BERGISCHE
UNIVERSITÄT
WUPPERTAL

Analysis III

JOCHEN GLÜCK

Vorlesungsmanuskript
Wintersemester 2025/2026

Version: 14. Juli 2026

Inhaltsverzeichnis

Einleitung

Notation

Folgende Notation wird im Manuskript durchgehend verwendet:

Notation	Bedeutung
$\mathbb{N}$	$\{1,2,3,\dots\}$
$\mathbb{N}_0$	$\{0,1,2,3,\dots\}$
$\mathbb{K}$	$\mathbb{R}$ or $\mathbb{C}$

Standardliteratur zur Vorlesung

Einige deutschsprachige Standardwerke zu den Themen der Vorlesung finden Sie im Literaturverzeichnis des Manuskripts unter den Einträgen [1, 2, 5, 7].

Vorsicht: Fehler!

Dieses Manuskript enthält mit an Sicherheit grenzender Wahrscheinlichkeit Fehler. Mit hoher Wahrscheinlichkeit enthält es sogar viele Fehler. Wenn Sie glauben, einen Fehler entdeckt zu haben: Geben Sie mir bitte Bescheid (wirklich!), am einfachsten per E-Mail.

E-Mail-Adresse: glueck@uni-wuppertal.de

Danksagung

Ich danke allen Studierenden, die mir Korrekturhinweise zur Verbesserung dieses Manuskripts geschickt haben, inbesondere Amélie Krempel und Max Messerschmidt.

Stand

Diese Version des Manuskript stammt vom 14. Juli 2026.

Kapitel 1

Maße

1.1 Messbarkeit und σ-Algebren

Vorlesung 1 (Mi, 15.10.) beginnt am Anfang von Abschnitt 1.1

Einstiegsfragen.

(a) Welche Teilmengen von $\mathbb{R}^3$ besitzen ein Volumen?

(b) Sie werfen einen Würfel zehn mal nacheinander. Sei $w \in \{1,\dots,6\}^{10}$ der Vektor, der im k-ten Eintrag die Augenzahl Ihres k-ten Wurfs angibt. Können Sie nach dem dritten Wurf bereits entscheiden, ob $w \in \{1,2\}^3 \times \{1,\dots,6\}^7$ gilt? Und können Sie nach dem dritten Wurf bereits entscheiden, ob $w \in \{1,2\}^{10}$ gilt?

Für welche Teilmengen $A \subseteq \{1,\dots,6\}^{10}$ können Sie nach dem dritten Wurf bereits entscheiden, ob $w \in A$ gilt?

(c) Erinnerung an die Analysis 2: Was hat Stetigkeit von Funktionen mit Urbildern zu tun?

Definition 1.1.1 (Komplement von Mengen)**.** Sei Ω eine Menge und $S \subseteq \Omega$. Dann nennt man die Menge $S^\complement := \Omega \setminus S$ das **Komplement von S**, oder genauer das **Komplement von S in Ω**.

Beachten Sie, dass sich die Notation $S^\complement$ nur dann sinnvoll verwenden lässt, wenn die umgebende Menge Ω aus dem Kontext klar ist.

Definition 1.1.2 (σ-Algebren und Messräume)**.**

(a) Sei Ω eine Menge. Eine Teilmenge $\mathcal{A} \subseteq 2^\Omega$ der Potenzmenge 2^Ω heißt eine **σ-Algebra** oder genauer eine **σ-Algebra auf Ω**, falls sie die folgenden Axiome erfüllt:

(I) Es ist $\emptyset \in \mathcal{A}$.

(II) Für jedes $A \in \mathcal{A}$ gilt auch $A^\complