ gelten muss, dass $f_n(p) < ct(p)$ ist ($f_n(p) = 0$ kann ausgeschlossen werden, da wir uns auf $\tilde{X}$ bewegen). Daraus folgt, dass $f(p) = \sup_{n\in\mathbb{N}} f_n(p) \leq ct(p) \leq cf(p)$ ist, d.h. $f(p) \leq cf(p)$, aber $0 < f(p) \leq cf(p)$ steht im Widerspruch zu $c \in]0,1[$. $\lightning$
Sei nun $\varphi(E) = \int_E t \, \mathrm{d}\mu$. Wir wissen, dass dies eine σ-additive Funktion ist. Außerdem ist $\varphi(\tilde{E}_n) \xrightarrow{n\to\infty} \varphi(\tilde{X})$. Weiterhin ist

$$c\varphi(\tilde{E}_n) = c \int_{\tilde{E}_n} t \, \mathrm{d}\mu \leq \int_{\tilde{E}_n} f_n \, \mathrm{d}\mu \leq \int_{\tilde{X}} f_n \, \mathrm{d}\mu$$

Für den Grenzübergang $n \to \infty$ erhalten wir damit:

$$c\varphi(\tilde{X}) = c \int_{\tilde{X}} t \, \mathrm{d}\mu \leq \int_{\tilde{X}} f_n \, \mathrm{d}\mu$$

Da $c \in]0,1[$ beliebig war, können wir $c \to 1$ betrachten. Daraus folgt dann $\int_{\tilde{X}} t \, \mathrm{d}\mu \leq \int_{\tilde{X}} f_n \, \mathrm{d}\mu$. Außerdem können wir das sup über alle Treppenfunktionen $0 \leq t \leq f(x)$ betrachten und erhalten $\int_{\tilde{X}} f \, \mathrm{d}\mu \leq \int_{\tilde{X}} f_n \, \mathrm{d}\mu$. Bildet man nun den Grenzwert, so erhält man

$$\int_{\tilde{X}} f \, \mathrm{d}\mu \leq \lim_{n\to\infty} \int_{\tilde{X}} f_n \, \mathrm{d}\mu.$$

Insgesamt haben wir damit Gleichheit und der Satz ist bewiesen. □

Satz 8.29 Additivität bzgl. des Integranden
Seien $f_1, f_2 \in \mathcal{L}(X, \mu)$. Dann folgt, dass $f_1 + f_2 \in \mathcal{L}(X, \mu)$ und

$$\int_X (f_1 + f_2)\ \mathrm{d}\mu = \int_X f_1\ \mathrm{d}\mu + \int_X f_2\ \mathrm{d}\mu.$$

Beweis. Falls f_1, f_2 Treppenfunktionen sind ist die Aussage trivial. Wir nehmen nun an, dass $f_1, f_2 \geq 0$ beliebige, messbare Funktionen sind. Wir wissen dann, dass es monotone Folgen von Treppenfunktionen $\{t_n'\} \uparrow, \{t_n''\} \uparrow$ mit $0 \leq t_n'(x) \overset{n\to\infty}{\longrightarrow} f_1(x)$ und $0 \leq t_n''(x) \overset{n\to\infty}{\longrightarrow} f_2(x)$ gibt. Sei nun $t_n(x) = t_n'(x) + t_n''(x)$. Dann folgt $0 \leq t_n(x) \overset{n\to\infty}{\longrightarrow} f(x) = f_1(x) + f_2(x)$. Wenden wir nun den Satz von Lebesgue zur monotonen Konvergenz an, so folgt

$$\begin{aligned}\int_X f\ \mathrm{d}\mu &= \lim_{n\to\infty} \int_X t_n\ \mathrm{d}\mu = \lim_{n\to\infty} \int_X (t_n' + t_n'')\ \mathrm{d}\mu \\ &= \lim_{n\to\infty} \left(\int_X t_n'\ \mathrm{d}\mu + \int_X t_n''\ \mathrm{d}\mu \right)\end{aligned}$$

und mit dem S.v.L. zur monotonen Konvergenz folgt auch die Konvergenz der Summanden:

$$= \lim_{n\to\infty} \int_X t_n'\ \mathrm{d}\mu + \lim_{n\to\infty} \int_X t_n''\ \mathrm{d}\mu = \int_X f_1\ \mathrm{d}\mu + \int_X f_2\ \mathrm{d}\mu$$

Sind beide Funktionen nicht-positiv, so kann man das gemeinsame Vorzeichen herausziehen und analog argumentieren. Betrachte nun also $f_1 \geq 0$, $f_2 \leq 0$ messbare Funktionen und $A = \{x \mid f(x) \geq 0\}$, $B = \{x \mid f(x) < 0\}$. Auf A sind daher $f, f_1, -f_2 \geq 0$. Dann folgt

$$\begin{aligned}\int_A f_1\ \mathrm{d}\mu &= \int_A (f - f_2)\ \mathrm{d}\mu = \int_A (f + (-f_2))\ \mathrm{d}\mu = \int_A f\ \mathrm{d}\mu + \int_A (-f_2)\ \mathrm{d}\mu \\ &= \int_A f\ \mathrm{d}\mu - \int_A f_2\ \mathrm{d}\mu.\end{aligned}$$

Auf der Menge B können wir nahezu analog verfahren. Addiert man diese beiden Resultate, so folgt wegen $A \cap B = \emptyset$, $A \cup B = X$ das, was wir wollen.
Es bleibt nun noch der Fall, dass sich f_1 und f_2 irgendwie verhalten. Zerlege hierfür $X = E_{++} \cup E_{+-} \cup E_{-+} \cup E_{--}$ mit $E_{++} = \{x \mid f_1(x) \geq 0 \ \wedge\ f_2(x) \geq 0\}$, $E_{+-} = \{x \mid f_1(x) \geq 0 \ \wedge\ f_2(x) < 0\}$ usw. Damit folgt der Satz dann für beliebige, messbare Funktionen. □

Satz 8.30
Sei $f_k \geq 0$ messbar, $f(x) = \sum_{k=1}^{\infty} f_k(x)$. Dann folgt

$$\int_X f\ \mathrm{d}\mu = \sum_{k=1}^{\infty} \int_X f_k\ \mathrm{d}\mu.$$

Beweis. Sei $S_n(x) = \sum_{k=1}^{n} f_k(x)$. Es ist $S_n(x) \uparrow$. Außerdem wissen wir, dass $\int_X S_n(x)\ \mathrm{d}\mu = \sum_{k=1}^{n} \int_X f_k(x)\ \mathrm{d}\mu$ (endliche Additivität). Betrachte nun den Grenzwert

$$\underbrace{\lim_{n\to\infty} \int_X S_n(x)\ \mathrm{d}\mu}_{\int_X f(x)\ \mathrm{d}\mu \text{ nach S.v.L. mon. Konv.}} = \underbrace{\lim_{n\to\infty} \sum_{k=1}^{n} \int_X f_k(x)\ \mathrm{d}\mu}_{\sum_{k=1}^{\infty} \int_X f_k(x)\ \mathrm{d}\mu}$$

Damit ist der Satz bewiesen. $\square$

Satz 8.31 Lemma von Fatou
Sei $f_n \geq 0$ eine Folge messbarer Funktionen. Wir betrachten punktweise $f(x) = \liminf_{n\to\infty} f_n(x)$.

$$\Rightarrow \int_X f\ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n\ \mathrm{d}\mu.$$

Beweis. Sei $g_n(x) = \inf_{k\geq n} f_k(x)$. Damit ist $0 \leq g_1(x) \leq g_2(x) \leq \cdots$ monoton in n. Außerdem ist für festes n immer $g_n(x) \leq f_n(x)$. Also ist $\int_X g_n\ \mathrm{d}\mu \leq \int_X f_n\ \mathrm{d}\mu$. Nach dem Satz von Lebesgue zur monotonen Konvergenz ist $\int_X g_n\ \mathrm{d}\mu \overset{n\to\infty}{\longrightarrow} \int_X f\ \mathrm{d}\mu$. Da diese Seite konvergiert ist „$\lim = \liminf = \limsup$". Wir können auf der rechten Seite also einfach den $\liminf$ betrachten und erhalten

$$\int_X f\ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n\ \mathrm{d}\mu.$$

Damit ist der Satz bewiesen. $\square$

Aufgabe. Finden Sie ein Beispiel für das wirklich „$<$" gilt.

Satz 8.32 Satz zur majorisierten Konvergenz (Lebesgue)
Sei $f_n(x) \overset{n\to\infty}{\longrightarrow} f(x)$ für alle $x \in X$ mit messbaren Funktionen f_n. Weiter existiere eine Funktion $g \in \mathcal{L}(X,\mu)$ mit $|f_n(x)| \leq g(x)$ für alle $x \in X$ und $n \in \mathbb{N}$.

$$\Rightarrow \lim_{n\to\infty} \int_X f_n\ \mathrm{d}\mu = \int_X f\ \mathrm{d}\mu,\ f \in \mathcal{L}(X,\mu).$$

Beweis. Wir wissen, dass $0 \leq |f_n(x)| \leq g(x)$ gilt. Aus $\int_X g\mathrm{d}\mu < \infty$ folgt $\int_X |f_n(x)|\ \mathrm{d}\mu < \infty$. Daraus wiederum folgt dann, dass $\int_X f_n^{\pm}(x)\ \mathrm{d}\mu < \infty$ gilt, d.h. $f_n \in \mathcal{L}(X,\mu)$. Weiterhin gilt im Grenzwert $n \to \infty$, dass $0 \leq |f(x)| \leq g(x)$ ist. Daraus folgt $\int_X |f|\ \mathrm{d}\mu < \infty \Rightarrow f \in \mathcal{L}(X,\mu)$.
Die Ungleichung $|f_n(x)| \leq g(x)$ impliziert $0 \leq g(x) + f_n(x)$ und $0 \leq g(x) - f_n(x)$. Für $n \to \infty$ gehen diese über in $0 \leq g(x) + f(x)$ und $0 \leq g(x) - f(x)$. Mit dem Lemma von Fatou folgt dann

$$\underbrace{\int_X (g(x) + f(x))\ \mathrm{d}\mu}_{\int_X g\ \mathrm{d}\mu + \int_X f\ \mathrm{d}\mu} \leq \liminf_{n\to\infty} \int_X (g + f_n)\ \mathrm{d}\mu \leq \int_X g\ \mathrm{d}\mu + \liminf_{n\to\infty} \int_X f_n\ \mathrm{d}\mu,$$

womit wir dann

$$\int_X f \ \mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X f_n \ \mathrm{d}\mu$$

erhalten. Analog folgt aus der zweiten Ungleichung

$$\int_X (g-f)\mathrm{d}\mu \leq \liminf_{n\to\infty} \int_X (g-f_n) \ \mathrm{d}\mu,$$

und mit der Additivität dann:

$$-\int_X f \ \mathrm{d}\mu \leq \underbrace{\liminf_{n\to\infty} \left(-\int_X f_n \ \mathrm{d}\mu\right)}_{-\limsup_{n\to\infty}\int_X f_n \ \mathrm{d}\mu}$$

Und damit erhalten wir dann

$$\limsup_{n\to\infty} \int_X f_n \ \mathrm{d}\mu \leq \int_X f \ \mathrm{d}\mu.$$

Wegen $\liminf \int \leq \limsup \int$ folgt aus diesen Ungleichungen dann letztlich die Behauptung. $\square$

8.6 Das Lebesgue- und das Riemann-Integral

Sei $X = [a,b] \subset \mathbb{R}$ und μ das Lebesgue-Maß auf $\mathbb{R}$. Weiterhin bezeichne $\int_a^b f(x) \ \mathrm{d}x$ das Riemann-Integral und $\int_{[a,b]} f \ \mathrm{d}\mu$ das Lebesgue-Integral. Bevor wir einen Satz über den Zusammenhang dieser Integralbegriffe formulieren benötigen wir zunächst ein Lemma:

Lemma 8.33
Sei $f \geq 0$ messbar und $\int_X f \ \mathrm{d}\mu = 0$, so folgt $f(x) = 0$ μ-fast überall.

Beweis. Sei $E_n = \left\{x \mid f(x) > \frac{1}{n}\right\}$ und $E = \{x \mid f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n$. Aus $\mu(E) > 0$ folgt, dass ein $n_0 \in \mathbb{N}$ existiert mit $\mu(E_{n_0}) > 0$. Daraus wiederum folgt dann

$$\int_X f \ \mathrm{d}\mu \geq \int_{E_{n_0}} f \ \mathrm{d}\mu \geq \frac{1}{n_0} \underbrace{\mu(E_{n_0})}_{>0} > 0 \ \lightning$$

So könnte das Integral nicht Null sein. Damit ist das Lemma bewiesen. $\square$

Satz 8.34 Satz von Lebesgue
Für Riemann- bzw. Lebesgue-Integrale gelten folgende Eigenschaften:

1. $f \in \mathcal{R}([a,b]) \;\Rightarrow\; f \in \mathcal{L}([a,b],\mu)$ und
$$\int_{[a,b]} f \;\mathrm{d}\mu = \int_a^b f(x)\;\mathrm{d}x$$
2. f messbar und $f \in \mathcal{R}([a,b]) \;\Leftrightarrow\; f$ ist auf $[a,b]$ μ-fast überall stetig.

Beweis. Wir betrachten das Intervall $[a,b]$ mit einer Zerlegungsfolge $\delta^{(n)} = \{x_k\}_{k=0}^m$ und $f \geq 0$. Wir erinnern uns, dass das Riemann-Integral über Darboux-Summen eingeführt wurde. Es bezeichne $\Delta_i x = [x_i, x_{i+1}]$, $m_i = \inf_{x\in\Delta_i x} f(x)$ und $M_i = \sup_{x\in\Delta_i x} f(x)$. Weiter sei $s(f,\delta^{(n)}) = \sum_i m_i \Delta x_i = I(f_u^{(n)})$ und $S(f,\delta^{(n)}) = \sum_i M_i\Delta x_i = I(f_o^{(n)})$ mit $\Delta x_i = x_{i+1} - x_i$. Schaut man sich das an, so erhält man

$$\begin{aligned}
f_u^{(n)} &= \sum_i m_i \chi_{[x_i,x_{i+1}[} \\
f_o^{(n)} &= \sum_i M_i \chi_{[x_i,x_{i+1}[} \\
I(f_u^{(n)}) &= \int_{[a,b]} f_u^{(n)}\;\mathrm{d}\mu \\
I(f_o^{(n)}) &= \int_{[a,b]} f_o^{(n)}\;\mathrm{d}\mu.
\end{aligned}$$

Wir wählen nun eine Zerlegungsfolge, welche die Bedingung $\delta^{(n)} \subset \delta^{(n+1)}$ für $n \in \mathbb{N}$ erfüllt. Ferner gelte für den Rang $\lambda(\delta^{(n)}) \to 0$ für $n \to \infty$. Wir haben $f_u^{(n)}(x) \leq f(x) \leq f_o^{(n)}(x)$ und $f_u^{(n)}(x) \uparrow$, $f_o^{(n)}(x) \downarrow$ für $x \in X$ und $n \to \infty$. Daraus folgt, dass punktweise ein Grenzwert existiert mit $f_u^{(n)}(x) \to f_u(x) \leq f(x)$ und $f_o^{(n)}(x) \to f_o(x) \geq f(x)$. Wir haben

$$\begin{aligned}
s(f,\delta^{(n)}) &= \int_{[a,b]} f_u^{(n)}\;\mathrm{d}\mu \overset{n\to\infty}{\longrightarrow} \int_{[a,b]} f_u\;\mathrm{d}\mu \\
S(f,\delta^{(n)}) &= \int_{[a,b]} f_o^{(n)}\;\mathrm{d}\mu \overset{n\to\infty}{\longrightarrow} \int_{[a,b]} f_o\;\mathrm{d}\mu.
\end{aligned}$$

Da $f \in \mathcal{R}([a,b])$ ist folgt, dass die oberen und unteren Darboux-Summen den gleichen Grenzwert besitzen. Damit folgt

$$\begin{aligned}
&\Rightarrow \int_{[a,b]} f_u\;\mathrm{d}\mu = \int_{[a,b]} f_o\;\mathrm{d}\mu \\
&\Rightarrow \int_{[a,b]} (f_o - f_u)\;\mathrm{d}\mu = 0 \\
&\Rightarrow f_o - f_u = 0 \qquad , \mu\text{-fast überall.}
\end{aligned}$$

Die letzte Implikation folgt dabei aus dem vorausgehenden Lemma.
Außerdem ist $f_u(x) \le f(x) \le f_o(x)$ mit $f_u, f_o \in \mathcal{L}([a,b],\mu)$ und $\int_{[a,b]} f_u \, \mathrm{d}\mu = \int_{[a,b]} f_o \, \mathrm{d}\mu = \int_{[a,b]} f \, \mathrm{d}\mu$ und daraus folgt $f \in \mathcal{L}([a,b],\mu)$. Aus dem Obigen folgt zudem $f_u(x) = f(x) = f_o(x)$ (μ-fast überall). Daraus folgt dann

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} s(f,\delta^{(n)}) = \lim_{n\to\infty} S(f,\delta^{(n)}) = \lim_{n\to\infty} \int_{[a,b]} f_u^{(n)}\,\mathrm{d}\mu = \int_{[a,b]} f_u\,\mathrm{d}\mu \overset{\text{S.v.mon.Konv.}}{=} \int_{[a,b]} f\,\mathrm{d}\mu.$$

Wir kommen nun zur zweiten Aussage. Zu zeigen ist $f_u(x) = f_o(x)$ μ-fast überall $\Leftrightarrow$ f ist auf $[a,b]$ μ-fast überall stetig. f_u, f_o können nur dann gleich sein, wenn f stetig ist. $\square$

Anmerkung. Ist $f \in \mathcal{L}([a,b],\mu)$ und $F(x) = \int_{[a,x]} f \, \mathrm{d}\mu$ ($a \le x \le b$), so folgt $\frac{\mathrm{d}F}{\mathrm{d}x} = f$ μ-fast überall. Die Umkehrung gilt mit μ-fast überall nicht! Betrachte z.B.

$$F(x) = \begin{cases} 1 & x < 0 \\ 0 & x \ge 0 \end{cases},$$

dann ist $\frac{\mathrm{d}F}{\mathrm{d}x} = 0$ μ-fast überall.

Zu den Unterschieden zwischen Riemann- und Lebesgue-Integral:
Wir wollen in einer Tabelle nochmal einige der wichtigsten Unterschiede der beiden Integralbegriffe formulieren, die häufig zu Denkfehlern führen:

Riemann-Integral	**Lebesgue-Integral**
gerichtet: $\int_a^b \mathrm{d}x = -\int_b^a \mathrm{d}x$	nicht gerichtet: $\int_{[a,b]} \mathrm{d}\mu$
$f \in \mathcal{R}[a,b]$ $\Rightarrow$ beschränkt	$f \in \mathcal{L}(X,\mu)$ $\Rightarrow$ $\lvert f\rvert$ endlich bis auf Menge von Maß 0
$X = [a,b]$	auch $\mu(X) = \infty$ möglich dabei: $f \in \mathcal{L}(X,\mu)$ $\Rightarrow$ $\{x \mid f(x) \neq 0\}\sigma$-finit (d.h. abzählbare Vereinigung v. Mengen mit endl. Maß)
$\int_{-\infty}^{+\infty} = \lim\limits_{R\to\infty} \int_{-R}^{+R}$ z.B. $\frac{\sin x}{x}$	aber: $\frac{\sin x}{x} \notin \mathcal{L}(\mathbb{R},\mu)$
	Konvergenzsätze: monotone/majorisierte Konvergenz, Lemma von Fatou Voraussetzungen müssen nur μ-f.ü. erfüllt sein

8.7 Die $\mathcal{L}^p$-Funktionsräume

Sei $(X,\mathcal{R},\mu)$ ein Maßraum und $f,g : X \to \mathbb{R}^*$ messbare Funktionen. Wir führen die folgende Äquivalenzrelation ein:

$$f \sim g \;:\Leftrightarrow\; f(x) = g(x) \qquad \mu\text{-fast überall}$$

Mit $\hat{f}$ bezeichnet man zunächst die Äquivalenzklassen. Elemente in L^p sind also nicht Funktionen, sondern zunächst mal Funktionenklassen. Aus Gründen der Bequemlichkeit verwendet man diese Begriffe aber häufig synonym. In diesem Sinne ist eine stetige Funktion in L^p zum Beispiel so zu verstehen, dass es in der Äquivalenzklasse eine stetige Funktion gibt, es müssen aber keineswegs alle Funktionen in dieser Klasse stetig sein.

Definition 8.35 Die Funktionenräume $\mathcal{L}^p$ und $\mathcal{L}^\infty$

Wir definieren die folgenden Funktionenräume:

$$\begin{aligned}\hat{f} \in \mathcal{L}^p(X,\mu) \;&:\Leftrightarrow\; |f|^p \in \mathcal{L}(X,\mu),\ f \in \hat{f}\\ \hat{f} \in \mathcal{L}^\infty(X,\mu) \;&:\Leftrightarrow\; \exists_{c\in\mathbb{R}} |f(x)| \leq c \qquad \mu\text{-fast überall}\end{aligned}$$

Man schreibt oft $f \in \mathcal{L}^p$, $f \in \mathcal{L}^\infty$. Darauf kann man die folgende Norm einführen:

$$\begin{aligned}\|f\|_p &:= \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}}\\ \|f\|_\infty &:= \inf\{c \in \mathbb{R} \mid |f(x)| \leq c \mu\text{-fast überall}\}\end{aligned}$$

Satz 8.36 Höldersche Ungleichung

Sei $1 \leq p \leq \infty$, $1 \leq q \leq \infty$ mit $p^{-1} + q^{-1} = 1$, sowie $f \in \mathcal{L}^p$, $g \in L^q$. Dann folgt, dass $f \cdot g \in L^1(X,\mu)$ ist, $\|fg\|_1 \leq \|f\|_p \|g\|_q$ und

$$\int_X |fg| \,\mathrm{d}\mu \leq \left(\int_X |f|^p \,\mathrm{d}\mu\right)^{\frac{1}{p}} \cdot \left(\int_X |g|^q \,\mathrm{d}\mu\right)^{\frac{1}{q}}$$

Beweis. Sei $p = 1$, $q = \infty$, $g \in \mathcal{L}^\infty$. Dann folgt $|g(x)| \leq \|g\|_\infty$ μ-fast überall. Also ist $|f(x)g(x)| \leq |f(x)| \|g\|_\infty$ μ-fast überall. Dann folgt durch Integration $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$.

Sei nun $1 < p, q < \infty$. Zu Zeigen ist $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ für $a, b \in \mathbb{R}$ und $a, b \geq 0$. Sei $h(t) = \frac{t^p}{p} + \frac{t^{-q}}{q}$, $t > 0$. Für $t \to \infty$ und $t \to 0$ ist $h(t) \to \infty$. Betrachte die Ableitung $h'(t) = t^{p-1} - t^{-q-1}$. Es ist $h(t) \downdownarrows$ für $t \in]0,1[$ und $h(t) \upuparrows$ für $t \in]1,\infty[$. Dann ist $\min h(t) = h(1) = p^{-1} + q^{-1} = 1$. Daraus folgt $h(t) \geq 1$ für $t > 0$. Wähle nun $t = \frac{a^{\frac{1}{q}}}{b^{\frac{1}{p}}} \Rightarrow \frac{a^{\frac{p}{q}}}{pb} + \frac{b^{\frac{q}{p}}}{qa} \geq 1$. Daraus folgt dann

$$ab \leq \frac{a^{\frac{p}{q}+1}}{p} + \frac{b^{\frac{q}{p}+1}}{q} = \frac{a^p}{p} + \frac{b^q}{q}$$

Dann ist

$$\begin{aligned}1 + \frac{p}{q} &= p\left(\frac{1}{p} + \frac{1}{q}\right) = p\\ 1 + \frac{q}{p} &= q\left(\frac{1}{q} + \frac{1}{p}\right) = q.\end{aligned}$$

O.B.d.A. sei $\|f\|_p \neq 0$ und $\|g\|_q \neq 0$. Wähle $a = \frac{|f(x)|}{\|f\|_p}$ und $b = \frac{|g(x)|}{\|g\|_q}$. Dann folgt

$$ab = \frac{|f(x)| \cdot |g(x)|}{\|f\|_p \|g\|_q} \leq \frac{a^p}{p} + \frac{b^q}{q} = \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q}$$

Durch Integration erhält man dann schließlich:

$$\frac{\|fg\|_1}{\|f\|_p \|g\|_q} \leq \frac{1}{p}\frac{\|f\|_p^p}{\|f\|_p^p} + \frac{1}{q}\frac{\|g\|_q^q}{\|g\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1$$

Also zuletzt $\|fg\|_1 \leq \|f\|_p \|g\|_q$. □

Satz 8.37 Minkowskische Ungleichung
Sei $1 \leq p \leq \infty$ und $f, g \in \mathcal{L}^p(X, \mu)$. Dann folgt, dass $f + g \in \mathcal{L}^p(X, \mu)$ ist und $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Beweis. Die Fälle $p \in \{1, \infty\}$ werden zur Übung überlassen. Sei nun $1 < p < \infty$. Dann ist

$$|f(x) + g(x)|^p \leq \begin{cases} 2^p|f(x)|^p & f \geq g \\ 2^p|g(x)|^p & f < g \end{cases}$$

Daraus folgt $|f(x) + g(x)|^p < 2^p(|f|^p + |g|^p) \in \mathcal{L}(X, \mu)$, also folgt $|f + g| \in \mathcal{L}^p(X, \mu)$. Nun ist

$$\begin{aligned} \|f + g\|_p^p &= \int_X |f + g|^p \, \mathrm{d}\mu = \int_X |f + g||f + g|^{p-1} \, \mathrm{d}\mu \\ &\leq \int_X |f||f + g|^{p-1} \, \mathrm{d}\mu + \int_X |g||f + g|^{p-1} \, \mathrm{d}\mu, \end{aligned}$$

woraus mit der Hölderschen Ungleichung dann

$$\begin{aligned} &\|f\|_p \left(\int_X |f + g|^{q(p-1)} \, \mathrm{d}\mu \right)^{\frac{1}{q}} + \|g\|_p \left(\int_X |f + g|^{q(p-1)} \, \mathrm{d}\mu \right)^{\frac{1}{q}} \\ &\qquad \leq \left(\|f\|_p + \|g\|_p \right) \left(\int_X |f + g|^p \, \mathrm{d}\mu \right)^{\frac{1}{q}} \end{aligned}$$

Daraus folgt dann $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. □

Satz 8.38
Der Raum $\mathcal{L}^p(X, \mu)$ ist mit der Norm $\|\cdot\|_p$ vollständig. Das heißt für $1 \leq p \leq \infty$ folgt, dass $\mathcal{L}^p(X, \mu)$ ein Banachraum ist.

Beweis. Sei also $1 \leq p < \infty$. Der Fall $p = \infty$ geht ähnlich, wird hier aber nicht bewiesen. Sei $\{f_n\} \in \mathrm{CF}(\mathcal{L}^p)$, d.h.

$$\forall_{\varepsilon>0}\ \exists_{N(\varepsilon)}\ \forall_{n,m\geq N(\varepsilon)}\ :\ \|f_n - f_m\|^p = \int_X |f_n - f_m|^p \,\mathrm{d}\mu < \varepsilon^p.$$

Wir betrachten nun eine Teilfolge $n_1 < n_2 < \ldots < n_l < \ldots$ mit $n_k \in \mathbb{N}$ so, dass $\left\|f_{n_k} - f_{n_{k+1}}\right\|_p < 2^{-k}$ ist. Weiterhin betrachten wir $g \in \mathcal{L}^q(X,\mu)$ mit $p^{-1} + q^{-1} = 1$. Dann folgt mit der Hölderlin-Ungleichung

$$\int_X |g \cdot (f_{n_k} - f_{n_{k+1}})| \,\mathrm{d}\mu \leq \|g\|_q \underbrace{\left\|f_{n_k} - f_{n_{k+1}}\right\|_p}_{<2^{-k}}.$$

Daraus folgt dann

$$\sum_{k=1}^{\infty} \int_X \left|g \cdot (f_{n_k} - f_{n_{k+1}})\right| \,\mathrm{d}\mu \leq \|g\|_q.$$

Über den Satz der monotonen Konvergenz folgt damit

$$\int_X \overbrace{\sum_{k=1}^{\infty} \left|g \cdot (f_{n_k} - f_{n_{k+1}})\right|}^{=:F} \,\mathrm{d}\mu < \infty.$$

Es gilt, dass $F(x)$ μ-fast überall endlich ist. Daraus folgt die Konvergenz der Reihe:

$$\sum_{k=1}^{\infty} \left|g(x)(f_{n_k}(x) - f_{n_{k+1}}(x))\right| \qquad \text{konvergent } \mu\text{-fast überall.}$$

Schrittweise folgt dann nun die Konvergenz der Reihen

$$\sum_{k=1}^{\infty} \left|f_{n_k}(x) - f_{n_{k+1}}(x)\right| \qquad \text{konvergent}$$

$$\sum_{k=1}^{\infty} \left(f_{n_k}(x) - f_{n_{k+1}}(x)\right) = f_{n_1}(x) - \lim_{k\to\infty} f_{n_k}(x).$$

Daraus folgt dann die Existenz von $f(x) = \lim_{k\to\infty} f_{n_k}(x)$ μ-fast überall (d.h. auf einer Teilfolge). f ist messbar. Mit dem Lemma von Fatou folgt dann

$$\left(\int \underbrace{|f - f_{n_k}|^p}_{=:h\geq 0} \,\mathrm{d}\mu\right)^{\frac{1}{p}} \leq \liminf_{j\to\infty} \left(\int \underbrace{\left|f_{n_j} - f_{n_k}\right|^p}_{=:h_j\geq 0} \,\mathrm{d}\mu\right)^{\frac{1}{p}},$$

das heißt es ist $h_j(x) \xrightarrow{(\cdot)} h$ μ-fast überall. Nun ist

$$\overbrace{\left\|f_{n_j} - f_{n_{j-1}}\right\|_p}^{<2^{-j}} + \cdots + \overbrace{\left\|f_{n_{k+1}} - f_{n_k}\right\|_p}^{<2^{-k}} \geq \left\|f_{n_j} - f_{n_k}\right\|_p.$$

Und mit dieser in Klammern angedeuteten Abschätzung auch $\left\|f_{n_j} - f_{n_k}\right\|_p \leq 2 \cdot 2^{-k}$ für alle $j > k$. Daraus folgt, dass $\|f - f_{n_k}\|_p \leq 2 \cdot 2^{-k}$ ist und daraus schließlich $f_{n_k} \xrightarrow{\mathcal{L}^p} f$. Wir wollen nun aber eine Abschätzung für jedes Folgenglied, statt nur für eine Teilfolge. Mit der Dreiecksungleichung ist

$$\|f_m - f\|_p \leq \overbrace{\|f_m - f_{n_k}\|_p}^{<\varepsilon} + \underbrace{\|f_{n_k} - f\|_p}_{\to 0}.$$

Damit ist der Satz bewiesen. □

Wichtige Spezialfälle:
Beispiel 1: $\mathcal{L}^1(X, \mu)$, $\mathcal{L}^2(X, \mu) \ni f, g$.
Für diese beiden Funktionen ist ein Skalarprodukt definiert durch

$$\langle f, g \rangle_{\mathcal{L}^2} = \int_X f\overline{g} \, \mathrm{d}\mu,$$

wobei $\overline{g}$ das komplexe Konjugat von g ist. Spricht man, wie wir bisher, nur von reellwertigen Funktionen, so ist dies unerheblich. Wir sehen, dass der Satz von Cauchy-Schwartz-Bunjakowskij damit ein Spezialfall ist:

$$|\langle f, g \rangle| = \left| \int f\overline{g} \, \mathrm{d}\mu \right| \leq \int |f||g| \, \mathrm{d}\mu \leq \|f\|_2 \|g\|_2 \qquad p, q = 2.$$

Damit ist $\mathcal{L}^2(X, \mu)$ ein Hilbertraum.
Beispiel 2: $X = \mathbb{N}$, $\mathcal{R} = 2^{\mathbb{N}}$ σ-Algebra, $M \subset \mathbb{N}$ und $\mu(M) = \#$ sei das Zählmaß, d.h. $\mu(M) = \mathrm{card}(M)$. Damit ist $\mu(\{n\}) = 1$.
Betrachte nun eine Folge f_n mit $f : \mathbb{N} \to \mathbb{R}$. Dann ist

$$\|f\|_p = \left(\int_{\mathbb{N}} |f|^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} = \left(\sum_{n=1}^{\infty} |f_n|^p \right)^{\frac{1}{p}}.$$

Dies bedeutet, dass $\mathcal{L}^p(\mathbb{N}, \#) = l^p$ ist.
Allgemein: Im Allgemeinen sind die $\mathcal{L}^p(X, \mu)$ Banachräume.

Anmerkung. Sei $X = [a, b]$ und μ das Lebesgue-Maß. Dann zeigt sich, dass $C_0^\infty[a, b]$ dicht in $\mathcal{L}^p([a, b], \mu)$ liegt für $1 \leq p < \infty$.

8.8 Weitere Konvergenzaussagen

Sei $(X, \mathcal{R}, \mu)$ ein Maßraum und $\mu(X) < \infty$. Es ist

$$f_n \rightrightarrows f \;\Leftrightarrow\; \forall_{\varepsilon>0}\, \exists_{N_\varepsilon}\, \forall_{n\geq N_\varepsilon}\, \forall_{x\in X} \;:\; |f_n(x) - f(x)| < \varepsilon,$$

sowie

$$f_n \xrightarrow{(\cdot)} f(\mu\text{-f.ü.}) \;\Leftrightarrow\; \mu\left(\{x \;:\; f_n(x) \nrightarrow f(x)\}\right) = 0.$$

Das Untere folgt aus dem Oberen, nicht aber umgekehrt.

Satz 8.39 Satz von Egorov

Sei $\mu(X) < \infty$ und $f_n \xrightarrow{(\cdot)} f$ μ-fast überall. Dann folgt, dass für alle $\delta > 0$ ein $E_\delta \in \mathcal{R}$ existiert, so dass die folgenden beiden Eigenschaften erfüllt sind:

1. $\mu(X \setminus E_\delta) < \delta$
2. $f_n|_{E_\delta} \rightrightarrows f|_{E_\delta}$

Beweis. Es sei $E_n^m = \bigcap_{l\geq n}\{x \;:\; |f_l(x) - f(x)| < m^{-1}\}$, d.h. $|f_l(x) - f(x)| < m^{-1}$ für alle $l \geq n$. Weiterhin sei nun $E^m = \bigcup_{n=1}^\infty E_n^m$. Es ist $E_1^m \subset E_2^m \subset \ldots$ und

$$\mu(E^m) = \lim_{k\to\infty} \mu\Big(\underbrace{\bigcup_{n=1}^k E_n^k}_{=E_k^m}\Big).$$

Das heißt dann, dass für $\delta > 0$ ein $\mathbb{N} \ni n_0 = n_0(\delta, m)$ existiert, so dass $\mu(E^m \setminus E_{n_0}^m) < \delta 2^{-m}$ ist. Dann sei nun $E_\delta = \bigcap_{m=1}^\infty E_{n_0}^m$. Wir zeigen nun, dass diese Menge diese erwünschten Eigenschaften erfüllt.

Eigenschaft 2): Es ist $\tilde{x} \in E_\delta \;\Leftrightarrow\; \tilde{x} \in E_{n_0}^m\; \forall_m$. Daraus folgt, dass $|f_l(\tilde{x}) - f(\tilde{x})| < m^{-1}$ genau für alle $l \geq n_0$ erfüllt ist. Dies bedeutet aber, dass ein Grenzwert existiert mit $\lim_{n\to\infty} f_n(\tilde{x}) = f(\tilde{x})$. Dabei ist die Wahl von $n_0 = n_0(\delta, m)$ unabhängig von $\tilde{x}$, das heißt es liegt wirklich gleichmäßige Konvergenz vor: $f_n|_{E_\delta} \rightrightarrows f|_{E_\delta}$.

Eigenschaft 1): Für alle m gilt $\mu(X \setminus E^m) = 0$, denn es ist $x_0 \notin E^m$ genau dann, wenn $x_0 \notin E_n^m$ für alle $n \in \mathbb{N}$ gilt. Dies ist aber genau dann der Fall, wenn $\neg|f_l(x_0) - f(x_0)| < m^{-1}$ für alle $l \geq n$ ist. Daraus folgt dann aber $f_l(x_0) \nrightarrow f(x_0)$. Dies geschieht nach Voraussetzung aber nur für $x_0 \in A$ mit einer Menge A, für die $\mu(A) = 0$ ist. Es ist also wirklich $\mu(X \setminus E^m) = 0$.

Damit ist dann $\mu(X \setminus E_{n_0(\delta,m)}^m) = \mu(E^m \setminus E_{n_0(\delta,m)}^m) < \delta 2^{-m}$. Damit ist dann

$$\begin{aligned}\mu(X \setminus E_\delta) &= \mu\left(X \setminus \bigcap_{m=1}^\infty E_{n_0}^m\right) = \mu\left(\bigcup_{m=1}^\infty (X \setminus E_{n_0}^m)\right)\\ &= \sum_{m=1}^\infty \underbrace{\mu(X \setminus E_{n_0}^m)}_{<\delta 2^{-m}} \leq \delta.\end{aligned}$$

Damit ist der Satz bewiesen. □

Konvergenz im Maß:

Definition 8.40 Konvergenz im Maß
Wir definieren

$$f_n \xrightarrow{\mu} f \;:\Leftrightarrow\; \forall_{\delta>0} \lim_{n\to\infty} \mu\left(\{x \;:\; |f_n(x) - f(x)| \geq \delta\}\right) = 0$$

und sagen dann, f_n konvergiert im Maß gegen f.

Satz 8.41
Sei $f_n \xrightarrow{(\cdot)} f$ μ-fast überall. Dann ist auch $f_n \xrightarrow{\mu} f$.

Beweis. $f_n(x) \xrightarrow{(\cdot)} f(x)$ für $x \in X \setminus A$ mit $\mu(A) = 0$. Es sei $E_k(\delta) = \{x \;:\; |f_k(x) - f(x)| \geq \delta\}$. Außerdem sei $R_n(\delta) = \bigcup_{k=n}^{\infty} E_k(\delta)$ und $M = \bigcap_{n=1}^{\infty} R_n(\delta)$. Offenbar ist $R_1(\delta) \supset R_2(\delta) \supset \ldots$. Daraus folgt $\mu(M) = \lim_{n\to\infty} \mu(R_n(\delta))$. Unser Ziel ist es, zu zeigen, dass $M \subset A$ ist, denn daraus folgt dann $\mu(M) = 0$. Sei $x_0 \in X \setminus A$, das heißt $f_n(x_0) \xrightarrow{n\to\infty} f(x_0)$. Daraus folgt $|f_n(x_0) - f(x_0)| < \delta$ für alle $n \geq N$, und daraus wiederum $x_0 \notin E_n(\delta)$ für $n \geq N$. Dies impliziert dann $x_0 \notin R_n(\delta)$ für $n \geq N$ und daraus letztlich $x_0 \notin M$. Dies heißt dann gerade $M \subset A$. Nach dem oben Gesagten ist dann $\mu(M) = 0$. Dann ist $\lim_{n\to\infty} \mu(R_n(\delta)) = \mu(M) = 0$. Da $E_n(\delta) \subset R_n(\delta)$ ist gilt $0 \leq \mu(E_n(\delta)) \leq \mu(R_n(\delta))$. Da nun die rechte Seite gegen 0 konvergiert folgt $\lim_{n\to\infty} \mu(E_n(\delta)) = 0$, das heißt

$$\lim_{n\to\infty} \mu\Big(\underbrace{\{x \;:\; |f_n(x) - f(x)| \geq \delta\}}_{=E_n(\delta)}\Big) = 0.$$

Damit ist der Satz bewiesen. □

Anmerkung. Die Umkehrung $f_n \xrightarrow{\mu} f \;\Rightarrow\; f_n \xrightarrow{(\cdot)} f$ μ-fast überall gilt nicht!

Satz 8.42
Es sei $f_n \xrightarrow{\mu} f$. Dann existiert eine Teilfolge f_{n_k}, so dass $f_{n_k} \xrightarrow{(\cdot)} f$ μ-fast überall.

Beweis. Der Satz wird hier nicht bewiesen. □

Aufgabe. Zeigen Sie: $f_n \xrightarrow{\mathcal{L}^1} f \;\Rightarrow\; f_n \xrightarrow{\mu} f$.

Zusammenfassung: Für $\mu(X) < \infty$ erhalten wir also folgendes Diagramm:

$$\begin{tikzcd}
 & f_n \rightrightarrows f \arrow[dl, Rightarrow] \arrow[dr, Rightarrow] & \\
\mathcal{L}^2\text{-Konvergenz} \arrow[d, Rightarrow] & & f_n \xrightarrow{(\cdot)} f\,\mu\text{-f.ü.} \arrow[ul, "\text{auf } E_\delta \text{ (Egorov)}"'] \arrow[dl, Rightarrow] \\
\mathcal{L}^1\text{-Konvergenz} \arrow[r, Rightarrow] & f_n \xrightarrow{\mu} f \arrow[ur, "\text{auf Teilfolge}"'] &
\end{tikzcd}$$

9 Zur Theorie der gewöhnlichen Differentialgleichungen

Anmerkung: Im Englischen werden gewöhnliche Differentialgleichungen „ordinary differential equations“ genannt. Daher lautet das Kürzel für diese „ODE“.

9.1 Motivation

ODEs beschreiben Probleme, welche deterministisch, endlichdimensional und differenzierbar sind. Doch was bedeuten diese Begriffe im Einzelnen? Dies wollen wir hier kurz erläutern:

1. **Deterministisch:** Ist der Zustand eines Systems zum Zeitpunkt $t = t_0$ (Gegenwart) bekannt, so kann er für alle t bestimmt werden.
2. **Endlichdimensional:** Der Zustand des Systems wird durch endlich viele Größen bestimmt.

Die Menge aller möglichen Zustände eines Systems nennt man den Phasenraum M.

Beispiel:
Ein Beispiel einer typischen Differentialgleichung ist $\dot{g}(t) = -\kappa g(t)$ $(\kappa > 0)$, dessen allgemeine Lösung sich durch $g(t) = g(0)e^{-\kappa t}$ beschreiben lässt. Der Zustand $g(t)$ des Systems ist also alleine durch das Wissen des Parameters $g_0 = g(0) \geq 0$ bestimmt. Der Phasenraum wäre dann $M = [0, +\infty[$.

Mathematische Beschreibung von „deterministisch“:
Sei $x = y(0) \in M$. Für $t \in \mathbb{R}$ sei $y(t) \in M$. Es sei $g^t : M \to M$ mit $g^t x = y(t)$. Die „Spur“, welche $g^t x$ in M für ein x hinterlässt wird mit Trajektorie oder Orbit bezeichnet. Es gelten die folgenden Eigenschaften:

1. $g^0 = \mathrm{id}$
2. $g^{t+s} = g^t g^s = g^{s+t} = g^s g^t$
3. $g^{-t} = (g^t)^{-1}$

Insgesamt ist $\{g^t\}$ also eine abelsche Gruppe. Das Paar $(M, \{g^t\})$ nennt man Phasenfluss.

Zum Begriff der Differenzierbarkeit:
$g^t x = g(t, x)$. Betrachte $g : \mathbb{R} \times M \to M$ mit $M \subset \mathbb{R}^n$. g sei in t differenzierbar. Betrachte nun

$$v(x) = \frac{\mathrm{d}}{\mathrm{d}\tau} (g^\tau x) \Big|_{\tau=0}$$

Man nennt $v(x)$ das Geschwindigkeitsfeld. Es ist

$$\begin{aligned}
\mathbb{v}(y(t)) &= \frac{\mathrm{d}}{\mathrm{d}\tau}(g^\tau y(t))\Big|_{\tau=0} = \frac{\mathrm{d}}{\mathrm{d}\tau}(g^\tau \circ g^t x)\Big|_{\tau=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}\tau}(g^{\tau+t} x)\Big|_{\tau=0} \overset{\tau \to s = r+t}{=} \frac{\mathrm{d}}{\mathrm{d}s}(g^s x)|_{s=t} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} y(t)
\end{aligned}$$

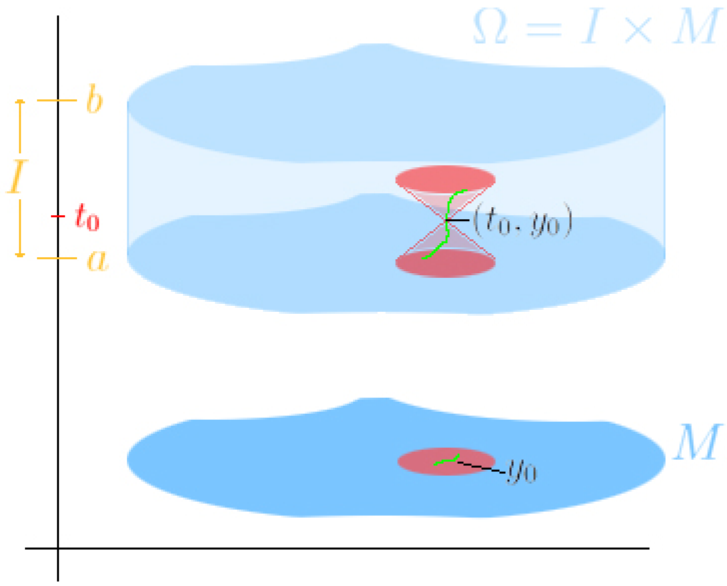
Daraus folgt also $\dot{y}(t) = v(y(t))$. Ist $v = v(x)$ zeitunabhängig, so spricht man von einer autonomen DGL. Ist dagegen $\dot{y}(t) = v(t, y(t))$, $v = v(t, x)$ zeitabhängig, so spricht man von einer nicht autonomen DGL.

DGL n-ter Ordnung:

$$y^{(n)}(t) = v\left(t, y(t), \ldots, y^{(n-1)}(t)\right)$$

9.2 Die Methode von Euler

Im Laufe des Kapitels sollte man sich stets die folgende Skizze (Erklärungen unter dem Bild) im Hinterkopf behalten:



Gegeben sei ein Phasenraum $M \subset \mathbb{R}^n$. Außerdem sei ein Zeitintervall $I = [a, b]$ und darin ein Punkt $t_0 \in I$ gegeben. Wir wollen das System in diesem Intervall bestimmen, wobei es dabei den Phasenraum nie verlassen soll. Veranschaulicht man den $\mathbb{R}^n$ als Fläche und die Zeit als dritte Achse, so ist $\Omega := I \times M$ ein Zylinder, dessen Grundfläche der Phasenraum M ist. In Ω suchen wir eine Integralkurve. Projeziert man diese wieder auf den $\mathbb{R}^n$ herunter, so erhält man gerade die Trajektorie in M. Sei nun $\mathbb{v} : I \times M \to \mathbb{R}^n$ das Geschwindigkeitsfeld. $\mathbb{v}$ sei stetig und es sei $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ und $\mathbb{y}(t) = \mathbb{y}_0$. Dies nennt man ein Cauchyproblem oder Anfangswertproblem (AWP).
Betrachte nun die Gleichung $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$. Durch Integrieren wollen wir die Ableitung eliminieren:

$$\mathbb{y}(s) = \mathbb{y}(t_0) + \int_{t_0}^{s} \mathbb{v}(t, \mathbb{y}(t))\ \mathrm{d}t$$

Allerdings hat man nun das Problem, dass $\mathbb{y}$ dafür im Integranden auftaucht. Wir haben nun also folgende zwei Probleme:

(1): $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ mit $\mathbb{y}(t)|_{t=t_0} = \mathbb{y}_0$

(2): $\mathbb{y}(s) = \mathbb{y}_0 + \int_{t_0}^{s} \mathbb{v}(t, \mathbb{y}(t))\ \mathrm{d}t$

Wir wollen an das Vektorfeld $\mathbb{v}$ nun einige Forderungen stellen, so dass die Differential- und die Integralgleichung nicht nur zu äquivalenten Problemen, sondern auch lösbar in unserem Verfahren werden:

1. $\mathbb{v} : \Omega \to \mathbb{R}^n$ stetig.
2. $\|\mathbb{v}(t, \mathbb{x})\| \leq C$, $(t, \mathbb{x}) \in \Omega$.
3. $\|\mathbb{v}(t, x') - \mathbb{v}(t, x'')\| \leq L \, \|x' - x''\|$.

Annahme: Es existiert eine Lösung von **(1)** $\Leftrightarrow$ **(2)**. Wie können wir diese dann bestimmen?

Unterteile zunächst I in $t_n, t_{n-1}, \ldots, t_0, t_{-1}, \ldots$.
Erster Schritt: $t \in [t_0, t_1]$. Wir approximieren

$$\textbf{(3):} \qquad \tilde{\mathbb{y}}(s) = \mathbb{y}(t_0) + \int_{t_0}^{s} \mathbb{v}(t, \mathbb{y}_0) \, \mathrm{d}t$$

Es sei $\tilde{\mathbb{y}}(t_1) = \mathbb{y}_1$.
k-ter Schritt: i.A. für $s \in [t_{k-1}, t_k]$:

$$\textbf{(3'):} \qquad \tilde{\mathbb{y}}(s) = \mathbb{y}_{k-1} + \int_{t_{k-1}}^{s} \mathbb{v}(t, \mathbb{y}_{k-1}) \, \mathrm{d}t$$

Wichtig: Im Allgemeinen ist

$$\mathbb{y}_k = \tilde{\mathbb{y}}(t_k) \overset{\text{i.A.}}{\neq} \mathbb{y}(t_k).$$

Vereinfachung: $\mathbb{v}(t, \mathbb{y}_{k-1}) \to \mathbb{v}(t_{k-1}, \mathbb{y}_{k-1})$. Dann ist

$$\hat{\mathbb{y}} = \hat{\mathbb{y}}_{k-1} + \underbrace{\int_{t_{k-1}}^{s} \mathbb{v}(t_{k-1}, \mathbb{y}_{k-1}) \, \mathrm{d}t}_{(s-t_{k-1})\mathbb{v}(t_{k-1}, \mathbb{y}_{k-1})} .$$

Frage: Unter welchen Bedingungen bleibt $\tilde{\mathbb{y}}$ bzw. $\hat{\mathbb{y}}$ in M? Aus der zweiten Forderung von oben erhält man

$$\|\mathbb{y}(s) - \mathbb{y}_0\| \leq |s - t_0| \cdot C$$

und fordert dann noch

$$\leq \operatorname{dist}(y_0, \partial M).$$

Daraus erhält man dann eine Bedingung an das Zeitintervall I:

$$|s - t_0| \leq C^{-1} \operatorname{dist}(y_0, \partial M).$$

Man kann all diese Schritte analog für $\tilde{\mathbb{y}}$ und $\hat{\mathbb{y}}$ durchführen und erhält die selben Abschätzungen. Nicht nur die Lösung, sondern auch unsere Approximation bleiben also in M.

Frage: Wie gut ist diese Approximation? Es sei $|b-0| = |b-t_0| \leq C^{-1}\operatorname{dist}(y_0, \partial M)$ und $b = t_N$, dann ist $|t_k - t_{k-1}| = \frac{b}{N}$. Für $s \in [t_0, t_1]$ berechnen wir zunächst **(2)**$-$**(3)**:

$$\mathbb{y}(s) - \tilde{\mathbb{y}}(s) = \int_{t_0}^{s} (\mathbb{v}(t, \mathbb{y}(t)) - \mathbb{v}(t, \mathbb{y}_0)) \, \mathrm{d}t$$

Damit könnnen wir dann berechnen:

$$\begin{aligned}\|\mathbb{y}(s) - \tilde{\mathbb{y}}(s)\| &\overset{3.}{\leq} L \int_{t_0}^{s} \|\mathbb{y}(t) - \mathbb{y}_0\| \, \mathrm{d}t \\ &\leq L \int_{t_0}^{s} (t - t_0) \cdot C \, \mathrm{d}t = \frac{LC}{2}(s - t_0)^2.\end{aligned}$$

Es sei $\Delta_1 := \|\mathbb{y}(t_1) - \mathbb{y}_1\|$. Es gilt dann insbesondere $\Delta_1 \leq \frac{LC}{2}\left(\frac{b}{N}\right)^2$.
Wir betrachten nun nochmal **(2)** und schreiben diese Gleichung wie folgt:

$$\textbf{(2'):} \qquad \mathbb{y}(s) = \mathbb{y}(t_{k-1}) + \int_{t_{k-1}}^{s} \mathbb{v}(t, \mathbb{y}(t)) \, \mathrm{d}t$$

Nun berechnen wir für $s \in [t_{k-1}, t_k]$ den Term **(2')**$-$**(3')**:

$$\mathbb{y}(s) - \tilde{\mathbb{y}}(s) = \mathbb{y}(t_{k-1}) - \mathbb{y}_{k-1} + \int_{t_{k-1}}^{s} (\mathbb{v}(t, \mathbb{y}(t)) - \mathbb{v}(t, \mathbb{y}_{k-1})) \, \mathrm{d}t,$$

und können dann wie folgt abschätzen:

$$\begin{aligned}&\|\mathbb{y}(s) - \tilde{\mathbb{y}}(s)\| \\ &\leq \underbrace{\|\mathbb{y}(t_{k-1}) - \mathbb{y}_{k-1}\|}_{\Delta_{k-1}} + \left\| \int_{t_{k-1}}^{s} (\mathbb{v}(t, \mathbb{y}(t)) \underbrace{- \mathbb{v}(t, \mathbb{y}(t_{k-1})) + \mathbb{v}(t, \mathbb{y}(t_{k-1}))}_{=0} - \mathbb{v}(t, \mathbb{y}_{k-1})) \, \mathrm{d}t \right\| \\ &\leq \Delta_{k-1} + \int_{t_{k-1}}^{s} \|\mathbb{v}(t, \mathbb{y}(t)) - \mathbb{v}(t, \mathbb{y}(t_{k-1}))\| \, \mathrm{d}t + \int_{t_{k-1}}^{s} \|\mathbb{v}(t, \mathbb{y}(t_{k-1})) + \mathbb{v}(t, \mathbb{y}_{k-1})\| \, \mathrm{d}t \\ &\leq \Delta_{k-1} + L \int_{t_{k-1}}^{s} \|\mathbb{y}(t) - \mathbb{y}(t_{k-1})\| \, \mathrm{d}t + L \int_{t_{k-1}}^{s} \underbrace{\|\mathbb{y}(t_{k-1}) - \mathbb{y}_{k-1}\|}_{\Delta_{k-1}} \, \mathrm{d}t \\ &\qquad\qquad \leq \Delta_{k-1} + \frac{LC}{2}(s - t_{k-1})^2 + L(s - t_{k-1})\Delta_{k-1}\end{aligned}$$

Für $s = t_k$ ergibt sich damit

$$\Delta_k = \|\mathbb{y}(t_k) - \tilde{\mathbb{y}}(t_k)\| \leq \Delta_{k-1}\left(1 + L\frac{b}{N}\right) + \frac{LC}{2}\left(\frac{b}{N}\right)^2.$$

Wie entwickelt sich nun der Fehler? Es ist

$$\begin{aligned}
\Delta_1 &\leq \frac{LC}{2}\left(\frac{b}{N}\right)^2 \\
\Delta_2 &\leq \frac{LC}{2}\left(\frac{b}{N}\right)^2\left(1+L\frac{b}{N}\right)+\frac{LC}{2}\left(\frac{b}{N}\right)^2 \\
&= \frac{LC}{2}\left(\frac{b}{N}\right)^2\left[1+\left(1+L\frac{b}{N}\right)\right] \\
\Delta_3 &\leq \frac{LC}{2}\left(\frac{b}{N}\right)^2\left[1+\left(1+L\frac{b}{N}\right)+\left(1+L\frac{b}{N}\right)^2\right]
\end{aligned}$$

Und für ein beliebiges k ergibt sich:

$$\begin{aligned}
\Delta_k &\leq \frac{LC}{2}\left(\frac{b}{N}\right)^2\sum_{l=0}^{k-1}\left(1+L\frac{b}{N}\right)^l \\
&= \frac{\cancel{L}C}{2}\left(\frac{b}{N}\right)^{\cancel{2}}\cdot\frac{\left(1+L\frac{b}{N}\right)^k-1}{\left(\cancel{1}+\cancel{L\frac{b}{N}}\right)\cancel{-1}}
\end{aligned}$$

Und damit dann:

$$\Delta_k \leq \frac{C}{2}\left(\frac{b}{N}\right)\left[\left(1+L\frac{b}{N}\right)^k-1\right].$$

Für $k=N$ folgt damit

$$\begin{aligned}
\|\mathrm{y}(b)-\tilde{\mathrm{y}}(b)\| = \Delta_N &\leq \frac{1}{N}\frac{Cb}{2}\left(\left(1+\frac{Lb}{N}\right)^N-1\right) \\
&\leq \frac{1}{N}\frac{Cb}{2}\left[e^{Lb}-1\right].
\end{aligned}$$

Wir sehen also, dass das Euler-Verfahren wirklich konvergiert und haben sogar eine Abschätzung an die Konvergenzgeschwindigkeit.

9.3 Lokale Existenz und Eindeutigkeit der Lösung des Cauchy-Problems

Wir betrachten das Cauchy-Problem $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ und $\mathbb{y}(t)|_{t=t_0} = \mathbb{y}_0$. Es sei $t, t_0 \in I = [a, b]$ und $\Omega = I \times M$, wobei $M \subset \mathbb{R}^n$ der Phasenraum ist.

Satz 9.1 Picard-Lindelöf
Erfüllt $\mathbb{v} : \Omega \to \mathbb{R}^n$ die drei Bedingungen von oben, das heißt erfüllt es

1. $\mathbb{v} : \Omega \to \mathbb{R}^n$ stetig,
2. $\|\mathbb{v}(t, \mathbb{x})\| \leq C$, $(t, \mathbb{x}) \in \Omega$,
3. $\|\mathbb{v}(t, x') - \mathbb{v}(t, x'')\| \leq L \|x' - x''\|$,

so besitzt das obige Cauchy-Problem genau eine Lösung für $t \in I_\varepsilon(t_0)$ mit $I_\varepsilon = \{t \in I : |t - t_0| \leq (1 - \varepsilon)\alpha\}$, wobei $\varepsilon > 0$ und $\alpha = \min\{C^{-1} \operatorname{dist}(y_0, \partial M), L^{-1}\}$ ist.

Beweis. Strategie: Anwendung des Fixpunktsatzes von Banach.
1. Schritt: $\mathbb{y} : I_\varepsilon(t_0) \to \mathbb{R}^n$ löst das Cauchyproblem genau dann, wenn $\mathbb{y}(t) = \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, \mathbb{y}(\tau)) \,\mathrm{d}\tau$ ist ($t \in I_\varepsilon(t_0)$). Dies ist äquivalent dazu, dass $\mathbb{y} = T\mathbb{y}$ mit $(T\mathbb{y})(t) = \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, \mathbb{y}(\tau)) \,\mathrm{d}\tau$ gilt. Dabei ist $\mathbb{y} \in C(I_\varepsilon(t_0), \mathbb{R}^n)$.
2. Schritt: Es sei $F = \{f \in C(I_\varepsilon(t_0), \mathbb{R}^n) | f(t) \in \overline{U_{r_\varepsilon}(\mathbb{y}_0)}\}$, d.h. $\forall t \in I_\varepsilon(t_0)$ ist $\|f(t) - \mathbb{y}_0\| \leq r_\varepsilon$ mit $r_\varepsilon > 0$. Für $f, g \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ erinnern wir uns an $\|f - g\|_C := \max_{t \in I_\varepsilon(t_0)} \|f(t) - g(t)\| = d_C(f, g)$. Wir behaupten, dass (F, d_C) ein *vollständiger* metrischer Raum ist. Wir wissen, dass $C(I_\varepsilon(t_0), \mathbb{R}^n)$ ein Banachraum ist, insbesondere ist er also vollständig. Für diesen Raum ist T aber nicht wohldefiniert, weshalb wir uns auf F beschränken. Wir betrachten daher nun eine Cauchyfolge $\{f_n\} \in \mathrm{CF}(F, d_C)$, das heißt $d_C(f_n, f_m) \to 0$. Es ist aber offenbar auch $\{f_n\} \in \mathrm{CF}(C(I_\varepsilon(t_0), \mathbb{R}^n))$ und da dieser vollständig ist existiert ein $f \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ mit $f_n \xrightarrow{d_C} f$. Da $f_n \in F$ gilt folgt, dass $\forall t \in I_\varepsilon(t_0)$ gilt, dass $\|f_n(t) - \mathbb{y}_0\| \leq r_\varepsilon$ ist. Und wegen $f_n \xrightarrow{d_C} f \Leftrightarrow f_n \rightrightarrows f$ folgt auch $f_n \xrightarrow{(\cdot)} f$ und daraus $\|f(t) - \mathbb{y}_0\| \leq r_\varepsilon$. Also ist $f \in F$ und es gilt $f_n \xrightarrow{d_C} f$, also ist F wirklich ein vollständiger metrischer Raum.
3. Schritt: Wir zeigen, dass für ein geeignetes $r_\varepsilon > 0$ wirklich $T : F \to F$ ist und dass T dann sogar eine Kontraktion ist. Für $\mathbb{y} \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ gilt $T\mathbb{y} \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ nach der ersten Voraussetzung im Satz. Es ist

$$\begin{aligned} \|(T\mathbb{y})(t) - \mathbb{y}_0\| &\leq \left| \int_{t_0}^t \underbrace{\|\mathbb{v}(\tau, \mathbb{y}(\tau))\|}_{\leq C} \,\mathrm{d}\tau \right| \leq |t - t_0| \cdot C \\ &\leq (1 - \varepsilon) \operatorname{dist}(\mathbb{y}_0, \partial M). \end{aligned}$$

Das bedeutet für $r_\varepsilon \geq (1 - \varepsilon) \operatorname{dist}(\mathbb{y}_0, \partial M)$ gilt $T\mathbb{y} \in F$ für alle $\mathbb{y} \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ mit $\mathbb{y}(t) \in M$, insbesondere also auch für $\mathbb{y} \in F$. Es bleibt noch zu zeigen, dass es sich wirklich um eine Kontraktion handelt. Dazu betrachten wir $f, g \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ mit

$f(t), g(t) \in M$. Es ist

$$\begin{aligned} \|(Tf)(t) - (Tg)(t)\| &= \left\| \int_{t_0}^{t} \Big(\mathbb{v}(\tau, f(\tau)) - \mathbb{v}(\tau, g(\tau)) \Big) \, \mathrm{d}\tau \right\| \\ &\leq \left| \int_{t_0}^{t} L \cdot \|f(\tau) - g(\tau)\| \, \mathrm{d}\tau \right| \\ &\leq |t - t_0| \cdot L \cdot \max_{\tau \in I_\varepsilon(t_0)} \|f(\tau) - g(\tau)\| \\ &\leq \underbrace{|t - t_0| \cdot L}_{\leq (1-\varepsilon)} \cdot d_C(f, g) \\ &\leq (1 - \varepsilon) d_C(f, g). \end{aligned}$$

Es ist also $d_C(Tf, Tg) \leq (1 - \varepsilon) d_C(f, g)$ für $\varepsilon > 0$. Wähle nun $\varepsilon > 0$ und $r_\varepsilon = (1 - \varepsilon) \operatorname{dist}(\mathbb{y}_0, \partial M)$, so ist also (F, d_C) ein vollständiger metrischer Raum und $T : F \to F$ eine Kontraktion. Mit dem Fixpunktsatz von Banach folgt dann, dass *genau* ein $\mathbb{y} \in F$ mit $T\mathbb{y} = \mathbb{y}$ gibt. □

Anmerkung: Durch wiederholtes Anwenden des Satzes von Picard-Lindelöf kann man die Lösung schrittweise eindeutig fortsetzen (und zwar sowohl vorwärts als auch rückwärts in der Zeit), bis sie Ω verlässt, das heißt, bis sie entweder das gesamte Zeitintervall ausfüllt (d.h. durch die obere Wand des „Zylinders“ stößt), oder bis sie den Phasenraum verlässt (d.h. durch die Wand des „Zylinders“ stößt).

Anmerkung: Die Voraussetzungen des Satzes sind nicht „optimal“, das heißt sie sind nicht die minimalsten Anforderungen, um die Aussage des Satzes zu erhalten. Der Satz von Picad-Lindelöf hat aber den Vorteil, dass die Bedingung der Lipschitz-Stetigkeit einfach zu verifizieren ist. In der folgenden Anmerkung soll dazu dennoch ein Beispiel gegeben werden.

Anmerkung: Es sei $M = \overline{M}$ konvex und beschränkt, sowie $\mathbb{v}$ auf $\operatorname{int}(M)$ sei Frechet-differenzierbar, sowie $\mathbb{v}'$ stetig auf $\overline{M}$ fortsetzbar. Konvexität bedeutet dabei, dass das Verbindungsstück zweier Punkte von M immer ganz in M liegt. Damit ist dann

$$\|\mathbb{v}(t, \mathbb{x}) - \mathbb{v}(t, \mathbb{y})\| \leq \sup_{\tilde{\mathbb{x}} \in \overline{\mathbb{x}, \mathbb{y}}} \|D_{\mathbb{x}} \mathbb{v}(t, \tilde{\mathbb{x}})\| \cdot \|\mathbb{x} - \mathbb{y}\| \leq L \|\mathbb{x} - \mathbb{y}\| .$$

Die Lipschitz-Stetigkeit ist hier also eine Folgerung und damit eine schwächere Bedingung als Differenzierbarkeit.

Der Fixpunktsatz von Banach gibt zudem eine Lösungsmethode:
Wähle $\mathbb{h}_0 \in C(I_\varepsilon(t_0), \mathbb{R}^n)$ mit $\mathbb{h}_0(t) = \mathbb{y}_0 \;\Rightarrow\; \mathbb{h}_0 \in F$. Dann sei für $t \in I_\varepsilon(t_0)$:

$$\mathbb{h}_j(t) = (T\mathbb{h}_{j-1})(t) = \mathbb{y}_0 + \int_{t_0}^{t} \mathbb{v}(\tau, \mathbb{h}_{j-1}(\tau)) \, \mathrm{d}\tau.$$

Daraus folgt dann $\mathbb{h}_j \xrightarrow{d_C} \mathbb{y}$, das heißt $\mathbb{h}_j \rightrightarrows \mathbb{y}$.

Frage: Wie schnell konvergiert diese Lösungsmethode?

$$\begin{aligned}\|\mathbb{y}(t) - \mathbb{h}_0\| &\leq \left\| \int_{t_0}^{t} \mathbb{v}(\tau, \mathbb{y}(\tau)) \, \mathrm{d}\tau \right\| \\ &\leq C|t - t_0|,\end{aligned}$$

also folgt dann:

$$\begin{aligned}\|\mathbb{y}(t) - \mathbb{h}_1(t)\| &\leq \left\| \int_{t_0}^{t} \Big(\mathbb{v}(\tau, \mathbb{y}(\tau)) - \mathbb{v}(\tau, \mathbb{h}_0) \Big) \, \mathrm{d}\tau \right\| \\ &\leq \left| \int_{t_0}^{t} L \, \|\mathbb{y}(\tau) - \mathbb{h}_0\| \, \mathrm{d}\tau \right| \\ &\leq CL \left| \int_{t_0}^{t} |\tau - t_0| \, \mathrm{d}\tau \right| \\ &\leq \frac{CL}{2} |t - t_0|^2.\end{aligned}$$

Führt man dies fort, so erhält man für $j \in \mathbb{N}$:

$$\|\mathbb{y}(t) - \mathbb{h}_j(t)\| \leq \frac{CL^j}{(j+1)!} |t - t_0|^{j+1}$$

für $t \in I_\varepsilon(t_0)$.

9.4 Der Satz von Peano

Der Satz von Picard-Lindelöf hat als wichtige Voraussetzung die Lipschitz-Stetigkeit ($\|\mathbb{v}(t, \mathbb{x}') - \mathbb{v}(t, \mathbb{x}'')\| \leq L \, \|\mathbb{x}' - \mathbb{x}''\|$). Dies ist aber in vielen Anwendungen nicht gegeben, wie das folgende Beispiel zeigen soll:
Beispiel: $\dot{y}(t) = \sqrt[3]{y^2(t)}$, $t \in \mathbb{R}$, $y(0) = 0$.
Das Geschwindigkeitsfeld ist autonom und lautet $v(t, y) = v(y) = y^{\frac{2}{3}}$. v ist aber *nicht* Lipschitz-stetig (in 0). Wir betrachten nun die Lösungen der Differentialgleichung.
Lösung 1): Die triviale Lösung lautet $y(t) = 0$ $(t \in \mathbb{R})$.
Lösung 2): Wir machen den folgenden Ansatz:

$$\frac{\dot{y}(t)}{\sqrt[3]{y^2(t)}} = 1 \qquad \Big| \int_{t_0}^{t}$$

Dies führt dann zu

$$\underbrace{\int_{t_0}^{t} \frac{\dot{y}(\tau)}{\sqrt[3]{y^2(\tau)}} \, \mathrm{d}\tau = \int_{y_0}^{y} \frac{\mathrm{d}y}{y^{\frac{2}{3}}} = 3y^{\frac{1}{3}} \Big|_{y(t_0)=y_0}^{y(t)}}_{\text{linke Seite}} = \underbrace{(t - t_0)}_{\text{rechte Seite}} .$$

Wir erhalten damit

$$3\left(y^{\frac{1}{3}}(t) - \underbrace{\cancel{y^{\frac{1}{3}}(t_0)}}_{=0}\right) = t - \underbrace{\cancel{t_0}}_{=0},$$

was uns schließlich zur Lösung $y(t) = \left(\frac{t}{3}\right)^3$ führt. Dies löst die Differentialgleichung und erfüllt die Anfangsbedingung.

Satz 9.2 Satz von Peano
Sei $I \subset \mathbb{R}$, $M \subset \mathbb{R}^n$, $\Omega = I \times M$, $(t_0, y_0) \in \Omega$. Zudem gelte

1. $\mathbb{v} : \Omega \xrightarrow{\text{stetig}} \mathbb{R}^n$.
2. $\|\mathbb{v}(t, \mathbb{x})\| \leq C$ für alle $(t, \mathbb{x}) \in \Omega$.

Dann folgt, dass das Cauchyproblem für $t \in \tilde{I}_\varepsilon(t_0)$ *mindestens* eine Lösung besitzt, wobei $\tilde{I}_\varepsilon(t_0) = \left\{t \in I : |t - t_0| \leq (1 - \varepsilon)\frac{1}{C} \operatorname{dist}(y_0, \partial M)\right\}$ ist.

Wir wollen uns dem Beweis dieses Satzes nun langsam nähern. Dazu formulieren wir zunächst eine Vorgehensweise.

Strategie des Beweises: Wir wollen einen vollständigen, metrischen Raum (F, d) finden, um den Fixpunktsatz von Banach anzuwenden. Klar ist, dass es keine kleine Änderung des Beweises des Satzes von Picard-Lindelöf sein kann, da wir die Eindeutigkeit der Lösung im Satz verlieren.

Definition 9.3 relative Kompaktheit
$G \subset F$ heißt relativ kompakt genau dann, wenn $\overline{G}$ (folgen-)kompakt ist.

Beispiel: $G \subset \mathbb{R}^n$, G beschränkt, so ist G relativ kompakt, denn $\overline{G}$ ist dann beschränkt und abgeschlossen (Satz von Bolzano).
Das Kriterium von Bolzano gilt aber *nicht* in unendlich-dimensionalen Räumen!

Anmerkung: Gegeben sei eine Folge $\{f_n\} \in G$, dann ist G relativ kompakt genau dann, wenn es eine Teilfolge $\{f_{n_k}\} \to f \in \overline{G}$ gibt.

Es sei $D \subset F$ und $T : D \to F$, wobei T im Allgemeinen nicht linear ist. Wir führen folgenden Begriff ein:

Definition 9.4 kompakte Abbildung
T heißt kompakt auf D genau dann, wenn $TD = \{y \in F | y = Tx,\ x \in D\}$ relativ kompakt ist.

Definition 9.5 Approximative Lösbarkeit des Fixpunktproblems
Das Fixpunktproblem lässt sich für T auf D approximativ lösen, falls eine Folge $\{x_n\}$ mit $x_n \in D$ existiert, so dass $d(Tx_n, x_n) \xrightarrow{n\to\infty} 0$.

Anmerkung: Während die x_n noch in D liegen, kann es bereits sein, dass Tx_n nicht mehr in D liegt. Approximative Lösbarkeit bedeutet jedoch, dass die Bilder nicht „sehr weit“ aus D herausführen.

Satz 9.6
Es sei (F, d) ein vollständiger metrischer Raum und $F \supset D$ abgeschlossen. Ferner sei eine Abbildung $T : D \to F$ gegeben, welche die folgenden Eigenschaften erfüllt:

1. $T : D \to F$ ist stetig.
2. T ist auf D kompakt.
3. Das Fixpunktproblem ist für T auf D approximativ lösbar.

Dann folgt, dass es *mindestens* ein $y \in D$ mit $Ty = y$ gibt.

Beweis. Aus **3.** folgt, dass es eine Folge $\{x_n\} \subset D$ gibt, so dass $d(Tx_n, x_n) \to 0$ ist. Wir bezeichnen $y_n := Tx_n \in TD$. Aus **2.** folgt außerdem, dass TD relativ kompakt ist, das heißt es gibt eine Teilfolge $\{y_{n_k}\}$, so dass $y_{n_k} \to y \in \overline{TD}$ ist. Also ist $d(y_{n_k}, y) \to 0$, aber auch $d(y_{n_k}, x_{n_k}) \to 0$. Dann folgt, dass auch $d(x_{n_k}, y) \to 0$ ist, also gilt $x_{n_k} \to y$. Da $x_{n_k} \in D$ gilt folgt damit $y \in \overline{D} = D$ (da D nach Voraussetzung abgeschlossen ist). Da außerdem T stetig ist, folgt

$$\underbrace{\lim_{k\to\infty} y_{n_k}}_{=y} = \lim_{k\to\infty} Tx_{n_k} \overset{\text{stetig}}{=} T \underbrace{\lim_{k\to\infty} x_{n_k}}_{=y}.$$

Also ist $y = Ty$. □

Wie wird diese Idee implementiert?
Es ist $C(\tilde{I}_\varepsilon(t_0), \mathbb{R}^n)$ ein Banachraum. Es sei $C(\tilde{I}_\varepsilon(t_0), \mathbb{R}^n) \supset \overline{U_{r_\varepsilon}(\mathrm{y}_0)} = \{f \in C(\tilde{I}_\varepsilon(t_0)) : \|f(t) - \mathrm{y}_0\| \le r_\varepsilon\}$ mit $r_\varepsilon = (1-\varepsilon)\operatorname{dist}(\mathrm{y}_0, \partial M)$. Es sei $F = \overline{U_{r_\varepsilon}(\mathrm{y}_0)}$ mit $d = d_C$. Dann folgt, dass (F, d_C) ein vollständiger, abgeschlossener metrischer Raum ist.
Es sei nun $T : D \to F$ mit $D = F = \overline{D}$ gegeben durch

$$(Tf)(t) = \mathrm{y}_0 + \int_{t_0}^{t} \mathrm{v}(\tau, f(\tau))\, \mathrm{d}\tau \qquad , t \in \tilde{I}_\varepsilon(t_0).$$

Wir müssen nun die drei Eigenschaften **1.**, **2.**, **3.** aus dem Fixpunktsatz zeigen. Dann folgt, dass mindestens ein $\mathrm{y} \in D = F$ existiert, so dass $\mathrm{y} = T\mathrm{y}$ ist. Durch Differenzieren etc. folgt dann $\dot{\mathrm{y}}(t) = \mathrm{v}(t, \mathrm{y}(t))$ und $\mathrm{y}(t_0) = \mathrm{y}_0$.

1. Eigenschaft: $T : D \to F$ ist stetig
Für die Abbildung $(D, d_C) \to (F, d_C)$ ist insbesondere die verwendete Metrik d_C gleich. Zu zeigen ist also:

$$\forall_{\varepsilon>0} \exists_{\delta=\delta_\varepsilon(g)} \forall_{f\in D} \underbrace{\|f-g\|_C}_{=d_C(f,g)} < \delta_\varepsilon \;\Rightarrow\; \underbrace{\|Tf - Tg\|_C}_{=d_C(Tf,Tg)} < \varepsilon.$$

Betrachte $\mathrm{v} : \tilde{I}_{\tilde{\varepsilon}}(t_0) \times \overline{U_{r_{\tilde{\varepsilon}}(\mathrm{y}_0)}} \to \mathbb{R}^n$. Insbesondere ist $\tilde{I}_{\tilde{\varepsilon}}(t_0) \times \overline{U_{r_{\tilde{\varepsilon}}(\mathrm{y}_0)}}$ eine kompakte Menge in $\mathbb{R} \times \mathbb{R}^n$. Da v nach Voraussetzung stetig ist folgt damit, dass v gleichmäßig stetig ist.

Frieren wir $\mathbb{v}$ im ersten Argument ein, so bedeutet dies insbesondere

$$\forall_{\varepsilon>0}\exists_{\tilde{\delta}_\varepsilon>0}\forall_{t\in\tilde{I}_{\tilde{\varepsilon}}(t_0)}\forall_{\mathbb{x}',\mathbb{x}''\in\overline{U_{r_{\tilde{\varepsilon}}}(y_0)}}\,\|\mathbb{x}'-\mathbb{x}''\|\leq\tilde{\delta}_\varepsilon \;\Rightarrow\; \|\mathbb{v}(t,\mathbb{x}')-\mathbb{v}(t,\mathbb{x}'')\|<\varepsilon.$$

Wähle $f,g\in D=C(\tilde{I}_\varepsilon(t_0))$ mit Werten in $\overline{U_{r_{\tilde{\varepsilon}}}(\mathbb{y}_0)}$ mit $\|f-g\|_C<\tilde{\delta}_\varepsilon$. Daraus folgt

$$\forall_{t\in\tilde{I}_{\tilde{\varepsilon}}(t_0)}\,\|f(t)-g(t)\|<\tilde{\delta}_\varepsilon.$$

Daraus wiederum sehen wir dann

$$(Tf)(t)-(Tg)(t)=\int_{t_0}^{t}\Big(\mathbb{v}(\tau,f(\tau))-\mathbb{v}(\tau,g(\tau))\Big)\,\mathrm{d}\tau,$$

womit wir dann wie folgt abschätzen können:

$$\begin{aligned}\|(Tf)(t)-(Tg)(t)\| &\leq \left|\int_{t_0}^{t}\Big\|\Big(\mathbb{v}(\tau,f(\tau))-\mathbb{v}(\tau,g(\tau))\Big)\Big\|\,\mathrm{d}\tau\right|\\ &\leq |t-t_0|\varepsilon\leq\varepsilon\frac{r_{\tilde{\varepsilon}}}{C}.\end{aligned}$$

3. Eigenschaft: $\exists\{x_n\}$ mit $x_n\in D$, so dass $d_C(Tx_n,x_n)\stackrel{n\to\infty}{\longrightarrow}0$.
Wir betrachten $t\geq t_0$. Es ist

$$x_n(t)=\begin{cases}\mathbb{y}_0 & \text{für } t\in[t_0,t_0+n^{-1}]\\ \mathbb{y}_0+\int_{t_0+n^{-1}}^{t}\mathbb{v}(\tau,x_n(\tau-n^{-1}))\,\mathrm{d}\tau & \text{für } t>t_0+n^{-1}\end{cases}$$

Wir wollen nun zeigen, dass diese Folge wirklich die gewünschte Eigenschaft erfüllt, das heißt, dass $\|Tx_n-x_n\|_C\to 0$ gilt. Es gilt

$$\begin{aligned}&\|(Tx_n)(t)-x_n(t)\| \stackrel{t>t_0+n^{-1}}{=} \left\|\int_{t_0}^{t}\mathbb{v}(\tau,x_n(\tau))\,\mathrm{d}\tau-\int_{t_0+n^{-1}}^{t}\mathbb{v}(\tau,x_n(\tau-n^{-1}))\,\mathrm{d}\tau\right\|\\ &\leq\left\|\int_{t_0}^{t-n^{-1}}\mathbb{v}(s,x_n(s))\,\mathrm{d}s-\int_{t_0}^{t-n^{-1}}\mathbb{v}(s+n^{-1},x_n(s))\,\mathrm{d}s\right\|+\left\|\int_{t-n^{-1}}^{t}\mathbb{v}(s,x_n(s))\,\mathrm{d}s\right\|\\ &\qquad\leq\int_{t_0}^{t-n^{-1}}\left\|\mathbb{v}(s,x_n(s))\,\mathrm{d}s-\mathbb{v}(s+n^{-1},x_n(s))\right\|\,\mathrm{d}s+Cn^{-1}.\end{aligned}$$

Da $\mathbb{v}$ gleichmäßig stetig ist gilt insbesondere

$$\forall_{\varepsilon>0}\exists_{\delta_\varepsilon>0}\forall_{\mathbb{x}\in\overline{U_{r_{\tilde{\varepsilon}}}(\mathbb{y}_0)}}\forall_{t',t''\in\tilde{I}_{\tilde{\varepsilon}}(t_0)}|t'-t''|<\delta_\varepsilon\;\Rightarrow\;\|\mathbb{v}(t',\mathbb{x})-\mathbb{v}(t'',\mathbb{x})\|<\varepsilon.$$

Wir wählen nun n so, dass $n^{-1}<\delta_\varepsilon$ ist. Damit ist dann
$\|\mathbb{v}(s,x_n(s))-\mathbb{v}(s+n^{-1},x_n(s))\|<\varepsilon$. Also folgt

$$\begin{aligned}\|(Tx_n)(t)-x_n(t)\| &\leq Cn^{-1}+\left|t-n^{-1}-t_0\right|\varepsilon\\ &\leq Cn^{-1}+\varepsilon\frac{r_{\tilde{\varepsilon}}}{C}\stackrel{n\to\infty}{\longrightarrow}0.\end{aligned}$$

Wir wollen das Fixpunktproblem $Ty = y$ $(y \in D)$ lösen. Wir haben jetzt $x_n \in D$: $Tx_n - x_n \xrightarrow{d_C} 0$, *aber:* Wir wissen nicht, ob $\{x_n\}$ konvergiert. *Falls* dem so ist, so ist $Tx_n - x_n = s_n$ und wir wissen $s_n \to 0$, $x_n \to y$ und $Tx_n \to Ty$, dann wäre also wirklich $Ty = y$. Wir werden zeigen, dass eine Teilfolge $x_{n_k} \to y$ konvergiert. Dies ist bereits ausreichend. Dazu benötigen wir noch die **2.** Voraussetzung des Fixpunktsatzes, das heißt, dass TD relativ kompakt ist.
Wir benötigen hierfür ein Kompaktheitskriterium in $C(\tilde{I}_\varepsilon(t_0), \mathbb{R}^n)$. Der Satz von Bolzano gilt in diesem Raum *nicht*.

Definition 9.7 Gleichgradige Stetigkeit
Eine Menge von Funktionen $G \subset C(\tilde{I}, \mathbb{R}^n)$ heißt gleichgradig stetig genau dann, wenn

$$\forall_{\varepsilon>0}\ \exists_{\delta_\varepsilon>0}\ \forall_{t',t''\in\tilde{I}}\ \forall_{f\in G}\ |t'-t''| < \delta_\varepsilon \ \Rightarrow\ \|f(t') - f(t'')\| < \varepsilon.$$

Lemma 9.8
Sei J eine dichte Teilmenge von $\tilde{I}$ und $\{f_n\} \subset G \subset C(\tilde{I}, \mathbb{R}^n)$, wobei G eine gleichgradig stetige Menge sei. Ferner gelte $\forall t \in J$ $f_n(t) \xrightarrow{(\cdot)} f(t)$. Dann folgt:

$$\exists f \in C(\tilde{I}, \mathbb{R}^n)\ :\ f_n \xrightarrow{\|\cdot\|_C} f.$$

Beweis. Dazu müssen wir zeigen, dass $\{f_n\} \in \mathrm{CF}(C(\tilde{I}, \mathbb{R}^n))$ gilt, da wir ja wissen, dass der Raum vollständig ist. Das heißt, wir müssen zeigen, dass

$$\|f_n(t) - f_m(t)\| < \varepsilon \qquad \text{für } m, n \geq N_\varepsilon \text{ und alle } t \in \tilde{I} \text{ gilt.}$$

Wähle $\varepsilon > 0$ und sei $\delta_\varepsilon > 0$ gegeben durch die Definition der gleichgradigen Stetigkeit von G. Für $\tilde{I} = [a, b]$ sei $x_0 = a < x_1 < \ldots < x_n = b$ eine Zerlegung von $\tilde{I}$. Wir wählen den Rang so, dass $\lambda(\text{Zerlegung}) = \max_{k=1,\ldots,n} |x_k - x_{k-1}| < \delta_\varepsilon$ ist. Wir wissen, dass J dicht in $\tilde{I}$ liegt, das heißt, wir finden für alle $k = 1, \ldots, n$ ein $t_k \in \Delta_k \cap J$. Betrachte nun $t \in \tilde{I}$. Dann existiert ein k, so dass $t \in \Delta_k$ ist, dann ist auch $|t - t_k| < \delta_\varepsilon$, und es gilt

$$\|f_n(t) - f_m(t)\| \leq \underbrace{\|f_n(t) - f_n(t_k)\|}_{<\varepsilon} + \underbrace{\|f_m(t_k) - f_m(t)\|}_{<\varepsilon} + \|f_n(t_k) - f_m(t_k)\|.$$

Für $k = 1, \ldots, n$ gilt $f_n(t_k) \to f(t_k)$. Daraus folgt, dass $\{f_n(t_k)\} \in \mathrm{CF}(\mathbb{R}^n)$ ist. Und dies bedeutet, dass für alle $\varepsilon > 0$ ein $N(\varepsilon, k)$ existiert, so dass für alle $\tilde{n}, \tilde{m} \geq N(\varepsilon, k)$ gilt, dass $\|f_{\tilde{n}}(t_k) - f_{\tilde{m}}(t_k)\| < \varepsilon$ ist. Da es nur endlich viele $N(\varepsilon, k)$ gibt wählen wir $N_\varepsilon := \max\{N(\varepsilon, 1), \ldots, N(\varepsilon, n)\}$ und dann gelten für $n, m \geq N_\varepsilon$ alle diese Ungleichungen gleichzeitig und es folgt neben dem Obigen sogar

$$\|f_n(t) - f_m(t)\| \leq \underbrace{\|f_n(t) - f_n(t_k)\|}_{<\varepsilon} + \underbrace{\|f_m(t_k) - f_m(t)\|}_{<\varepsilon} + \underbrace{\|f_n(t_k) - f_m(t_k)\|}_{<\varepsilon}$$

für $n, m \geq N_\varepsilon$. Die linke Seite wird also klein und damit ist das Lemma bewiesen. □

Satz 9.9 Lemma von Arzela-Ascoli
$C(\tilde{I}, \mathbb{R}^n) \supset G$ ist relativ kompakt genau dann, wenn die folgenden Voraussetzungen erfüllt sind:

1. G ist beschränkt, d.h. es existiert ein C, so dass für alle $f \in G$ und $t \in \tilde{I}$ gilt: $\|f(t)\| \leq C$.
2. G ist gleichgradig stetig.

Beweis. Wir werden uns zunächst auf den Beweis einer der Richtungen beschränken, da wir nur die eine Richtung für die Anwendung beim Satz von Peano benötigen. Wir beweisen also zunächst die Richtung „$\Leftarrow$“:
Wir betrachten eine Folge $\{f_n\}$ mit $f_n \in G$. Zu zeigen ist, dass diese Folge eine Teilfolge enthält, die in $C(\tilde{I}, \mathbb{R}^n)$ konvergiert. Es sei $J = \tilde{I} \cap \mathbb{Q}$, diese Menge liegt dicht in $\tilde{I}$. Dann können wir $J = \{\tau_1, \tau_2, \dots\}$ schreiben, da $\mathbb{Q}$ abzählbar ist.
1. Schritt: Betrachte die Folge $\{f_n(\tau_1)\}$. Diese Menge ist nach Voraussetzung **1** in $\mathbb{R}^n$ beschränkt. Nach dem Satz von Bolzano ist sie damit relativ kompakt, d.h. es existiert eine Teilfolge $f_{n_k} =: f_k^{(1)}$, so dass $f_k^{(1)}(\tau_1)$ konvergiert. Dieser Grenzwert sei $f(\tau_1)$.
2. Schritt: Betrachte die Teilfolge $\{f_k^{(1)}(\tau_2)\}$, diese ist wieder beschränkt und wir können wieder eine Teilfolge $f_{k_l}^{(1)} =: f_l^{(2)}$ auswählen, so dass $f_l^{(2)}(\tau_2)$ gegen $f(\tau_2)$ (und auch in τ_1) konvergiert. Wir führen diese Konstruktion nun für $f_k^{(1)}$, $f_l^{(2)}$, $f_m^{(3)}$, $\dots$ fort und erhalten damit eine Diagonalfolge $f_r^{(r)}$. Diese Folge konvergiert auf allen $\tau_r \in J$. Aus dem letzten Lemma folgt damit $f_r^{(r)} \rightrightarrows f$, da G nach Voraussetzung **2** gleichgradig stetig ist. □

Wir können uns damit dem letzten Schritt zum Beweis des Satzes von Peano zuwenden - der zweiten der drei nachzuweisenden Eigenschaften:
2. Eigenschaft: TD ist relativ kompakt.
Dazu müssen wir nun zeigen:

- TD ist beschränkt (2’).
- TD ist gleichgradig stetig (2”).

(2’): Es ist $(Tf)(t) = \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, f(\tau))\, \mathrm{d}\tau$. Wir wollen die Norm davon nun abschätzen:

$$\begin{aligned}\|Tf\|_{C(\tilde{I},\mathbb{R}^n)} &\leq \max_t \left\| \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, f(\tau))\, \mathrm{d}\tau \right\| \\ &\overset{\|\mathbb{v}\| \leq C}{\leq} \|\mathbb{y}_0\| + |\tilde{I}| \cdot C.\end{aligned}$$

Damit ist TD durch drei endliche Konstanten beschränkt, die nicht von f abhängen.
(2”): Wir schätzen ab:

$$\begin{aligned}\|(Tf)(t') - (Tf)(t'')\| &\leq \left\| \int_{t'}^{t''} \mathbb{v}(\tau, f(\tau))\, \mathrm{d}\tau \right\| \\ &\leq |t' - t''| \cdot C \\ &< \varepsilon,\end{aligned}$$

falls $|t' - t''| < \delta_\varepsilon = \varepsilon C^{-1}$, wobei δ_ε nicht mehr von f abhängt.
Damit sind nun alle nachzuweisenden Eigenschaften und damit der Satz von Peano bewiesen. □

Kommentare:

1. Der Beweis des Satzes ist nicht konstruktiv, d.h. während wir beim Satz von Picard-Lindelöf im Laufe des Beweises ein Iterationsverfahren zur Konstruktion der Lösung erhalten haben, erhalten wir beim Satz von Peano keine solche Vorschrift. Dies liegt daran, dass wir aufgrund des Kompaktheitarguments einen Algorithmus benötigen würden, der uns sagt, wie eine Teilfolge konstruiert werden kann. Wir kennen allerdings lediglich die Existenz einer solchen Teilfolge.
2. Der Beweis der Sätze von Picard-Lindelöf und Peano zeigt, wie wichtig Fixpunktsätze sind. In der modernen Mathematik stellen diese Fixpunktsätze ein sehr wichtiges Werkzeug dar.

Satz 9.10 Satz von Schauder
Sei B ein Banachraum und $B \supset D$ sei abgeschlossen und konvex. Ferner sei $T : D \to D$ stetig und auf D kompakt. Dann existiert (mindestens) ein y mit $Ty = y$.

Spezialfall: Sei $\overline{U_1(0)} = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ die abgeschlossene Einheitskugel im endlichdimensionalen $\mathbb{R}^n$. Weiter sei $T : \overline{U_1(0)} \to \overline{U_1(0)}$ eine stetige Abbildung. Damit lässt sich der Satz von Schauder anwenden, d.h. es gibt ein $y \in \overline{U_1(0)}$ mit $Ty = y$.
Anschaulich bedeutet dies: Bildet man die abgeschlossene Einheitskugel stetig auf sich selbst ab, so gibt es mindestens einen Punkt, der dabei nicht bewegt wird. Dieser Spezialfall ist als Brouwerscher Fixpunktsatz bekannt.

Vervollständigung des Beweises von Arzela-Ascoli:
Wir wollen uns im Folgenden wieder dem Satz von Arzela-Ascoli widmen, von dem wir zunächst nur einen Teil bewiesen hatten.

Definition 9.11 ε-Netz
Sei (B, d) ein metrischer Raum und $G \subset B$ eine Teilmenge. Eine Teilmenge $G_\varepsilon \subset G$ heißt ε-Netz genau dann, wenn für alle $x \in G$ ein $y = y(\varepsilon, x) \in G_\varepsilon$ existiert, so dass $d(x, y) < \varepsilon$ ist.

Lemma 9.12
Es sei $G \subset B$ mit G relativ kompakt und B vollständig. Dann existiert für alle $\varepsilon > 0$ ein *endliches* ε-Netz.

Beweis. Angenommen, es gibt ein $\varepsilon_0 > 0$, so dass es kein endliches ε_0-Netz gibt. Betrachte $y_0 \in G$, dann ist $G \setminus U_{\varepsilon_0}(y_0) \neq \emptyset$. Betrachte dann $y_1 \in G \setminus U_{\varepsilon_0}(y_0)$, so ist $G \setminus (U_{\varepsilon_0}(y_0) \cup U_{\varepsilon_0}(y_1)) \neq \emptyset$. Wir erhalten also eine Folge $y_1, \ldots, y_k, \ldots$ aus G mit der Eigenschaft, dass $d(y_k, y_l) \geq \varepsilon_0$ für alle $l \neq k$ ist. Dann gilt dies aber auch für jede Teilfolge y_{k_n}, womit keine Teilfolge existieren kann, die eine Cauchyfolge ist. Dann kann keine in B konvergente Teilfolge existieren, womit G nicht relativ kompakt ist. □

Beweis von „⇒“ von Arzela-Ascoli:
Jede kompakte Menge ist beschränkt. Da $G \subset \overline{G}$ und $\overline{G}$ kompakt ist, ist damit auch G beschränkt. Es bleibt also noch die gleichgradige Stetigkeit nachzuweisen.
Sei $\varepsilon > 0$ und $f_1, \dots, f_{N(\varepsilon)}$ das endliche ε-Netz $G_\varepsilon \subset G$, das heißt für alle $f \in G$ existiert ein $f_k \in G_\varepsilon$, so dass $\|f - f_k\|_{C(\tilde{I},\mathbb{R}^n)} < \varepsilon$ ist (*). Alle $f_1, \dots, f_{N(\varepsilon)}$ sind stetig auf dem kompakten Intervall $\tilde{I}$ und damit gleichmäßig stetig. Für das gewählte $\varepsilon > 0$ existieren also $\delta(\varepsilon, k) > 0$, so dass für alle $t', t'' \in \tilde{I}$ mit $|t'-t''| < \delta(\varepsilon, k)$ gilt: $|f_k(t') - f_k(t'')| < \varepsilon$. Da es nur endlich viele solcher $\delta(\varepsilon, k)$ gibt können wir das Minimum $\min_{k=1,\dots,N(\varepsilon)} \delta(\varepsilon, k) =: \delta(\varepsilon) > 0$ wählen. Dann gilt $|t'-t''| < \delta(\varepsilon) \Rightarrow |f_k(t') - f_k(t'')| < \varepsilon$ für alle $k = 1, \dots, N(\varepsilon)$.
Wähle nun $t', t'' \in \tilde{I}$ mit $|t' - t''| < \delta(\varepsilon)$ und wähle $f \in G$, so gibt es ein f_k mit der obigen Eigenschaft (*). Dann folgt

$$\|f(t') - f(t'')\| \leq \underbrace{\|f(t') - f_k(t')\|}_{<\varepsilon} + \underbrace{\|f(t'') - f_k(t'')\|}_{<\varepsilon} + \underbrace{\|f_k(t') - f_k(t'')\|}_{<\varepsilon}.$$

Damit ist der Satz vollständig bewiesen. □

9.5 Stetigkeit der Lösung des Cauchyproblems bezüglich der Anfangsdaten

Gegeben sei das Cauchyproblem $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ und $\mathbb{y}(t_0) = \mathbb{y}_0$. Es sei $(t_0, \mathbb{y}_0) \in \mathrm{int}(I) \times \mathrm{int}(M)$. Außerdem seien die Voraussetzungen **1.** - **3.** des Satzes von Picard-Lindelöf gegeben. Dann existiert $\mathbb{y}(t) = \mathbb{y}(t; t_0, \mathbb{y}_0)$ auf I_ε und ist dort eindeutig.

Satz 9.13
$\mathbb{y}(t; t_0, \mathbb{y}_0)$ ist stetig in allen drei Argumenten.

Beweis. Wir werden den Beweis nur skizzieren. Wir betrachten noch einmal das Verfahren aus dem Beweis des Satzes von Picard-Lindelöf. Es ist $h_0(t) = \mathbb{y}_0$ stetig in $(t, t_0, \mathbb{y}_0)$. Nun war $h_1(t) = \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, h_0(\tau)) \, d\tau$, und dies ist ebenfalls wieder stetig in $(t, t_0, \mathbb{y}_0)$. Dies kann man fortführen, d.h. $h_j(t) = \mathbb{y}_0 + \int_{t_0}^t \mathbb{v}(\tau, h_{j-1}(\tau)) \, d\tau$ ist immer stetig in $(t, t_0, \mathbb{y}_0)$. Wir kennen neben dem Lösungsverfahren aber auch eine Fehlerabschätzung:

$$\begin{aligned}
\|\mathbb{y}(t, t_0, \mathbb{y}_0) - h_j(t, t_0, \mathbb{y}_0)\| &\leq \frac{1}{(j+1)!} C L^j |t - t_0|^{j+1} \\
&\leq C \cdot |I_\varepsilon| \cdot \frac{(L \cdot |I_\varepsilon|)^j}{(j+1)!} \\
&\xrightarrow{j\to\infty} 0,
\end{aligned}$$

wobei diese Konvergenz gleichmäßig bezüglich $t \in I_\varepsilon$ (aber auch t_0 und $\mathbb{y}_0$) ist. Damit ist der Grenzwert stetig. □

9.6 Techniken zum Lösen von Differentialgleichungen

Wir werden uns im Folgenden mit einigen Lösungstechniken für Differentialgleichungen beschäftigen. Dabei beschränken wir uns auf Differentialgleichungen erster Ordnung. Für das Verfahren der Trennung der Veränderlichen betrachten wir also

$$y'(x) = f(x, y)$$

und sagen, dass diese Differentialgleichung trennbare Veränderliche besitzt, wenn man $f(x, y) = h(y)g(x)$ mit $h(y) \not\equiv 0$ schreiben kann. Dann gilt

$$\begin{aligned} \frac{\mathrm{d}y}{\mathrm{d}x} &= h(y)g(x) \\ \Leftrightarrow \int \frac{1}{h(y)}\,\mathrm{d}y &= \int g(x)\,\mathrm{d}x + c. \end{aligned}$$

Beispiel: Es sei die Differentialgleichung $(x^2 - 1)y' = -2xy^2$ mit der Anfangsbedingung $y(0) = 1$ gegeben. Wir sehen, dass diese Differentialgleichung der obigen Form genügt, wenn man $g(x) = \frac{-2x}{x^2-1}$ und $h(y) = y^2$ schreibt. Es folgt also

$$\begin{aligned} \int \frac{1}{y^2}\,\mathrm{d}y &= -\int \frac{2x}{x^2-1}\,\mathrm{d}x + c \\ \Leftrightarrow -\frac{1}{y} &= -\ln|x^2 - 1| + c. \end{aligned}$$

Damit erhalten wir $y(x) = \frac{1}{\ln|x^2-1|-c}$. Einsetzen der Anfangsbedingung liefert $c = -1$ und wir erhalten die Lösung

$$y(x) = \frac{1}{\ln|x^2 - 1| + 1}.$$

Beispiel: Es sei die Differentialgleichung $y'(x) = 2\sqrt{y(x)}$ mit der Anfangsbedingung $y(0) = 1$ gegeben. In diesem Fall ist $g(x) = 2$ und $h(y) = \sqrt{y}$ und wir erhalten

$$\begin{aligned} \int \frac{1}{\sqrt{y}}\,\mathrm{d}y &= \int 2\,\mathrm{d}x + c \\ \Leftrightarrow 2\sqrt{y} &= 2x + c. \end{aligned}$$

Dies liefert uns also $y(x) = \left(x + \frac{c}{2}\right)^2$ und mit der Anfangsbedingung folgt $c = \pm 2$, wir erhalten also als Lösung

$$y(x) = (x \pm 1)^2 .$$

Dies sind aber nicht alle Lösungen, denn es gibt tatsächlich noch zwei weitere Lösungen:

$$y(x) = \begin{cases} (x \pm 1)^2 & \text{für } -1 < x < 1 \\ 0 & \text{sonst} \end{cases}$$

Beispiel: Es sei die Differentialgleichung $xy'(x)+y(x) = (y(x))^2$ mit der Anfangsbedingung $y(1) = \frac{1}{2}$ und der Einschränkung $x \geq 1$ gegeben. Wir trennen die Variablen mit $g(x) = \frac{1}{x}$ und $h(y) = y^2 - y$ und erhalten

$$\begin{aligned} \int \frac{1}{y^2-y}\,\mathrm{d}y &= \int \frac{1}{x}\,\mathrm{d}x + c \\ \Leftrightarrow \int \left(\frac{1}{y-1} + \frac{-1}{y}\right)\mathrm{d}y &= \int \frac{1}{x}\,\mathrm{d}x + c \\ \Leftrightarrow \ln|y-1| - \ln|y| &= \ln|x| + c \\ \Leftrightarrow \frac{|y-1|}{|y|} &= \tilde{c}|x| = \tilde{c}x. \end{aligned}$$

Mit der Anfangsbedingung folgt, dass $0 \leq y \leq 1$ gilt, wir müssen also keine Fallunterscheidung durchführen und erhalten $y(x) = \frac{1}{1+\tilde{c}x}$. Aus der Anfangsbedingung folgt dann $\tilde{c} = 1$ und wir erhalten als Lösung

$$y(x) = \frac{1}{1+x}.$$

Lineare homogene und inhomogene Differentialgleichungen

Die allgemeine Form lautet

$$a_1(x)y'(x) + a_0(x)y(x) = g(x),$$

wobei $a_1 \not\equiv 0$ und $x \in D$ gilt. Dabei ist D der gemeinsame Definitionsbereich von a_1, a_0 und g.

Wir nennen diese Differentialgleichung

- linear in y,
- homogen, falls $g \equiv 0$,
- inhomogen, falls $g \not\equiv 0$.

Satz 9.14

Jede Lösung $y(x)$ der Differentialgleichung $a_1y' + a_0y = g$ ist gegeben durch $y(x) = y_h(x) + y_p(x)$, wobei y_h Lösung der homogenen und y_p spezielle Lösung der inhomogenen Differentialgleichung ist.

Beweis. Wir definieren zunächst den Operator $L[y] := a_1y' + a_0y$. Für die Richtung „$\Rightarrow$" sei y_h Lösung der homogenen DGL und y_p spezielle Lösung der inhomogenen DGL, d.h. $L[y_h] = 0$ und $L[y_p] = g$. Damit folgt aus Linearitätsgründen für $y = y_h + y_p$, dass $L[y] = L[y_h + y_p] = L[y_h] + L[y_p] = 0 + g = g$, also löst y die Differentialgleichung.
Für die Richtung „$\Leftarrow$" seien $y(x)$ und $\tilde{y}(x)$ Lösungen der Differentialgleichung. Dann gilt wieder aus Linearitätsgründen $L[y - \tilde{y}] = L[y] - L[\tilde{y}] = g - g = 0$. Die Differenz $y - \tilde{y}$ löst also die homogene DGL, es ist damit also $y - \tilde{y} = y_h \;\Leftrightarrow\; y = y_h + \tilde{y}$. $\square$

Lösungen der homogenen Differentialgleichung: Es ist also $g(x) \equiv 0$ und wir haben damit die Differentialgleichung $a_1(x)y'(x) + a_0(x)y(x) = 0$ vorliegen. Hier können wir aber eine Trennung der Veründlichen durchführen, indem wir schreiben:

$$\begin{aligned} y'(x) &= -\frac{a_0(x)}{a_1(x)}y(x) \\ \Leftrightarrow \int \frac{1}{y}\,\mathrm{d}y &= -\int \frac{a_0(x)}{a_1(x)}\mathrm{d}x + \tilde{c} \\ \Leftrightarrow y_h(x) &= c\exp\left(-\int \frac{a_0(x)}{a_1(x)}\,\mathrm{d}x\right) = c\tilde{y}_h(x). \end{aligned}$$

Wir führen jetzt eine Variation der Konstanten durch. Dazu setzen wir $y_p(x) = c(x)\tilde{y}_h(x)$. Es soll gelten: $a_1 y_p' + a_0 y_p = g$. Das ist äquivalent zu

$$\begin{aligned} a_1(x)\left(c'(x)\tilde{y}_h(x) + c(x)\tilde{y}_h'(x)\right) + a_0(x)c(x)\tilde{y}_h(x) &= g(x) \\ \Leftrightarrow a_1(x)c'(x)\tilde{y}_h(x) + c(x)\underbrace{\left[a_1(x)\tilde{y}_h'(x) + a_0(x)\tilde{y}_h(x)\right]}_{=0} &= g(x) \\ \Leftrightarrow a_1(x)c'(x)\tilde{y}_h(x) &= g(x) \\ \Leftrightarrow c(x) &= \int \frac{g(x)}{a_1(x)\tilde{y}_h(x)}\,\mathrm{d}x \end{aligned}$$

Also ist $\tilde{y}_h(x) = \exp\left(-\int \frac{a_0(x)}{a_1(x)}\,\mathrm{d}x\right)$ und $y_p(x) = \int \frac{g(x)}{\tilde{y}_h(x)a_1(x)}\,\mathrm{d}x \cdot \tilde{y}_h(x)$. Die allgemeine Lösung ist $y(x) = c\tilde{y}_h(x) + y_p(x)$.

Beispiel: Gegeben sei $y' + yx^2 = 2x^2$, dann ist $a_1(x) = 1$, $a_0(x) = x^2$ und $g(x) = 2x^2$. Wir berechnen die Lösung der homogenen Differentialgleichung:

$$y_h(x) = c\exp\left(-\int x^2\,\mathrm{d}x\right) = c\exp\left(-\frac{1}{3}x^3\right).$$

Die Variation der Konstanten liefert

$$c(x) = \int \frac{2x^2}{\exp\left(-\frac{1}{3}x^3\right)}\,\mathrm{d}x = 2\int x^2\exp\left(\frac{1}{3}x^3\right)\,\mathrm{d}x = 2\exp\left(\frac{1}{3}x^3\right).$$

Damit erhalten wir als Lösung

$$\begin{aligned} y(x) &= c\exp\left(-\frac{1}{3}x^3\right) + 2\exp\left(\frac{1}{3}x^3\right)\exp\left(-\frac{1}{3}x^3\right) \\ &= c\exp\left(-\frac{1}{3}x^3\right) + 2. \end{aligned}$$

Beispiel: Gegeben sei $xy'(x) + (1+x)y(x) = 3x^2e^{-x}$. Hier ist $a_1(x) = x$, $a_0(x) = 1 + x$ und $g(x) = 3x^2e^{-x}$. Die Lösung der homogenen Differentialgleichung ist dann

$$y_h(x) = c\exp\left(-\int \frac{1+x}{x}\,\mathrm{d}x\right) = c\exp\left(-\ln|x| - x\right) = \frac{c}{x}e^{-x}.$$

Die Variation der Konstanten liefert uns

$$c(x) = \int \frac{3x^2 e^{-x}}{x \cdot \frac{1}{x} e^{-x}} \, \mathrm{d}x = \int 3x^2 \, \mathrm{d}x = x^3.$$

Also ist $y_p(x) = x^3 \cdot \frac{1}{x} e^{-x} = x^2 e^{-x}$ und wir erhalten als allgemeine Lösung

$$y(x) = c\frac{e^{-x}}{x} + x^2 e^{-x}.$$

Nichtlineare Differentialgleichungen erster Ordnung

Die BERNOULLI-Gleichungen haben die allgemeine Form

$$y'(x) + a(x)y(x) = b(x)y^n(x) \qquad n \in \mathbb{N}_0.$$

Für $n = 0$ oder $n = 1$ haben wir eine lineare Differentialgleichung, von der wir die Lösung bereits kennen. Für $n \geq 2$ schreiben wir die Gleichung um in

$$\frac{y'(x)}{y^n(x)} + a(x)y^{1-n}(x) = b(x)$$

und führen jetzt die Substitution $z(x) = y^{1-n}(x)$ durch.
Dann ist $z'(x) = (1-n)y^{-n}(x)y'(x)$, das heißt die Differentialgleichung ist äquivalent zu

$$\frac{z'(x)}{1-n} + a(x)z(x) = b(x).$$

Das ist eine lineare Differentialgleichung, die wir lösen können.
Beispiel: Gegeben sei $y'(x) + xy(x) = xy^3$. Hier ist $n = 3$ und $a(x) = b(x) = x$. Wir substituieren also $z(x) = y^{-2}(x)$ und bestimmen jetzt die Lösung der homogenen Differentialgleichung in z:

$$\frac{z_h'(x)}{-2} + xz_h(x) = 0.$$

Also ist

$$z_h(x) = c \exp\left(-\int \frac{x}{-\frac{1}{2}} \, \mathrm{d}x\right) = c \exp\left(x^2\right).$$

Die Variation der Konstanten liefert uns dann

$$c(x) = \int \frac{x}{\left(-\frac{1}{2}\right)\exp(x^2)} \, \mathrm{d}x = -\int 2x \exp(-x^2) \, \mathrm{d}x = \exp(-x^2).$$

Damit ist dann also

$$z(x) = c\exp(x^2) + \exp(-x^2)\exp(x^2) = c\exp(x^2) + 1,$$

wobei wir diese Lösung in z nun noch resubstituieren müssen. Damit ist

$$y(x) = \frac{1}{\sqrt{z(x)}} = \frac{1}{\sqrt{ce^{x^2}+1}}.$$

Lineare Differentialgleichungen zweiter Ordnung $(y'' + a_1 y' + a_0 y = g(x))$

1. Homogene Gleichung

Die homogene DGL lautet in diesem Fall

$$y'' + a_1 y' + a_0 y = 0 \qquad a_1, a_0 \in \mathbb{R}.$$

Definition 9.15 Fundamentalsystem

$y_1(x)$, $y_2(x)$ seien Lösungen der homogenen Differentialgleichung von oben, wobei y_1, y_2 linear unabhängig sind, d.h.

$$c_1 y_1 + c_2 y_2 = 0 \;\Rightarrow\; c_1 = c_2 = 0.$$

Dann heißt $\{y_1(x), y_2(x)\}$ ein Fundamentalsystem der homogenen DGL und

$$y(x) = c_1 y_1(x) + c_2 y_2(x) \qquad c_1, c_2 \in \mathbb{R}$$

ist die allgemeine Lösung der homogenen Differentialgleichung.

Die charakteristische Gleichung lautet

$$\lambda^2 + a_1\lambda + a_0 = 0.$$

1. $\lambda_1 \neq \lambda_2 \;\Rightarrow\; y_1(x) = e^{\lambda_1 x}$, $y_2(x) = e^{\lambda_2 x}$.
2. $\lambda_1 = \lambda_2 \;\Rightarrow\; y_1(x) = e^{\lambda x}$, $y_2(x) = xe^{\lambda x}$.

Beispiel: Betrachte $y''+2y'-3y = 0$. Die charakteristische Gleichung lautet $\lambda^2+2\lambda-3 = 0$, damit folgt $\lambda_1 = 1$ und $\lambda_2 = -3$. Wir erhalten die Fundamentallösungen $y_1(x) = e^x$ und $y_2(x) = e^{-3x}$. Die allgemeine Lösung lautet also

$$y(x) = c_1 e^x + c_2 e^{-3x} \qquad c_1, c_2 \in \mathbb{R}.$$

Beispiel: Gegeben sei die DGl $y'' - 2y' + y = 0$. Für die charakteristische Gleichung $\lambda^2 - 2\lambda + 1 = 0$ folgt $\lambda_1 = \lambda_2 = 1$. Damit haben wir $y_1(x) = e^x$ und $y_2(x) = xe^x$ als Fundamentalsystem und erhalten als allgemeine Lösung

$$y(x) = c_1 e^x + c_2 x e^x \qquad c_1, c_2 \in \mathbb{R}.$$

2. Inhomogene Gleichung

Gegeben sei nun die inhomogene Differentialgleichung

$$y'' + a_1 y' + a_0 y = g(x).$$

Die allgemeine Lösung y lässt sich als Summe der homogenen Lösung und einer partikulären Lösung schreiben:

$$y(x) = y_h(x) + y_p(x).$$

Wir wollen nun verschiedene Verfahren zur Lösung vorstellen.

1) Variation der Konstanten

Satz 9.16
Es sei $\{y_1(x), y_2(x)\}$ ein Fundamentalsystem der zugehörigen homogenen DGL. Gelten die folgenden Bedingungen:

1. $c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0$
2. $c_1'(x)y_1'(x) + c_2'(x)y_2'(x) = g(x)$,

so folgt, dass

$$y_p(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

eine Lösung der inhomogenen Differentialgleichung ist.

Beweis. Es ist $y_p'(x) = \cancel{c_1'(x)y_1(x)} + c_1(x)y_1'(x) + \cancel{c_2'(x)y_2(x)} + c_2(x)y_2'(x)$ und $y_p''(x) = \ldots = g(x) + c_1(x)y_1''(x) + c_2(x)y_2''(x)$. Damit erhalten wir

$$\begin{aligned} y_p'' + a_1y_p' + a_0y_p &= g(x) + c_1(x)y_1''(x) + c_2(x)y_2''(x) \\ &+ a_1\Big(c_1(x)y_1'(x) + c_2(x)y_2'(x)\Big) + a_0\Big(c_1(x)y_1(x) + c_2(x)y_2(x)\Big) \\ &= g(x) + c_1(x)\Big(\underbrace{y_1''(x) + a_1y_1'(x) + a_0y_1(x)}_{=0}\Big) + c_2(x)\Big(\underbrace{y_2''(x) + a_1y_2'(x) + a_0y_2(x)}_{=0}\Big) \\ &= g(x). \end{aligned}$$

Damit ist der Satz bewiesen. □

Beispiel: Gegeben sei $y'' - 2y' + y = \frac{e^x}{x}$ für $x \neq 0$. Die charakteristische Gleichung $\lambda^2 - 2\lambda + 1 = 0$ hat die Lösungen $\lambda_1 = \lambda_2 = 1$. Damit erhalten wir $y_1(x) = e^x$ und $y_2(x) = xe^x$ als Fundamentalsystem. Wir wollen nun eine partikuläre Lösung bestimmen. Es muss gelten

1. $c_1'(x)e^x + c_2'(x)xe^x = 0$
2. $c_1'(x)e^x + c_2'(x)(e^x + xe^x) = \frac{e^x}{x}$.

Dieses Gleichungssystem können wir lösen und erhalten dann $c_2'(x) = \frac{1}{x}$ und damit $c_2(x) = \ln|x|$. Ebenso folgt $c_1'(x) = -1$ und damit $c_1(x) = -x$. Damit erhalten wir $y(x) = y_h(x) + y_p(x) = k_1e^x + k_2xe^x - xe^x + \ln(|x|)xe^x$ mit $k_1, k_2 \in \mathbb{R}$.

2) Ansatzmethode

Es sei $y'' + a_1y' + a_0y = g(x)$, wobei $g(x) = e^{qx}(\alpha_0 + \alpha_1x + \ldots + \alpha_mx^m)$. Dann ist $y_p(x) = x^le^{qx}(\beta_0 + \ldots + \beta_mx^m)$.

1. Ist q keine Lösung des charakteristischen Polynoms, so ist $l = 0$.
2. Ist q Lösung des charakteristischen Polynoms mit Vielfachheit n, so ist $l = n$.

Beispiel: Es sei $y'' - 2y' + 2y = 2x^2 e^x$, $q = 1$. Das charakteristische Polynom $\lambda^2 - 2\lambda + 2 = 0$ hat die Lösungen $\lambda_1 = 1 + i$, $\lambda_2 = 1 - i$, es ist also $l = 0$. Damit folgt

$$\begin{aligned} y_p(x) &= e^x(\beta_0 + \beta_1 x + \beta_2 x^2) \\ y_p'(x) &= e^x(\beta_0 + \beta_1 x + \beta_2 x^2) + e^x(\beta_1 + 2\beta_2 x) \\ y_p''(x) &= e^x(\beta_0 + \beta_1 x + \beta_2 x^2) + 2e^x(\beta_1 + 2\beta_2 x) + e^x \cdot 2\beta_2. \end{aligned}$$

Wir setzen dies in die DGL ein und erhalten dadurch letztlich $\beta_2 = 2$, $\beta_1 = 0$ und $\beta_0 = -4$. Damit ist dann also $y_p(x) = e^x(-4 + 2x^2)$ und wir erhalten die allgemeine Lösung durch $y(x) = y_h(x) + y_p(x)$.

9.7 Charakterisierung der Lösungen des Cauchyproblems

Gegeben sei ein Cauchyproblem $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ mit $\mathbb{y}(t_0) = \mathbb{y}_0$, für welches die Voraussetzungen von Picard-Lindelöf erfüllt sind. Es sei $\mathbb{y}(t; t_0, \mathbb{y}_0)$ eine Lösung des Cauchyproblems mit $(t_0, \mathbb{y}_0)$. Ferner sei $\mathbb{y}_1 = \mathbb{y}(t_1; t_0, \mathbb{y}_0)$. Da wir den Satz von Picard-Lindelöf anwenden können ist dann wegen der Eindeutigkeit aber
$\mathbb{y}(t; t_0, \mathbb{y}_0) = \mathbb{y}(t; t_1, \mathbb{y}_1)$. Setzen wir nun $t = t_0$ ein, so erhalten wir

$$\underbrace{\mathbb{y}(t_0; t_0, \mathbb{y}_0)}_{=\mathbb{y}_0} = \mathbb{y}(t_0; t_1, \mathbb{y}(t_1; t_0, \mathbb{y}_0)) \qquad \text{für alle } t_1 \in I.$$

Es sei nun eine Abbildung ψ gegeben mit

$$\begin{aligned} \psi\colon I \times M &\to \mathbb{R}^n \\ \psi(\tau, z) &= \mathbb{y}(t_0; \tau, z), \end{aligned}$$

und wir bilden

$$\psi(t, z)\Big|_{z = \mathbb{y}(t; t_0, \mathbb{y}_0)} = \underbrace{\mathbb{y}(t_0; t, \mathbb{y}(t; t_0, \mathbb{y}_0))}_{=\mathbb{y}_0},$$

das heißt, dass ψ auf den Lösungen des Cauchyproblems konstant ist. Es gibt durchaus mehr solcher Funktionen, von denen andere physikalisch sinnvoller sind. Diese Konstruktion soll jedoch die nun folgende Definition motivieren.

Definition 9.17 (allgemeines) Integral
Ein (allgemeines) Integral einer Differentialgleichung $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ nennt man eine Abbildung $\psi\colon I \times M \to \mathbb{R}^n$ ($M \subset \mathbb{R}^n$), welche auf allen Integralkurven einen konstanten Wert annimmt.
Ist $\psi \not\equiv \text{const}$, so nennt man ψ nicht-trivial.

Definition 9.18 erstes Integral
Ein erstes Integral (der Bewegung) nennt man eine Abbildung $\psi_k\colon I \times M \to \mathbb{R}$, welche die selbe Eigenschaft wie in der letzten Definition hat, d.h., die auf allen Integralkurven einen konstanten Wert annimmt.

Definition 9.19 abhängige erste Integrale
Zwei erste Integrale ψ_k, ψ_l heißen abhängig, falls

$$\exists\, g\colon \mathbb{R} \to \mathbb{R} \ :\ \psi_l = g \circ \psi_k.$$

Beispiel: Betrachte die Auslenkung q einer Feder mit der Masse m. Der Impuls ist $p(t) = m\dot{q}(t)$ und es ist $\mathbb{y}(t) = \begin{pmatrix} q(t) \\ p(t) \end{pmatrix}$. Damit erhalten wir

$$\dot{\mathbb{y}}(t) = \begin{pmatrix} \dot{q}(t) \\ \dot{p}(t) \end{pmatrix} = \underbrace{\underbrace{\begin{pmatrix} 0 & \frac{1}{m} \\ -k & 0 \end{pmatrix}}_{=:A} \underbrace{\begin{pmatrix} q(t) \\ p(t) \end{pmatrix}}_{=:\mathbb{y}}}_{\mathbb{v}(\mathbb{y}(t)) = A\mathbb{y}(t)}.$$

Damit ist

$$E = \frac{p^2(t)}{2m} + \frac{k}{2} q^2(t)$$

ein erstes Integral. Dies sieht man, indem man diese Lösung in die allgemeine DGL einsetzt, worauf wir hier verzichten wollen. Nun ist

$$\begin{aligned} E(t) &= \frac{p^2(t)}{2m} + \frac{kq^2(t)}{2} \qquad \Big|\ \frac{\mathrm{d}}{\mathrm{d}t} \\ \Longrightarrow \quad \frac{\mathrm{d}E(t)}{\mathrm{d}t} &= \frac{p(t)\dot{p}(t)}{m} + kq(t)\dot{q}(t) \\ &= \frac{p(t)(-kq(t))}{m} + \frac{kq(t)p(t)}{m} \\ &= 0, \end{aligned}$$

wobei die folgenden Identitäten verwendet werden:

$$\begin{cases} \dot{p}(t) = -kq(t) \\ \dot{q}(t) = \frac{1}{m} p(t) \end{cases}$$

Wir erinnern uns nun zunächst daran, was unsere Kurzschreibweise für das Cauchyproblem bedeutete:

$$\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t)) \;\Leftrightarrow\; \begin{cases} \dot{y}_1(t) = v_1(t, y_1(t), \ldots, y_n(t)) \\ \ldots \\ \ldots \\ \dot{y}_n(t) = v_n(t, y_1(t), \ldots, y_n(t)) \end{cases}$$

Gegeben sei nun $\psi(t, \mathbb{y}(t)) = \text{const}$, wobei $\mathbb{y}(t)$ eine Lösung der Differentialgleichung sein soll. Dann wissen wir:

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\psi(t, \mathbb{y}(t)) \\ &= \frac{\partial \psi}{\partial t} \cdot 1 + \frac{\partial \psi}{\partial y_1} \cdot \frac{\mathrm{d}y_1}{\mathrm{d}t} + \ldots + \frac{\partial \psi}{\partial y_n}\frac{\mathrm{d}y_n}{\mathrm{d}t} \\ &= \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial y_1} v_1 + \ldots + \frac{\partial \psi}{\partial y_n} v_n. \end{aligned}$$

Wir haben also ein System von Funktionen $\psi_i(t, y_1(t), \ldots, y_n(t)) = c_i$ mit $1 \leq i \leq n$. Falls der Satz über implizite Funktionen anwendbar ist, folgt insbesondere $\det \frac{\mathrm{D}\psi}{\mathrm{D}\mathbb{y}} \neq 0$, das heißt, dass $\mathbb{y} = \mathbb{y}(t)$ lokal auflösbar ist.

9.8 Differenzierbarkeit der Lösung nach den Anfangsbedingungen

Gegeben sei ein (parameterabhängiges) Cauchyproblem $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t), \lambda)$ mit $\mathbb{y}(t_0) = \eta$. Die Lösung hängt damit ab von $\mathbb{y} = \mathbb{y}(t; t_0, \eta, \lambda)$. Sind die Voraussetzungen des Satzes von Picard-Lindelöf gleichmäßig in $\lambda \in \mathbb{R}$ erfüllt, so ist $\mathbb{y}(t; t_0, \eta, \lambda)$ stetig in t_0, η und λ.

Sei nun $\lambda \in D$. Es gelten

1. $\mathbb{v} \in C(I \times M \times D, \mathbb{R}^n)$,
2. $\|\mathbb{v}(t, \mathbb{x}, \lambda)\| \leq C$ für alle $(t, \mathbb{x}, \lambda)$,
3. $\|\mathbb{v}(t, \mathbb{x}', \lambda) - \mathbb{v}(t, \mathbb{x}'', \lambda)\| \leq L \|\mathbb{x}' - \mathbb{x}''\|$ für alle $(t, \mathbb{x}', \lambda)$ und $(t, \mathbb{x}'', \lambda)$.

Dann folgt die Stetigkeit in (t_0, η, λ).

Frage: $\mathbb{v}$ ist differenzierbar in $\mathbb{y}$, $\lambda \overset{?}{\Longrightarrow} \mathbb{y}(t; t_0, \eta, \lambda)$ differenzierbar.

Annahme: $\mathbb{v}$ sei differenzierbar und die Ableitung sei mit dem folgenden Integral vertauschbar:

$$\mathbb{y}(t; t_0, \eta, \lambda) = \eta + \int_{t_0}^{t} \mathbb{v}(\tau, \mathbb{y}(\tau, t_0, \eta, \lambda), \lambda)\ \mathrm{d}\tau.$$

Formales Differenzieren ergibt

$$\underbrace{\frac{\partial \mathbb{y}}{\partial t_0}}_{=g} = 0 + \int_{t_0}^{t} \frac{\mathrm{D}\,\mathbb{v}}{\mathrm{D}\,\mathbb{y}} \underbrace{\frac{\partial \mathbb{y}}{\partial t_0}}_{=g}\ \mathrm{d}\tau - \mathbb{v}(t_0, \mathbb{y}(t_0, t_0, \eta, \lambda), \lambda),$$

sowie

$$\underbrace{\frac{\mathrm{D}\,\mathbb{y}}{\mathrm{D}\,\eta}}_{=\omega} = \mathbb{1} + \int_{t_0}^{t} \frac{\mathrm{D}\,\mathbb{v}}{\mathrm{D}\,\mathbb{y}} \underbrace{\frac{\mathrm{D}\,\mathbb{y}}{\mathrm{D}\,\eta}}_{=\omega} \,\mathrm{d}\tau,$$

wobei $\mathbb{1}$ die $(n \times n)$-Einheitsmatrix ist.

Satz 9.20
Die Voraussetzungen des Satzes von Picard-Lindelöf seien gleichmäßig erfüllt, d.h. es gelten die drei Bedingungen 1. - 3., die zu Anfang deses Unterkapitels definiert wurden. Ferner sei $\frac{\partial v_k}{\partial y_l}$ stetig in $I \times M \times D$. Dann folgt, dass $\mathbb{y}(t, t_0, \eta, \lambda)$ im Existenzbereich bezüglich t_0, η, λ differenzierbar ist. Diese (partiellen) Ableitungen sind stetig und erfüllen die beiden obigen Ableitungsgleichungen nach t_0 und η.
Falls $\frac{\partial v_k}{\partial \lambda}$ stetig ist, so ist auch $\frac{\partial \mathbb{y}}{\partial \lambda}$ stetig und es gilt die Ableitungsgleichung nach λ (die oben nicht aufgeschrieben und zur Übung überlassen wird).

Beweis. Der Beweis findet sich in der Literatur. □

Satz 9.21
Ist zudem $\mathbb{v}$ *analytisch* in $\lambda \in D \subset \mathbb{C}$, so ist im Existenzbereich auch $\mathbb{y}$ analytisch in λ.

9.9 Lineare DGL: Existenz und Eindeutigkeit der Lösung

Sei E ein endlichdimensionaler Vektorraum mit $\dim E = n$. Damit sei der metrische Raum $(\mathcal{L}(E,E), \|\cdot\|_{\mathcal{L}})$ gegeben. Ferner sei $I = [a,b] \subset \mathbb{R}$ und $A(\cdot)\colon [a,b] \to \mathcal{L}(E,E)$, das heißt $A(t) \in \mathcal{L}(E,E)$ mit $t \in I$. Weiter sei $f\colon [a,b] \xrightarrow{\text{stetig}} E$.
Sei zum Beispiel $E = \mathbb{R}^n$ und $e_j = (0,\dots,0,1,0,\dots,0)$ der j-te Einheitsvektor, sowie

$$A(t) = \begin{pmatrix} \alpha_{11}(t) & \cdots & \alpha_{1n}(t) \\ \vdots & & \vdots \\ \alpha_{n1}(t) & \cdots & \alpha_{nn}(t) \end{pmatrix}$$

Aufgabe: $A(\cdot)$ ist stetig in t genau dann, wenn für alle $k, l = 1, \dots, n$ gilt, dass $\alpha_{kl}(\cdot)\colon I \to \mathbb{R}$ stetig ist.
Es sei nun $f(t) = (f_1(t), \dots, f_n(t))$. Für alle $k = 1, \dots, n$ sei $f_k(\cdot)\colon I \to \mathbb{R}$ stetig. Betrachte $\mathbb{y}(t) = (y_1(t), \dots, y_n(t))$. Es sei $\dot{\mathbb{y}}(t) = A(t)\mathbb{y}(t) + f(t)$, was die Kurzschreibweise für das folgende System ist:

$$\begin{cases} \dot{y}_1(t) = \alpha_{11}(t)y_1(t) + \ldots + \alpha_{1n}(t)y_n(t) + f_1(t) \\ \qquad \vdots \\ \dot{y}_n(t) = \alpha_{n1}(t)y_1(t) + \ldots + \alpha_{nn}(t)y_n(t) + f_n(t) \end{cases}$$

Man nennt dies eine lineare, nicht autonome und inhomogene Differentialgleichung. Sie ist nicht autonom, da sowohl A als auch f von der Zeit t abhängen. Sie ist inhomogen, weil im Allgemeinen $f(t) \not\equiv 0$.

Wir können nun schreiben $\dot{\mathbb{y}}(t) = \mathbb{v}(t, \mathbb{y}(t))$ mit $\mathbb{v}(t, \mathbb{x}) = A(t)\mathbb{x} + f(t)$.

- $\mathbb{v}(t, \mathbb{x})$ ist stetig in $(t, \mathbb{x})$.
- $\|\mathbb{v}(t, \mathbb{x})\| \leq C$ für $(t, \mathbb{x}) \in I \times U_R(0)$, denn $\|A(t)\|_{\mathcal{L}} \leq C_1$ und $\|f(t)\|_E \leq C_2$ für alle $t \in I$. Damit folgt dann

$$\begin{aligned} \|\mathbb{v}(t, \mathbb{x})\|_E &\leq \|A(t)\mathbb{x}\|_E + \|f(t)\|_E \leq \|A(t)\|_{\mathcal{L}} \|\mathbb{x}\|_E + \|f(t)\|_E \\ &\leq C_1 R + C_2. \end{aligned}$$

- $\|\mathbb{v}(t, \mathbb{x}') - \mathbb{v}(t, \mathbb{x}'')\|_E \leq \|A(t)(\mathbb{x}' - \mathbb{x}'')\|_E \leq C_1 \|\mathbb{x}' - \mathbb{x}''\|_E$, also gilt Lipschitz-Stetigkeit.

Mit dem Satz von Picard-Lindelöf folgt damit zum Einen lokal die Existenz und Eindeutigkeit der Lösung, zum Anderen die globale Fortsetzbarkeit bis an den Rand des Phasenraums.

Frage: Für $M = \mathbb{R}^n$ scheint dies zunächst problemfrei zu sein, da es „keinen Rand gibt“. Die Frage ist nun allerdings, ob die Lösung in endlicher Zeit „explodieren“ kann, das heißt, ob sie in endlicher Zeit bereits ins Unendliche „verschwindet“, denn dann lässt sich die Differentialgleichung nur bis zu diesem Zeitpunkt lösen.

Satz 9.22

Das Cauchyproblem für die homogene lineare Differentialgleichung $\dot{\mathbb{y}}(t) = A(t)\mathbb{y}(t)$ mit $\mathbb{y}(t_0) = \mathbb{y}_0$ ($t_0 \in [a, b]$) besitzt eine eindeutige Lösung für alle $t \in [a, b]$.

Beweis. $\mathbb{y}(t)$ sei eine Lösung vom Cauchyproblem des Satzes. Falls es ein $t_1 \in I$ gibt, so dass $\mathbb{y}(t_1) = 0$ ist, so folgt $\mathbb{y}(t) \equiv 0$ für alle $t \in I$, denn dass die Nulllösung dann das Problem löst sieht man sofort und Picard-Lindelöf sagt uns wegen der Eindeutigkeit, dass dies dann wirklich die einzige Lösung ist. Für diese Lösung ist aber klar, dass der Satz gilt. Es sei nun also $\mathbb{y}(t) \neq 0$ für ein (und damit alle) $t \in I$. Die Funktion $g(t) = \ln \|\mathbb{y}(t)\|^2 = \ln\langle \mathbb{y}(t), \mathbb{y}(t)\rangle$ ist stetig differenzierbar. Es gilt

$$\frac{\mathrm{d}g(t)}{\mathrm{d}t} = \frac{1}{\|\mathbb{y}(t)\|^2} \frac{\mathrm{d}}{\mathrm{d}t} \langle \mathbb{y}(t), \mathbb{y}(t)\rangle = 2\frac{\langle \dot{\mathbb{y}}(t), \mathbb{y}(t)\rangle}{\|\mathbb{y}(t)\|^2}$$

und damit dann

$$\left|\frac{\mathrm{d}g(t)}{\mathrm{d}t}\right| \leq 2\frac{\|\dot{\mathbb{y}}(t)\|}{\|\mathbb{y}(t)\|} \leq 2\frac{\|A(t)\mathbb{y}(t)\|}{\|\mathbb{y}(t)\|} \leq 2\frac{C_1 \|\mathbb{y}(t)\|}{\|\mathbb{y}(t)\|} = 2C_1.$$

Außerdem gilt $g(t_0) = \ln \|\mathbb{y}_0\|^2$. Damit folgt

$$g(t) = g(t_0) + \int_{t_0}^{t} \frac{\mathrm{d}g(\tau)}{\mathrm{d}\tau} \,\mathrm{d}\tau$$

und damit

$$\begin{aligned}\underbrace{\cancel{|}g(t)\cancel{|}}_{\ln\|y(t)\|^2} &\leq \cancel{|}g(t_0)\cancel{|} + \int_{t_0}^{t} \underbrace{\left|\frac{\mathrm{d}g(\tau)}{\mathrm{d}\tau}\right|}_{\leq 2C_1} \,\mathrm{d}\tau \\ &\leq \underbrace{\cancel{|}g(t_0)\cancel{|}}_{\ln\|y_0\|^2} + 2C_1|b-a| \qquad \text{mit } t, t_0 \in I,\end{aligned}$$

wobei wir die Betragsstriche weglassen können. Damit ist

$$\|y(t)\|^{\cancel{2}} \leq \|y_0\|^{\cancel{2}}\, e^{\cancel{2}C_1|b-a|},$$

das heißt für $R = \|y_0\|\, e^{C_1|b-a|}$ ist $y(t)$ auf $I \times U_R(0)$ fortsetzbar. □

9.10 Struktur der Lösungen der homogenen Gleichung

Es seien $y^{(1)}(t)$ und $y^{(2)}(t)$ Lösungen von $\dot{y}(t) = A(t)y(t)$ (ohne Festlegen von Anfangsbedingungen). Dann ist auch jede Linearkombination $y(t) = \beta_1 y^{(1)}(t) + \beta_2 y^{(2)}(t)$ eine Lösung von $\dot{y}(t) = A(t)y(t)$, denn es ist

$$\begin{aligned}\dot{y}(t) &= \beta_1 \dot{y}^{(1)}(t) + \beta_2 \dot{y}^{(2)}(t) \\ &= \beta_1 A(t) y^{(1)}(t) + \beta_2 A(t) y^{(2)}(t) \\ &= A(t)\Big(\underbrace{\beta_1 y^{(1)}(t) + \beta_2 y^{(2)}(t)}_{y(t)}\Big).\end{aligned}$$

Dann folgt, dass die Menge $\mathcal{N}$ der Lösungen von $\dot{y}(t) = A(t)y(t)$ linear ist und es gilt $\mathcal{N} \subset C^1(I, \mathbb{R}^n)$.
Sei $B = \{b_1, \ldots, b_n\}$ eine Basis in E und $Y_B = \{y_{b_1}, \ldots, y_{b_n}\}$ seien die Lösungen der Cauchyprobleme $\dot{y}_{b_j}(t) = A(t) y_{b_j}(t)$ mit der Anfangsbedingung $y_{b_j}(t_0) = b_j$.

Satz 9.23
Es sei $A(\cdot) \in C(I, \mathcal{L})$. Dann folgt, dass Y_B eine Basis in $\mathcal{N}$ ist. Daraus folgt, dass $\dim \mathcal{N} = n = \dim E$.

Beweis.
1. z.Z.: Y_B ist linear unabhängig, das heißt für alle $t \in I$ gilt

$$\beta_1 y_{b_1}(t) + \ldots + \beta_n y_{b_n}(t) = 0 \;\Leftrightarrow\; \beta_1 = \ldots = \beta_n = 0.$$

Sei nun $\beta_1 y_{b_1}(t) + \ldots + \beta_n y_{b_n}(t) = 0$ für alle $t \in I$, dann gilt es insbesondere auch für $t = t_0$, also ist $\beta_1 b_1 + \ldots + \beta_n b_n = 0$. Da B eine Basis war folgt damit aber auch $\beta_1 = \ldots = \beta_n = 0$.

2. z.Z.: Y_B ist vollständig in $\mathcal{N}$. Sei $\mathrm{y}(t)$ eine Lösung von $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$, das heißt $\mathrm{y} \in \mathcal{N}$. Sei $\mathrm{y}(t_0) = b = \beta_1 b_1 + \ldots + \beta_n b_n \in E$. Betrachte $\tilde{\mathrm{y}}(t) = \beta_1 \mathrm{y}_{b_1}(t) + \ldots + \beta_n \mathrm{y}_{b_n}(t) \in \mathcal{N}$. Dann ist $\mathrm{y}(t) - \tilde{\mathrm{y}}(t) \in \mathcal{N}$. Außerdem gilt $\mathrm{y}(t_0) - \tilde{\mathrm{y}}(t_0) = 0$. Daraus folgt dann aber aus der Eindeutigkeit von Picard-Lindelöf, dass $\mathrm{y}(t) - \tilde{\mathrm{y}}(t) \equiv 0$ gilt. □

Wir wissen nun also, dass das homogene, nicht autonome Cauchyproblem einen n-dimensionalen Lösungsraum besitzt, in welchem man eine Basis bilden kann, indem man eine Basis von E als Anfangswerte betrachtet und die zugehörigen Cauchyprobleme löst.

Lemma 9.24
Das System von Vektoren $B^{(t_1)} = \{\mathrm{y}_{b_1}(t_1), \ldots, \mathrm{y}_{b_n}(t_n)\} \subset E$ bildet für alle $t_1 \in I$ eine Basis von E.

Beweis. Der Beweis wird zur Übung überlassen. □

Definition 9.25 Fundamentalsystem
Ein vollständiges, linear unabhängiges System von Lösungen der homogenen Gleichung $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$ nennt man ein Fundamentalsystem.

Bemerkung: $\mathcal{N} \supset Y = \{\mathrm{y}^{(1)}(t), \ldots, \mathrm{y}^{(n)}(t)\}$ ist ein Fundamentalsystem genau dann, wenn es ein t_1 gibt, so dass $\{\mathrm{y}^{(1)}(t_1), \ldots, \mathrm{y}^{(n)}(t_1)\}$ linear unabhängig ist.

9.11 Die Wronski-Determinante und die Formel von Liouville

Definition 9.26 Wronski-Determinante
Die Abbildungen $\varphi_1, \ldots, \varphi_n \colon I \to E = \mathbb{R}^n$ seien gegeben durch $\varphi_j(\tau) = (\varphi_j^1(\tau), \ldots, \varphi_j^n(\tau))^t$. Die Wronski-Determinante $W(\varphi_1, \ldots, \varphi_n)(\cdot) \colon I \to \mathbb{R}$ ist dann gegeben durch

$$W(\varphi_1, \ldots, \varphi_n)(\tau) = \det \begin{pmatrix} \varphi_1^1(\tau) & \cdots & \varphi_n^1(\tau) \\ \vdots & & \vdots \\ \varphi_1^n(\tau) & \cdots & \varphi_n^n(\tau) \end{pmatrix}.$$

Lemma 9.27
$Y = \{\mathrm{y}_1(t), \ldots, \mathrm{y}_n(t)\}$ ist ein Fundamentalsystem für $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$ mit $A(\cdot) \in C(I, \mathcal{L})$, $I = [a, b]$ und $\mathrm{y} \colon I \to \mathbb{R}^n$

$$\iff W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t) \neq 0 \text{ für alle } t \in I$$
$$\iff W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(\tilde{t}) \neq 0 \text{ für ein } \tilde{t} \in I.$$

Beweis. Der Lösungsraum für die homogene Gleichung hat die $\dim \mathcal{N} = n$. Damit ist Y eine Basis in $\mathcal{N}$ genau dann, wenn Y linear unabhängig ist. Dies ist wiederum genau dann der Fall, wenn $Y_{\tilde{t}}$ linear unabhängig in $\mathbb{R}^n$ für alle bzw. für ein $\tilde{t} \in I$ ist. Dies ist äquivalent dazu, dass $\det Y_{\tilde{t}} = W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(\tilde{t}) \neq 0$. □

Anmerkung: Es sei $A = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{n1} \\ \vdots & & \vdots \\ \alpha_{1n} & \cdots & \alpha_{nn} \end{pmatrix}$.
Dann ist die Spur $\mathrm{Sp}(A) := \sum_{k=1}^{n} \alpha_{kk}$ und es gilt $\mathrm{Sp}(BAC) = \mathrm{Sp}(ACB) = \mathrm{Sp}(CBA)$. Falls B invertierbar ist gilt insbesondere $\mathrm{Sp}(B^{-1}AB) = \mathrm{Sp}(ABB^{-1}) = \mathrm{Sp}(A)$, da $BB^{-1} = \mathbb{1}$, das heißt die Spur ist invariant bezüglich Ähnlichkeitstransformationen. Durch Zurückführen per Diagonalisierung oder auf Jordannormalform folgt daraus, dass die Spur gleich der Summe der Eigenwerte ist. Da Basiswechsel Ähnlichkeitstransformationen sind, ist die Spur auch invariant bezüglich eines Basiswechsels.

Satz 9.28 Formel von Liouville
$Y = \{\mathrm{y}_1(t), \ldots, \mathrm{y}_n(t)\} \subset \mathcal{N}$ sei ein System von Lösungen der o.g. homogenen Differentialgleichung. Dann gilt:

$$\frac{\mathrm{d}}{\mathrm{d}t} W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t) = (\mathrm{Sp}\, A(t)) \cdot W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t),$$

woraus sich die folgende Gleichung ergibt:

$$W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t) = W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t_0) e^{\int_{t_0}^{t} \mathrm{Sp}(A(\tau))\ \mathrm{d}\tau}.$$

Beweis.
Fall 1: Y ist linear abhängig $\Leftrightarrow W(t) \equiv 0$.
Fall 2: Y sei linear unabhängig. Wir vereinbaren, dass wir die Argumente von W nicht immer schreiben müssen. Dann ist

$$W(\mathrm{y}_1, \ldots, \mathrm{y}_n)(t) = \sum_{\pi} (-1)^{\mathrm{sign}\,\pi} y_1^{\pi(1)}(t) \cdot \ldots \cdot y_n^{\pi(n)}(t)$$

und dann

$$\dot{W} = \sum_{\pi} (-1)^{\mathrm{sign}\,\pi} \dot{y}_1^{\pi(1)}(t) y_2^{\pi(2)}(t) \cdot \ldots \cdot y_n^{\pi(n)}(t) + \ldots + \sum_{\pi} (-1)^{\mathrm{sign}\,\pi} y_1^{\pi(1)}(t) \cdot \ldots \cdot \dot{y}_n^{\pi(n)}(t),$$

das heißt die Ableitung wird summandenweise „durchgeschoben“. Damit ist dann

$$\dot{W} = \det(\dot{\mathrm{y}}_1(t), \mathrm{y}_2(t), \ldots, \mathrm{y}_n(t)) + \ldots + \det(\mathrm{y}_1(t), \ldots, \mathrm{y}_{n-1}(t), \dot{\mathrm{y}}_n(t)).$$

Wir fixieren jetzt einen Zeitpunkt $t = \tilde{t}$. Betrachte $Y_{\tilde{t}} = \{\mathrm{y}_1(\tilde{t}), \ldots, \mathrm{y}_n(\tilde{t})\}$ als Basis im $\mathbb{R}^n$ und stelle $A(t)$ für $t = \tilde{t}$ in dieser Basis dar:

$$A(\tilde{t}) \longleftrightarrow (\alpha_{ij}(\tilde{t}))$$
$$A(\tilde{t}) \mathrm{y}_k(\tilde{t}) = \sum_{l} \alpha_{kl}(\tilde{t}) \mathrm{y}_l(\tilde{t}) = \dot{\mathrm{y}}_k \Big|_{t=\tilde{t}}.$$

Damit ist dann

$$
\begin{aligned}
\dot{W}\Big|_{t=\tilde{t}} &= \det\left(\sum_l \alpha_{1l}(\tilde{t})\mathrm{y}_l(\tilde{t}), \mathrm{y}_2(\tilde{t}), \ldots, \mathrm{y}_n(\tilde{t})\right) \\
&\qquad + \ldots + \det\left(\mathrm{y}_1(\tilde{t}), \ldots, \mathrm{y}_{n-1}(\tilde{t}), \sum_l \alpha_{nl}(\tilde{t})\mathrm{y}_l(\tilde{t})\right) \\
&= \alpha_{11}(\tilde{t}) \det\big(\mathrm{y}_1(\tilde{t}), \ldots, \mathrm{y}_n(\tilde{t})\big) \\
&\qquad + \ldots + \alpha_{nn}(\tilde{t}) \det\big(\mathrm{y}_1(\tilde{t}), \ldots, \mathrm{y}_n(\tilde{t})\big) \\
&= \underbrace{\big(\alpha_{11}(\tilde{t}) + \ldots + \alpha_{nn}(\tilde{t})\big)}_{=\operatorname{Sp} A(\tilde{t})} \cdot W(\tilde{t}).
\end{aligned}
$$

Damit ist der Satz bewiesen. □

9.12 Der Evolutionsoperator

Wir betrachten wieder $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$ mit $A(\cdot)\colon I \overset{\text{stetig}}{\longrightarrow} \mathcal{L}(E,E)$ und $E \sim \mathbb{R}^n$. Ferner betrachten wir einen Operator $U(t_1,t_0)\colon E \to E$ mit $t_0, t_1 \in I$. Dabei soll $\mathrm{y}_1 = U(t_1,t_0)\mathrm{y}_0$ (für $\mathrm{y}_0, \mathrm{y}_1 \in E \sim \mathbb{R}^n$) genau dann gelten, wenn $\mathrm{y}_1 = \mathrm{y}(t_1)$, wobei $\mathrm{y}(t)$ das Cauchyproblem $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$ mit $\mathrm{y}(t_0) = \mathrm{y}_0$ löst.
Die Existenz und Eindeutigkeit des Cauchyproblems sichern uns, dass für jedes y_0 wirklich genau ein solches y_1 existiert. Damit ist die Abbildung wohldefiniert. Ferner ist $U(t_1,t_0)\colon E \to E$ eine lineare Abbildung. Betrachte dazu ein $\mathrm{y}_0 = \beta_I \mathrm{y}_0^I + \beta_{II}\mathrm{y}_0^{II}$. Angenommen, $\mathrm{y}^I(t)$ und $\mathrm{y}^{II}(t)$ lösen das Cauchyproblem mit $\mathrm{y}^I(t_0) = \mathrm{y}_0^I$ und $\mathrm{y}^{II}(t_0) = \mathrm{y}_0^{II}$. Dann folgt, dass $\mathrm{y}(t) = \beta_I \mathrm{y}^I(t) + \beta_{II}\mathrm{y}^{II}(t)$ ebenfalls das Cauchyproblem löst, allerdings mit $\mathrm{y}(t_0) = \beta_I \mathrm{y}_0^I + \beta_{II}\mathrm{y}_0^{II}$. Damit folgt

$$
\begin{aligned}
U(t_1,t_0)(\beta_I \mathrm{y}_0^I + \beta_{II}\mathrm{y}_0^{II}) &= \mathrm{y}(t_1) = \beta_I \mathrm{y}^I(t_1) + \beta_{II}\mathrm{y}^{II}(t_1) \\
&= \beta_I U(t_1,t_0)\mathrm{y}_0^I + \beta_{II} U(t_1,t_0)\mathrm{y}_0^{II}.
\end{aligned}
$$

Betrachte nun $E \sim \mathbb{R}^n$ und den k-ten Einheitsvektor $e_k = (0,\ldots,0,1,0,\ldots,0)$. Dann ist

$$
\begin{aligned}
\mathrm{y}_0 &= \sum_k \langle \mathrm{y}_0, e_k\rangle e_k = (y_0^1, \ldots, y_0^n)^t \\
\mathrm{y}(t) &= \sum_k \langle \mathrm{y}_0, e_k\rangle \mathrm{y}_k(t),
\end{aligned}
$$

wobei die $\mathrm{y}_k(t)$ Lösungen des Cauchyproblems $\dot{\mathrm{y}}_k(t) = A(t)\mathrm{y}_k(t)$ mit $\mathrm{y}_k(t_0) = e_k$ seien. Dies löst dann ebenfalls wieder das Cauchyproblem mit $\mathrm{y}(t_0) = \mathrm{y}_0$. Also ist

$$
\mathrm{y}_1 = \mathrm{y}(t_1) = \sum_k \langle \mathrm{y}_0, e_k\rangle \mathrm{y}_k(t_1).
$$

Damit folgt dann

$$U(t_1, t_0) = \begin{pmatrix} y_1^1(t_1) & \cdots & y_n^1(t_1) \\ \vdots & & \vdots \\ y_1^n(t_1) & \cdots & y_n^n(t_1) \end{pmatrix}.$$

Auf der linken Seite taucht ein t_0 auf, nicht jedoch auf der rechten Seite. Wo ist dieses t_0 also „hinverschwunden“? Es ist $\mathrm{y}_k(t) = \mathrm{y}_k(t; t_0)$ abhängig vom Zeitpunkt der Anfangsbedingung. Also hängt die rechte Seite sehr wohl von t_0 ab.
Es gilt

$$\det U(t_1, t_0) = W(\mathrm{y}_1, \ldots, \mathrm{y}_k)(t_1),$$

wobei alle $\mathrm{y}_1, \ldots, \mathrm{y}_k$ von t_0 abhängen.
Eigenschaften:

1. $U(t, t) = \mathbb{1}$.
2. $U(t_1, t_0) = U(t_1, t)U(t, t_0)$.
3. $\frac{\mathrm{d}}{\mathrm{d}t}U(t, t_0) = (\dot{\mathrm{y}}_1, \ldots, \dot{\mathrm{y}}_n) = (A(t)\mathrm{y}_1(t), \ldots, A(t)\mathrm{y}_n(t))$ $= A(t)U(t, t_0)$.
4. $U(t, t_0)\mathrm{y}_0 = 0 \;\Leftrightarrow\; \mathrm{y}_0 = 0$, denn $\det U(t, t_0) = W(t) \neq 0$, da $W(t_0) = \det \mathbb{1} \neq 0$ (Injektivität).
5. $U(t, t_0)E = E$ (Surjektivität).
6. $U(t_0, t)U(t, t_0) = U(t_0, t_0) = \mathbb{1}$ (Spezialfall von **2.**).

Satz 9.29
Das (inhomogene) Cauchyproblem $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t) + f(t)$ mit $\mathrm{y}(t_0) = \mathrm{y}_0$ und $A(t)$, $f(t)$ stetig für alle $t \in I$ besitzt für alle $t \in I$ die Lösung

$$\mathrm{y}(t) = U(t, t_0)\mathrm{y}_0 + \int_{t_0}^{t} U(t, \tau) f(\tau)\, \mathrm{d}\tau.$$

Die Lösung des Cauchyproblems ist eindeutig.

Beweis. Es ist allgemein

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{t_0}^{t} \varphi(t, \tau)\, \mathrm{d}\tau = 1 \cdot \varphi(t, t) + \int_{t_0}^{t} \frac{\partial \varphi(t, \tau)}{\partial t}\, \mathrm{d}\tau.$$

Damit folgt aus der Formel für die Lösung im Satz

$$\begin{aligned}\dot{\mathrm{y}}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(U(t,t_0)\mathrm{y}_0\right) + U(t,t)f(t) + \int_{t_0}^{t} \frac{\partial}{\partial t} U(t,\tau)f(\tau)\ \mathrm{d}\tau \\ &= A(t)U(t,t_0)\mathrm{y}_0 + f(t) + \int_{t_0}^{t} A(t)U(t,\tau)f(\tau)\ \mathrm{d}\tau \\ &= A(t)\Bigg[\underbrace{U(t,t_0)\mathrm{y}_0 + \int_{t_0}^{t} U(t,\tau)f(\tau)\ \mathrm{d}\tau}_{\mathrm{y}(t)}\Bigg] + f(t),\end{aligned}$$

das heißt $\mathrm{y}(t)$ erfüllt $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t) + f(t)$. Für $t = t_0$ ist $\int_{t_0}^{t_0} \cdots\ \mathrm{d}\tau = 0$ und $U(t_0,t_0)\mathrm{y}_0 = \mathbb{1}\mathrm{y}_0 = \mathrm{y}_0$, also insgesamt $\mathrm{y}(t_0) = \mathrm{y}_0$. □

Anmerkung zur Eindeutigkeit: Seien $\mathrm{y}^{(I)}$ und $\mathrm{y}^{(II)}$ Lösungen des inhomogenen Cauchyproblems (aus dem Satz) mit den Anfangswerten $\mathrm{y}^{(I)}(t_0) = \mathrm{y}^{(II)}(t_0) = \mathrm{y}_0$. $\mathrm{y}^{(I)}(t) - \mathrm{y}^{(II)}(t) = \mathrm{y}(t)$ löst das homogene Problem $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t)$ mit $\mathrm{y}(t_0) = 0$. Die Lösung dieses Cauchyproblems ist aber gerade $\mathrm{y}(t) \equiv 0$.

Allgemeine Struktur der Lösungen von $\dot{\mathrm{y}} = A(t)\mathrm{y} + f(t)$

Satz 9.30
Es seien $A(\cdot) \in C(I,\mathcal{L})$ und $f \in C(I,E)$. Die allgemeine Lösung der o.g. Differentialgleichung $\dot{\mathrm{y}} = A(t)\mathrm{y} + f(t)$ ist die Summe einer Partiallösung y_p der inhomogenen Gleichung und der allgemeinen Lösung $\mathrm{y}_h \in \mathcal{N}$ der homogenen Gleichung.

Lösungsweg für nicht-autonome inhomogene Systeme
Gegeben sei $\dot{\mathrm{y}}(t) = A(t)\mathrm{y}(t) + f(t)$ mit $A(\cdot) \in C(I,\mathcal{L})$ und $f \in C(I,E)$.

1. Finde eine Lösung $\mathrm{y}_n \in \mathcal{N}$ des homogenen Systems.
2. Konstruiere $U(t,t_0)$ aus den Lösungen y_k mit $\mathcal{N} \ni \mathrm{y}_k(t_0) = e_k$.
3. Berechne

$$\mathrm{y}_p = U(t,t_0)\mathrm{y}_0 + \int_{t_0}^{t} U(t,\tau)f(\tau)\ \mathrm{d}\tau$$

9.13 Lineare autonome Systeme

Nun sei $A(t) \equiv A \in \mathcal{L}(E,E)$ unabhängig von t.

Spezialfall: $A = \alpha \cdot \mathbb{1}$ mit $\alpha \in \mathbb{C}$. Wir haben also das System $\dot{\mathrm{y}}(t) = A\mathrm{y}(t)$ mit $\mathrm{y}(t_0) = \mathrm{y}_0$. Dann folgt $\mathrm{y}(t) = e^{\alpha(t-t_0)}\mathrm{y}_0$.

Lösungsidee im allgemeinen Fall: Es sei $A \in \mathcal{L}(E,E)$ beliebig. Für das Cauchyproblem wie oben könnte man dann $\mathrm{y}(t) = e^{(t-t_0)A}\mathrm{y}_0$ schreiben. Dazu muss man den

Ausdruck „e^{Matrix}" aber sinnvoll definieren. Dies geschieht über

$$e^A = \exp(A) := \sum_{k=0}^{\infty} \frac{A^k}{k!}.$$

1. Diese Reihe konvergiert im Raum $(\mathcal{L}(E,E), \|\cdot\|_{\mathcal{L}})$ absolut, denn es ist

$$\|A\|_{\mathcal{L}} = \sup_{x\neq 0} \frac{\|Ax\|_E}{\|x\|_E},$$
$$\|AB\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}} \|B\|_{\mathcal{L}}.$$

   Damit ist dann $\left\|A^k\right\|_{\mathcal{L}} \leq \|A\|_{\mathcal{L}}^k$ und es folgt

$$\sum_{k=0}^{\infty} \left\| \frac{A^k}{k!} \right\|_{\mathcal{L}} \leq \sum_{k=0}^{\infty} \frac{\|A\|_{\mathcal{L}}^k}{k!} = e^{\|A\|_{\mathcal{L}}}.$$

2. Multiplikationseigenschaft: „$e^{A+B} = e^A \cdot e^B$" ist im Allgemeinen falsch! Sie gilt jedoch, falls A und B kommutieren, das heißt, wenn $AB = BA$ gilt. Es ist

$$\begin{aligned} e^A e^B &= \left(\sum_{k=0}^{\infty} \frac{A^k}{k!} \right) \left(\sum_{n=0}^{\infty} \frac{B^n}{n!} \right) = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{A^k B^n}{k! n!} \\ &\overset{n=p-k}{=} \sum_{p=0}^{\infty} \frac{1}{p!} \underbrace{\sum_{k=0}^{p} \frac{p!}{k!(p-k)!} A^k B^{p-k}}_{=(A+B)^p} = e^{A+B}, \end{aligned}$$

   wobei die angewendete binomische Formel im Matrizenfall eben nur für kommutierende Matrizen A, B gilt.

3. Die Funktion $t \mapsto \exp(t \cdot A) = e^{tA}$ mit $t \in \mathbb{R}$ ist eine Abbildung von $\mathbb{R}$ nach $\mathcal{L}(E,E)$ und es gilt

$$\frac{\mathrm{d}}{\mathrm{d}t} e^{tA} = Ae^{tA}.$$

   Betrachte dazu

$$\lim_{\Delta t \to 0} \frac{e^{\Delta t A} - \mathbb{1}}{\Delta t} = \lim_{\Delta t \to 0} \frac{\Delta t A + o(\Delta t)}{\Delta t} = A.$$

   Damit ist dann

$$\begin{aligned} Ae^{tA} &= \lim_{\Delta t \to 0} \left(\frac{e^{\Delta t A} - \mathbb{1}}{\Delta t} \right) e^{tA} = \lim_{\Delta t \to 0} \left(\frac{e^{\Delta t A} e^{tA} - e^{tA}}{\Delta t} \right) \\ &= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \left(e^{(t+\Delta t)A} - e^{tA} \right), \end{aligned}$$

   wobei im letzten Schritt ausgenutzt wurde, dass $\Delta t A$ und tA kommutieren.

4. Es sei $B \in \mathcal{L}(E, E)$ und B invertierbar. Dann gilt

$$e^A = Be^{B^{-1}AB}B^{-1}.$$

Betrachte dazu $(B^{-1}AB)^n = (B^{-1}A\cancel{B})(\cancel{B^{-1}}A\cancel{B})\dots(\cancel{B^{-1}}AB) = B^{-1}AB$. Damit ist

$$\begin{aligned} e^{B^{-1}AB} &= \sum_{n=0}^{\infty} \frac{(B^{-1}AB)^n}{n!} = \sum_{n=0}^{\infty} B^{-1}\frac{A^n}{n!}B \\ &= B^{-1}\left(\sum_{n=0}^{\infty} \frac{A^n}{n!}\right) B = B^{-1}e^A B. \end{aligned}$$

Sei nun $\mathrm{y}(t) = e^{(t-t_0)A}\mathrm{y}_0$. Dann ist

$$\begin{aligned} \dot{\mathrm{y}}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{(t-t_0)A}\right)\mathrm{y}_0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(e^{tA}\right)e^{-t_0A}\mathrm{y}_0 = Ae^{tA}e^{-t_0A}\mathrm{y}_0 = Ae^{(t-t_0)A}\mathrm{y}_0 \\ &= A\mathrm{y}(t). \end{aligned}$$

Für $t = t_0$ folgt $e^{(t-t-0)A} = e^0 = \mathbb{1}$. Also ist $\mathrm{y}(t_0) = \mathbb{1}\mathrm{y}_0 = \mathrm{y}_0$ und die Anfangsbedingung damit erfüllt.

Frage: Wie berechnet man $e^{(t-t_0)A}$? Für den Spezialfall, dass A diagonalisierbar ist, gibt es eine Matrix B mit $B^{-1}AB = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$. Damit ist dann

$$\begin{aligned} e^{\operatorname{diag}\{\lambda_1,\dots,\lambda_n\}} &= \sum_{k=0}^{\infty} \frac{\operatorname{diag}\{\lambda_1,\dots,\lambda_n\}^k}{k!} = \sum_{k=0}^{\infty} \frac{\operatorname{diag}\{\lambda_1^n,\dots,\lambda_n^n\}}{k!} \\ &= \operatorname{diag}\left\{\sum_{k=0}^{\infty}\frac{\lambda_1^k}{k!},\dots,\sum_{k=0}^{\infty}\frac{\lambda_n^k}{k!}\right\} \\ &= \operatorname{diag}\left\{e^{\lambda_1},\dots,e^{\lambda_n}\right\}. \end{aligned}$$

Damit ist dann also

$$e^{(t-t_0)A} = B \operatorname{diag}\left\{e^{(t-t_0)\lambda_1},\dots,e^{(t-t_0)\lambda_n}\right\} B^{-1}.$$

Allgemeiner Fall: Was passiert nun, wenn A nicht notwendigerweise diagonalisierbar ist? Dann lässt sich A jedoch auf Jordannormalform bringen, also existiert eine invertierbare Matrix B mit $B^{-1}AB = J$, wobei J eine Jordanmatrix mit den Jordanblöcken $J_\nu(\lambda)$ ist, für die gilt:

$$J_\nu(\lambda) = \underbrace{\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}}_{\nu}$$

Es ist $J_\nu(\lambda) = \lambda \cdot \mathbb{1}_\nu + T_\nu$, wobei T_ν auf der ersten Nebendiagonalen Einsen stehen hat:

$$T_\nu = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$$

Wichtig: $\lambda \cdot \mathbb{1}_\nu$ und T_ν kommutieren, also gilt $e^{\lambda \mathbb{1}_\nu + T_\nu} = e^\lambda e^{T_\nu}$. Ferner verschiebt sich die Diagonale bei den Potenzen T_ν^k immer um eins nach oben, also hat $T_\nu^{\nu-1}$ nur noch genau eine Eins im oberen rechten Eintrag. Es ist $T_\nu^\nu = T_\nu^m = 0$ für alle $m \geq \nu$. Ferner ist

$$e^{T_\nu} = \sum_{k=0}^{\infty} \frac{T_\nu^k}{k!} = \sum_{k=0}^{\nu-1} \frac{T_\nu^k}{k!} = \begin{pmatrix} 1 & 1 & \frac{1}{2} & \cdots & \frac{1}{(\nu-1)!} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{1}{2} \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}$$

Im Allgemeinen ist

$$e^{(t-t_0)J_\nu(\lambda)} = e^{(t-t_0)\lambda} e^{(t-t_0)T_\nu} = \begin{pmatrix} 1 & (t-t_0) & \cdots & \frac{(t-t_0)^{\nu-1}}{(\nu-1)!} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & (t-t_0) \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

$$\mathrm{y}(t) = \underbrace{U(t,t_0)\mathrm{y}_0}_{\tilde{\mathrm{y}}(t)} + \underbrace{\int_{t_0}^t U(t,\tau) f(\tau)\ \mathrm{d}\tau}_{\hat{\mathrm{y}}(t)}$$

$$\underbrace{\begin{cases} \dot{\tilde{\mathrm{y}}} = A(t)\tilde{\mathrm{y}} \\ \tilde{\mathrm{y}}(t_0) = \mathrm{y}_0 \end{cases}}_{\substack{\text{homogen} \\ \text{Anfangsbedingung } \mathrm{y}_0}} \qquad \underbrace{\begin{cases} \dot{\hat{\mathrm{y}}} = A(t)\hat{\mathrm{y}} + f \\ \hat{\mathrm{y}}(t_0) = 0 \end{cases}}_{\substack{\text{inhomogen} \\ \text{Anfangsbedingung } 0}}$$

Falls $A(t) \equiv A$ ist gilt $U(t,t_0) = e^{(t-t_0)A}$. Die Funktion

$$\mathrm{y}(t) = e^{(t-t_0)A}\mathrm{y}_0 + \int_{t_0}^t e^{(t-\tau)A} f(\tau)\ \mathrm{d}\tau$$

löst also das Cauchyproblem

$$\begin{cases} \dot{\mathrm{y}}(t) = A\mathrm{y}(t) + f(t) \\ \mathrm{y}(t_0) = \mathrm{y}_0 \end{cases}$$

9.14 Lineare Differentialgleichungen höherer Ordnung

Es sei $a_j \in C([a,b],\mathbb{R})$ für $j = 0,\dots,n-1$. Gegeben sei die lineare inhomogene Differentialgleichung der Ordnung n:

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_0(t)y(t) = f(t) \qquad \textbf{(*)}$$

Es sei $\mathbb{y}(t) = (y^1(t),\dots,y^n(t))^t = (y(t), y^{(1)}(t),\dots,y^{(n-1)}(t))^t$, das heißt $y^j(t) = y^{(j-1)}(t)$. Es gelten die Gleichungen

$$y^k(t) = y^{(k-1)}(t) \qquad \text{für } k = 2,\dots,n.$$

und

$$\begin{aligned} \dot{y}^n(t) &= \dot{y}^{(n-1)}(t) = y^{(n)}(t) \\ &= -a_0(t)y^1(t) - \ldots - a_{n-1}(t)y^n(t) + f(t). \end{aligned}$$

Also ist $\dot{\mathbb{y}}(t) = A(t)\mathbb{y}(t) + \mathbb{f}(t)$ (**(**)**) mit

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -a_0(t) & -a_1(t) & \cdots & -a_{n-1}(t) \end{pmatrix} \qquad \textbf{(***)}$$

und $\mathbb{f}(t) = (0,\dots,0,f(t))^t$. Also ist **(*)** $\Leftrightarrow$ **(**)** und wir können die gesamte bisherige Theorie hier wiederverwenden.

Was ist ein Cauchyproblem für **(**)** $\Rightarrow$ Cauchyproblem für **(*)**?

Ein Cauchyproblem für **(**)** ist $\dot{\mathbb{y}}(t) = A(t)\mathbb{y}(t) + \mathbb{f}(t)$ mit $\mathbb{y}(t_0) = \mathbb{c}$. Dies ist äquivalent zu einem Cauchyproblem für **(*)** mit $y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0y = f$ mit $y^k(t_0) = c_{k+1}$ für $k = 0,\dots,n-1$. Wichtig ist, dass all diese Gleichungen für einen gemeinsamen Zeitpunkt t_0 gelten.

Das Cauchyproblem für **(*)** besitzt für alle $t \in I = [a,b]$ eine eindeutige Lösung.

$$\text{Lösungsstruktur: } y(t) = \underbrace{y_h(t)}_{\text{allg. Lsg. d. hom. Gl.}} + \underbrace{y_p(t)}_{\text{part. Lsg. d. inh. Gl.}}$$

Die autonome Differentialgleichung höherer Ordnung:

$f(t)$ bleibt zeitabhängig, die Koeffizienten $a_j(t) \equiv a_j$ seien nun aber unabhängig von der Zeit. Die Differentialgleichung lautet also

$$y^{(n)}(t) + a_{n-1}y^{(n-1)}(t) + \ldots + a_0y(t) = f(t).$$

Wir definieren das charakteristische Polynom mit

$$P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0.$$

Es gibt aber auch noch das charakteristische Polynom der Matrix: $d_A(x) = \det(A - \lambda\mathbb{1})$, wobei $A(t) = A$ zeitunabhängig und wie weiter oben durch **(***)** gegeben ist.

Satz 9.31
Es gilt $d_A(\lambda) = (-1)^n P(\lambda)$.

Beweis. Es ist

$$d_A(\lambda) = \begin{pmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\lambda & 1 \\ -a_0 & -a_1 & \cdots & \cdots & -a_{n-1} - \lambda \end{pmatrix}$$

Entwickelt man dies nach der letzten Zeile, so erhält man

$$d_A(\lambda) = (-1)^{n+1}(-a_0)\det M_{n,1}(A) + \ldots + (-1)^{n+n}(-a_{n-1} - \lambda)\det M_{n,n}(A).$$

Man sieht nun leicht, dass $\det M_{n,k}(A) = (-\lambda)^{k-1}$ ist. Setzt man dies nun ein so erhält man

$$\begin{aligned} &= (-1)^n(a_0 + a_1\lambda + \ldots + \lambda^n) \\ &= (-1)^n P(\lambda). \end{aligned}$$

Damit ist der Beweis vollständig. □

Es seien $\lambda_1, \ldots, \lambda_k$ Nullstellen von $P(\lambda)$ (und damit auch von $d_A(\lambda)$) der Ordnung $\nu_1, \ldots, \nu_k$. Jede Nullstelle λ_j erzeugt dann ein Jordankästchen $J_{\nu_j}(\lambda_j)$ für A.

Satz 9.32
Sei λ_j eine Nullstelle der Ordnung ν_j von $P(\lambda)$, dann löst jede Linearkombination von $Y_j = \{e^{\lambda_j t}, te^{\lambda_j t}, \ldots, t^{\nu_j - 1}e^{\lambda_j t}\}$ die homogene autonome Differentialgleichung $y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y = 0$.
Die allgemeine Lösung des homogenen Problems ist der Raum aufgespannt durch Y_j $(j = 1, \ldots, k)$.

Beweis. Es ist

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y = \left(\frac{\mathrm{d}}{\mathrm{d}t} - \lambda_1\right)^{\nu_1} \cdot \ldots \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} - \lambda_k\right)^{\nu_k} y.$$

Daraus folgt die Behauptung. □

9.15 Die Laplace-Transformation

Betrachte $f : [0, +\infty[\to \mathbb{C}$. Die Laplace-Transformation $\mathcal{L}[f](p) = \tilde{f}(p)$ ist dann gegeben durch

$$\mathcal{L}[f](p) = \tilde{f}(p) = \int_0^{+\infty} f(t)e^{-tp}\,\mathrm{d}t, \qquad p \in \mathbb{C}.$$

Beispiel:

- $f(t) = 1 \;\Rightarrow\; \tilde{f}(p) = \frac{1}{p}$ für $\operatorname{Re} p > 0$.
- $f(t) = t^n \;\Rightarrow\; \tilde{f}(p) = \frac{n!}{p^{n+1}}$ für $\operatorname{Re} p > 0$ und $n \in \mathbb{N}$.
- $f(t) = e^{-ta} \;\Rightarrow\; \tilde{f}(p) = \frac{1}{p+a}$ für $\operatorname{Re}(p+a) > 0$.
- $f(t) = \cos \omega t \;\Rightarrow\; \tilde{f}(p) = \frac{p}{p^2+\omega^2}$ für $\operatorname{Re} p > |\operatorname{Im} \omega|$.
- $f(t) = \sin \omega t \;\Rightarrow\; \tilde{f}(p) = \frac{\omega}{p^2+\omega^2}$ für $\operatorname{Re} p > |\operatorname{Im} \omega|$.

Für die Laplace-Transformation gilt Linearität, das heißt

$$\mathcal{L}[\alpha f + \beta g](p) = \alpha \mathcal{L}[f](p) + \beta \mathcal{L}[g](p).$$

Ferner gilt

$$\mathcal{L}\left[f\left(\frac{t}{a}\right)\right](p) = a\mathcal{L}[f](pa), \qquad a > 0$$

und

$$\mathcal{L}[e^{-at}f(t)](p) = \mathcal{L}[f](p+a), \qquad a \in \mathbb{C}.$$

Die Heavysidefunktion lautet

$$H(t) = \chi_{[0,+\infty[}(t) = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}.$$

Verschiebt man ein Signal $f(t)$ um a, so gleicht man das „Unwissen“ über die Stellen links von a durch die Heavysidefunktion aus, man verschiebt also zu $f(t-a)H(t-a)$. Für die Laplace-Transformation gilt dann

$$\mathcal{L}[f(t-a)H(t-a)](p) = e^{-ap}\mathcal{L}[f](p), \qquad a \geq 0.$$

Laplace-Transformation und Ableitung

Es sei $f \in C^{(n)}([0,+\infty[)$ und die Transformationen $\mathcal{L}[f^{(k)}](p)$ existieren für alle $k = 0, \ldots, n$. Außerdem nehmen wir an, dass $f^{(k)}(t)e^{-tp} \xrightarrow{t\to\infty} 0$ gilt. Dann gilt

$$\mathcal{L}[f^{(n)}](p) = p^n \mathcal{L}[f](p) - p^{n-1}f(0) - p^{n-2}f^{(1)}(0) - \ldots - p^0 f^{(n-1)}(0).$$

Beweis. Der Beweis erfolgt durch partielles Integrieren:

$$\begin{aligned} \mathcal{L}[f^{(n)}](p) &= \int_0^{+\infty} f^{(n)}(t)e^{-pt} \,\mathrm{d}t \\ &= f^{(n-1)}(t)e^{-pt}\Big|_0^{+\infty} + p\int_0^{+\infty} f^{(n-1)}(t)e^{-pt} \,\mathrm{d}t \\ &= -f^{(n-1)}(0) \cdot 1 + p\mathcal{L}[f^{(n-1)}](p). \end{aligned}$$

Dies führt man dann immer weiter. □

Nun existiere $\mathcal{L}[f](p)$ für $\operatorname{Re} p > c$. Dann folgt, dass $\mathcal{L}[f](p)$ in allen Punkten p mit $\operatorname{Re} p > c$ analytisch ist. Ferner gilt

$$\mathcal{L}[t^n f(t)](p) = (-1)^n \frac{\mathrm{d}^n}{\mathrm{d}p^n} \mathcal{L}[f](p).$$

Definition 9.33 Faltung
Es seien $f, g\colon \mathbb{R} \to \mathbb{C}$. Dann ist die Faltung definiert durch

$$(f * g)(t) = \int_{\mathbb{R}} f(\tau) g(t-\tau) \ \mathrm{d}\tau.$$

Bei uns setzen wir $f, g\colon [0, +\infty[\to \mathbb{C}$ mit 0 fort. Damit verschwindet $f(\tau)$ für $\tau < 0$ und $g(t-\tau)$ für $\tau > t$. Also lautet die Faltung in diesem Fall

$$(f * g)(t) = \int_0^t f(\tau) g(t-\tau) \ \mathrm{d}\tau.$$

Angenommen, die Transformationen $\mathcal{L}[f](p)$ und $\mathcal{L}[g](p)$ existieren für $\operatorname{Re} p > c$. Dann gilt, dass $\tilde{h}(p) = \tilde{f}(p) \cdot \tilde{g}(p)$ die Laplace-Transformation von $h = f * g$ ist. Die Laplace-Transformation überführt eine Faltung also in ein Produkt.
Beweisskizze: Es ist

$$\begin{aligned}
\mathcal{L}[f](p) \cdot \mathcal{L}[g](p) &= \left(\int_0^{+\infty} e^{-tp} f(t) \ \mathrm{d}t \right) \left(\int_0^{+\infty} e^{-sp} g(s) \ \mathrm{d}s \right) \\
&= \int_0^{+\infty} \int_0^{+\infty} \underbrace{e^{-pt} e^{-sp}}_{e^{-p(t+s)}} f(t) g(s) \ \mathrm{d}s \ \mathrm{d}t \\
&\overset{t+s=y}{=} \int_0^{+\infty} \left(\int_t^{+\infty} f(t) g(y-t) e^{-py} \ \mathrm{d}y \right) \ \mathrm{d}t \\
&= \int_0^{+\infty} \left(\int_0^y f(t) g(y-t) \ \mathrm{d}t \right) e^{-py} \ \mathrm{d}y \\
&= \int_0^{+\infty} (f * g)(y) e^{-py} \ \mathrm{d}y \\
&= \mathcal{L}[f * g](p).
\end{aligned}$$

Laplace-Transformation und Lineare Differentialgleichung
Gegeben sei das System

$$\begin{cases} y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_0 y = f \\ y(0) = 0, \ldots, y^{(n-1)}(0) = 0 \end{cases}$$

mit *konstanten* Koeffizienten $a_k(t) = a_k$.

$$\begin{aligned}\mathcal{L}[f](p) &= \mathcal{L}[y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_0 y](p) \\ &= \mathcal{L}[y^{(n)}](p) + a_{n-1}\mathcal{L}[y^{(n-1)}](p) + \ldots + a_0\mathcal{L}[y](p) \\ &= p^n\mathcal{L}[y](p) + a_{n-1}p^{n-1}\mathcal{L}[y](p) + \ldots + a_0\mathcal{L}[y](p) \\ &= \underbrace{(p^n + a_{n-1}p^{n-1} + \ldots + a_0)}_{P(p)}\mathcal{L}[y](p)\end{aligned}$$

$$\Rightarrow \mathcal{L}[y](p) = \frac{\mathcal{L}[f](p)}{P(p)}$$

$$\Rightarrow y = \mathcal{L}^{-1}\left[\frac{\mathcal{L}[f](p)}{P(p)}\right].$$

Wir suchen y in der Form $y = Q * f$ mit $\mathcal{L}[Q](p) = \frac{1}{P(p)}$. Damit ist nämlich

$$\mathcal{L}[y](p) = \mathcal{L}[Q](p) \cdot \mathcal{L}[f](p) = \frac{\mathcal{L}[f](p)}{P(p)}.$$

Wie kann man ein solches Q finden? Es ist

$$\frac{1}{P(p)} = (p-\lambda_1)^{-\nu_1} \cdot \ldots \cdot (p-\lambda_k)^{-\nu_k}.$$

Es ist

$$(p-\lambda_l)^{-\nu_l} = \mathcal{L}[j_{\nu_l}(\lambda_l;t)](p) \quad \text{mit} \quad j_\nu(\lambda;t) = \frac{e^{\lambda t}t^{\nu-1}}{(\nu-1)!}H(t).$$

Dann ist

$$Q(t) = j_{\nu_1}(\lambda_1;t) * \ldots * j_{\nu_k}(\lambda_k;t).$$

Damit ist

$$y(t) = (Q * f)(t) = (j_{\nu_1}(\lambda_1,\cdot) * \ldots * j_{\nu_k}(\lambda_k,\cdot) * f)(t).$$

Beispiel: $\ddot{y}(t) - y(t) = f(t)$, dann ist $P(p) = p^2 - 1 = (p-1)(p+1)$. Damit ist dann

$$\mathcal{L}^{-1}\left[\frac{1}{p+1}\right] = e^{-t}H(t)$$
$$\mathcal{L}^{-1}\left[\frac{1}{p-1}\right] = e^{t}H(t).$$

Es folgt

$$\begin{aligned}Q(t) &= (e^{-t}H(t)) * (e^tH(t)) \\ &= \int_0^t e^{-\tau}e^{t-\tau}\,\mathrm{d}\tau = \sinh t, \qquad t \geq 0.\end{aligned}$$

Also ist

$$y(t) = \int_0^t \sinh(t-\tau)f(\tau)\,\mathrm{d}\tau.$$

9.16 Zum Langzeitverhalten autonomer Systeme

$\dot{\mathbb{y}} = \mathbb{v}(\mathbb{y})$. Frage: Gibt es konstante Lösungen? $\Rightarrow$ $\dot{\mathbb{y}} = 0 = \mathbb{v}(\mathbb{y})$.
Beispiel:

$$\dot{\mathbb{y}} = \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = A\mathbb{y} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Für $\det A \neq 0$ folgt für die Eigenwerte $\lambda_1, \lambda_2 \neq 0$. $\mathbb{b}_1, \mathbb{b}_2$ seien Eigenvektoren. Kritischer Punkt: $\mathbb{y} = 0$. Es ist $y(t) = \alpha_1 \mathbb{b}_1 e^{\lambda_1 t} + \alpha_2 \mathbb{b}_2 e^{\lambda_2 t}$.

10 Oberflächen- und Volumenintegrale. Elemente der Vektoranalysis

10.1 Produktmaß. Der Satz von Fubini

Gegeben sei $(\mathbb{R}^n, \mathrm{d}^n x)$. Dann ist ein Integralbegriff $\int_{\mathbb{R}^n} f(x_1, \ldots, x_n)\, \mathrm{d}^n x$ definiert (mit $x_j \in \mathbb{R}$). Wir wollen diese Integrale nun praktisch berechnen. Der Satz von Newton-Leibniz greift allerdings nur im Eindimensionalen, weshalb wir dieses Integral auf ein eindimensionale Integrale zurückführen wollen.
Ein Ansatz könnte so aussehen:

$$\int_{\mathbb{R}^n} f(x_1, \ldots, x_n)\, \mathrm{d}^n x = \int \cdots \left(\int f(x_1, \ldots, x_n)\, \mathrm{d}x_1 \right) \cdots\, \mathrm{d}x_n.$$

Da der Ausdruck auf der linken Seite wohldefiniert ist, müssen wir jedoch untersuchen, inwieweit auch die rechte Seite wohldefiniert und unabhängig von der Reihenfolge ist.

Es seien $(X, \mathcal{A}_X, \mu)$ und $(Y, \mathcal{A}_Y, \nu)$ Maßräume. Es sei $(x, y) \in X \times Y$ (also $x \in X$ und $y \in Y$). Wir wollen ein neues Maß auf dem Produkt $X \times Y$ definieren, wobei dieses insofern „sinnvoll" geschehen soll, als dass es mit den bereits bekannten Maßen zusammenhängt. Für $a_x \in \mathcal{A}_X$ und $a_y \in \mathcal{A}_Y$ soll daher das Produkt $a_x \times a_y$ messbar sein. Ferner kann man sich überlegen, dass die Forderung der Gültigkeit von $(\mu \otimes \nu)(a_x \times a_y) = \mu(a_x) \cdot \nu(a_y)$ sinnvoll ist, also das Verhalten auf „Rechtecken". Kann man dies widerspruchsfrei auf $\mathcal{A}_{X \times Y}$ erweitern?
Dabei soll $\mathcal{A}_{X \times Y}$ die kleinste σ-Algebra sein, welche alle „Rechtecke" $a_x \times a_y$ mit $a_x \in \mathcal{A}_X$ und $a_y \in \mathcal{A}_Y$ enthält.

Aufgabe. B_n und B_m seien die Borelalgebren auf $\mathbb{R}^n$ bzw. $\mathbb{R}^m$. Dann ist das Produkt (im Sinne der kleinsten σ-Algebra) $B_{\mathbb{R}^n \times \mathbb{R}^m} \stackrel{!}{=} B_{n+m}$ die Borelalgebra auf $\mathbb{R}^{n+m}$. Dies muss gezeigt werden, da beide Ausdrücke zunächst verschiedene Konstruktionen sind.

Satz 10.1
Gegeben sei eine Menge $E \in \mathcal{A}_{X \times Y}$ auf der *kleinsten* σ-Algebra wie oben. Fixiere einen Punkt $x \in X$. Man erhält dann eine „Projektion" $E_x = \{y | (x, y) \in E\} \subset Y$.

Analog kann man ein $y \in Y$ fixieren und erhält dann eine „Projektion“ $E_y = \{x|(x,y) \in E\} \subset X$. Dann folgt für alle $x \in X$, dass $E_x \in \mathcal{A}_Y$ und für alle $y \in Y$, dass $E_y \in \mathcal{A}_X$ gilt.

Beweis. Es sei $\mathcal{D} \subset \mathcal{A}_{X\times Y}$ die Klasse aller Mengen, für die für alle $x \in X$ gilt, dass $E_x \in \mathcal{A}_Y$ ist. Wir wollen zeigen: $\mathcal{D} = \mathcal{A}_{X\times Y}$. Wir wissen:

1. $\mathcal{R} \subset \mathcal{D}$, $\mathcal{R} = \{a_x \times a_y, a_x \in \mathcal{A}_X, a_y \in \mathcal{A}_Y\}$.
2. $\mathcal{D}$ ist eine σ-Algebra.

Daraus folgt $\mathcal{R} \subset \mathcal{A}_{X\times Y} \subset \mathcal{D}$, da $\mathcal{A}_{X\times Y}$ die kleinste σ-Algebra ist, die $\mathcal{R}$ enthält. Dann folgt aber $\mathcal{A}_{X\times Y} = \mathcal{D}$. □

Da wir nun wissen, dass jede Menge E_x messbar ist in Y ist, können wir $\nu(E_x) =: f(x)$ berechnen. Analog für $\mu(E_y) =: g(y)$. Da $E \in \mathcal{A}_{X\times Y}$ war sind diese Ausdrücke insbesondere wohldefiniert. Ferner gilt $0 \le f(x)$ und $0 \le g(y)$.
Es sei $E = a_x \times a_y$ mit $a_x \in \mathcal{A}_X$ und $a_y \in \mathcal{A}_Y$ (E ist also ein „Rechteck“). Dann ist

$$f(x) = \begin{cases} 0 & x \notin a_x \\ \nu(a_y) & x \in a_x \end{cases}, \qquad g(y) = \begin{cases} 0 & y \notin a_y \\ \mu(a_x) & y \in a_y \end{cases}$$

Es gilt dann

$$\int_X f(x)\,\mathrm{d}\mu = \int_{a_x} \nu(a_y)\,\mathrm{d}\mu(x) = \mu(a_x)\cdot\nu(a_y) = \int_{a_y} \mu(a_x)\,\mathrm{d}\nu(y) = \int_Y g(y)\,\mathrm{d}\nu(y).$$

Wir stellen uns nun die Frage, ob die Identität $\int_X f(x)\,\mathrm{d}\mu = \int_Y g(y)\,\mathrm{d}\nu$ für alle $E \in \mathcal{A}_{X\times Y}$ gilt.

Definition 10.2 **σ-finiter Maßraum**
Ein Maßraum heißt σ-finit, wenn er sich durch höchstens abzählbar viele Mengen von endlichem Maß überdecken lässt.

Satz 10.3
μ, ν seien σ-finite Maße. Dann gilt

$$\int_X f(x)\,\mathrm{d}\mu \overset{(*)}{=} \int_Y g(y)\,\mathrm{d}\nu \qquad \forall E \in \mathcal{A}_{X\times Y}.$$

Beweis. Sei $\mathcal{D} \subset \mathcal{A}_{X\times Y}$ die Klasse aller Mengen $E \in \mathcal{A}_{X\times Y}$, für welche $(*)$ gilt. Wir wollen zeigen, dass $\mathcal{D} = \mathcal{A}_{X\times Y}$ gilt.
Schritt 1: Für Rechtecke haben wir die Gleichung bereits bewiesen, d.h. $\mathcal{R} \subset \mathcal{D}$, wobei $\mathcal{R} = \{a_x \times a_y\}$ mit $a_x \in \mathcal{A}_X$, $a_y \in \mathcal{A}_Y$. Nun sei $\mathcal{F}$ die Menge aller endlichen, disjunkten Vereinigungen von „Rechtecken“ aus $\mathcal{R}$. Man kann dann technisch recht leicht zeigen, dass $\mathcal{F} \subset \mathcal{D}$ gilt. Gegeben seien nun disjunkte $E_n \in \mathcal{D}$ für $n = 1, \dots, N$. Es sei $E := \bigcup_{n=1}^N E_n$.

Dann ist $E_x = \bigcup_{n=1}^N (E_n)_x$ disjunkt. Dann folgt $\nu(E_x) = \sum_{n=1}^N \nu((E_n)_x)$ und damit

$$\begin{aligned}\int_X \nu(E_x)\ \mathrm{d}\mu &= \int_X \sum_{n=1}^N \nu((E_n)_x)\ \mathrm{d}\mu = \sum_{n=1}^N \int_X \nu((E_n)_x)\ \mathrm{d}\mu \\ &\overset{E_n \in \mathcal{D}}{=} \sum_{n=1}^N \int_Y \mu((E_n)_x)\ \mathrm{d}\nu = \int_Y \sum_{n=1}^N \mu((E_n)_x)\ \mathrm{d}\nu = \int_Y \mu(E_x)\ \mathrm{d}\nu.\end{aligned}$$

Also ist $E = \bigcup_{n=1}^N E_n \in \mathcal{D}$.

Schritt 2: Gegeben sei ein monotones System von Mengen mit $E_1 \subset E_2 \subset \ldots \subset E_n \subset \ldots$ mit $E_n \in \mathcal{D}$. Weiter sei $E_1' \supset E_2' \supset \ldots \supset E_n' \supset \ldots$ ein monotones System von Mengen mit $E_n' \in \mathcal{D}$. Dann ist $\lim_{n\to\infty} E_n = E = \bigcup_{n=1}^\infty E_n$ und $\lim_{n\to\infty} E_n' = E' = \bigcap_{n=1}^\infty E_n'$. Wir wollen zeigen, dass $E, E' \in \mathcal{D}$ ist. Man nennt $\mathcal{D}$ in diesem Fall eine monotone Mengenklasse. Es sei $f_n(x) = \nu((E_n)_x)$ und $g_n(y) = \mu((E_n)_y)$. Diese Folgen sind beide monoton und messbar. Wir können nun annehmen, dass $\nu(Y) < \infty$ und $\mu(X) < \infty$ ist. Dann folgt $0 \le f_n(x) \le \nu(Y) < \infty$ und $0 \le g_n(y) \le \mu(X) < \infty$. Dann existiert ein $\lim_{n\to\infty} f_n(x) =: f(x)$ und ein $\lim_{n\to\infty} g_n(y) =: g(y)$. Auch dies sind dann messbare Funktionen. Es gilt dann

$$\nu(E_x) = \nu\left(\bigcup_n (E_n)_x\right) = \lim_{n\to\infty} \nu((E_n)_x) = \lim_{n\to\infty} f_n(x) = f(x)$$

und analog $\mu(E_y) = \ldots = g(y)$. Damit ist mit dem Satz von Lebesgue zur monotonen Konvergenz

$$\begin{aligned}\int_X f(x)\ \mathrm{d}\mu = \int_X \nu(E_x)\ \mathrm{d}\mu &= \int_X \lim_{n\to\infty} f_n(x)\ \mathrm{d}\mu \\ &\overset{\text{S.v.L.}}{=} \lim_{n\to\infty} \int_X f_n(x)\ \mathrm{d}\mu \\ &\overset{E_n \in \mathcal{D}}{=} \lim_{n\to\infty} \int_Y g_n(y)\ \mathrm{d}\nu \\ &\overset{\text{S.v.L.}}{=} \int_Y \lim_{n\to\infty} g_n(y)\ \mathrm{d}\nu = \int_Y g(y)\ \mathrm{d}\nu.\end{aligned}$$

Dies gilt völlig analog für den zweiten Fall mit den E_n' und E'. Also ist wirklich $E, E' \in \mathcal{D}$.

Schritt 3: Enthält ein monotones Mengensystem ($\mathcal{D}$) einen Ring ($\mathcal{F}$), so enthält es auch den minimalen σ-Ring ($\mathcal{F}^*$), welcher durch den Ring erzeugt wird. Dann folgt $\mathcal{R} \subset \mathcal{F} \subset \mathcal{F}^* \subset \mathcal{D}$. Nach Definition ist $A_{X\times Y}$ die kleinste σ-Algebra, welche $\mathcal{R}$ enthält und damit folgt $\mathcal{R} \subset A_{X\times Y} \subset \mathcal{F}^* \subset \mathcal{D}$. Insbesondere gilt dann also $A_{X\times Y} \subset \mathcal{D}$. Zusammen mit $\mathcal{D} \subset A_{X\times Y}$ folgt dann $\mathcal{D} = A_{X\times Y}$. Zum Beweis dieses Schrittes benötigen wir Hilfsmittel. In diesem Beweis steht daher nur die Idee, im Folgenden wird dieser Schritt dann ausgeführt. □

Wir wenden uns nun Schritt 3 zu.

Lemma 10.4
Die monotone Mengenklasse M sei ein Ring. Dann ist M ein σ-Ring.

Beweis. Es seien $F_n \in M$ für $n \in \mathbb{N}$. Zu zeigen ist dann $\bigcup_{n\in\mathbb{N}} F_n = F \in M$. Es ist

$$\bigcup_{n\in\mathbb{N}} F_n = \bigcup_{k=1}^{\infty} \underbrace{\left(\bigcup_{n=1}^{k} F_n\right)}_{=:E_k}.$$

Die E_k sind monoton. Aus der Monotonieeigenschaft von M folgt dann die Behauptung.
□

Der Durchschnitt beliebig vieler monotoner Mengenklassen ist selbst monoton. Daraus folgt, dass es eine minimale monotone Mengenklasse $M(\mathcal{F})$ gibt, welche $\mathcal{F}$ enthält. Dann folgt, dass $\mathcal{F} \subset M(\mathcal{F}) \subset \mathcal{D}$ ist.
Idee: Wir zeigen, dass $M(\mathcal{F})$ ein Ring ist. Dann folgt aus obigem Lemma auch, dass es ein σ-Ring ist, das heißt dies ist der von $\mathcal{F}$ erzeugte σ-Ring, $M(\mathcal{F}) \subset \mathcal{D}$.
Für $G \in M(\mathcal{F})$ sei

$$\Theta_G = \{E \in M(\mathcal{F}) \mid E \setminus G \in M(\mathcal{F}) \;\wedge\; E \cap G \in M(\mathcal{F}) \;\wedge\; E \cup G \in M(\mathcal{F})\}.$$

Dann gilt:

$$M(\mathcal{F}) \text{ ist ein Ring } \;\Leftrightarrow\; \forall_G \Theta_G = M(\mathcal{F}).$$

Eigenschaften von Θ_G:

1. $E \in \Theta_G \;\Leftrightarrow\; G \in \Theta_E$.
2. Θ_G ist eine monotone Mengenklasse. Sei $\{E_n\}$ eine monotone Folge aus Θ_G. Es sei $E = \lim_{n\to\infty} E_n$. Dann ist

$$E \setminus G = \left(\lim_{n\to\infty} E_n \setminus G\right) = \lim_{n\to\infty} (E_n \setminus G) \in M(\mathcal{F})$$

   Analog geht dies für $E \cup G$ und $E \cap G$.

Sei zunächst $G \in \mathcal{F}$. Für $E \in \mathcal{F}$ folgt $G \cup E, G \cap E, G \setminus E \in \mathcal{F} \subset M(\mathcal{F})$. Dies ist äquivalent zu $\mathcal{F} \subset \Theta_G$. Daraus folgt dann $\mathcal{F} \subset M(\mathcal{F}) \subset \Theta_G$, da $M(\mathcal{F})$. Nach Definition gilt aber auch $\Theta_G \subset M(\mathcal{F})$ und damit ist $\Theta_G = M(\mathcal{F})$ für $G \in \mathcal{F}$.

$$\begin{aligned}
&\forall_{G\in\mathcal{F}} \Theta_G = M(\mathcal{F}) \\
\Leftrightarrow &\forall_{G\in\mathcal{F}} \forall_{E\in M(\mathcal{F})} E \in \Theta_G \\
\Leftrightarrow &\forall_{E\in M(\mathcal{F})} \forall_{G\in\mathcal{F}} G \in \Theta_E \\
\Rightarrow &\forall_{E\in M(\mathcal{F})} F \subset \Theta_E \\
\Rightarrow &\forall_{E\in M(\mathcal{F})} F \subset M(\mathcal{F}) \subset \Theta_E \\
\Rightarrow &\forall_{E\in M(\mathcal{F})} \Theta_E = M(\mathcal{F}).
\end{aligned}$$

Satz 10.5
Seien $(X, \mathcal{A}_X, \mu)$ und $(Y, \mathcal{A}_Y, \nu)$ σ-finite Maßräume. Definiere

$$\begin{aligned}(\mu \otimes \nu) \colon \mathcal{A}_{X\times Y} &\to [0, +\infty] \\ E &\mapsto \int_X \nu(E_x)\ \mathrm{d}\mu = \int_Y \mu(E_y)\ \mathrm{d}\nu.\end{aligned}$$

Dies ist ein σ-additives Maß.

Anmerkung: Sei $E = a_X \times a_Y$. Dann folgt $(\mu \otimes \nu)(E) = \mu(a_X)\nu(a_Y)$.

Beweis. Es ist $\mu(X) < \infty$, $\nu(Y) < \infty$. Es seien $E_n \in A_{X\times Y}$ $(n \in \mathbb{N})$ disjunkt und $E = \bigcup_{n\in\mathbb{N}} E_n \in A_{X\times Y}$. Dann ist auch $E_x = \bigcup_{n\in\mathbb{N}}(E_n)_x$, $E_y = \bigcup_{n\in\mathbb{N}}(E_n)_y$ disjunkt. Es folgt

$$\begin{aligned}&0 \le f_n(x) = \nu((E_n)_x) \le \nu(Y) \\ &0 \le g_n(y) = \mu((E_n)_y) \le \mu(X).\end{aligned}$$

Die σ-Additivität auf $\mathcal{A}_X$ bzw. $\mathcal{A}_Y$ ergibt:

$$\begin{aligned}&0 \le \nu(E_x) = \sum_{n=1}^{\infty} \nu((E_n)_x) = \sum_{n=1}^{\infty} f_n(x) \le \nu(Y) < \infty \\ &0 \le \mu(E_y) = \sum_{n=1}^{\infty} g_n(y) \le \mu(X) < \infty.\end{aligned}$$

$\sum_{n=1}^{k} f_n(x)$ ist monoton und beschränkt durch $\nu(Y) < \infty$. Daraus folgt mit dem Satz von Lebesgue

$$\begin{aligned}(\mu \otimes \nu)(E) = \int_X \nu(E_x)\ \mathrm{d}\mu = \int_X \sum_{n=1}^{\infty} f_n(x)\ \mathrm{d}\mu &= \int_X \lim_{k\to\infty} \sum_{n=1}^{k} f_n(x)\ \mathrm{d}\mu \\ &\overset{\text{S.v.L.}}{=} \lim_{k\to\infty} \sum_{n=1}^{k} \int_X f_n(x)\ \mathrm{d}\mu \\ &= \sum_{n=1}^{\infty} \int_X \nu((E_n)_x)\ \mathrm{d}\mu \\ &= \sum_{k=1}^{\infty} (\mu \otimes \nu)(E_n).\end{aligned}$$

Damit ist die σ-Additivität bewiesen. ☐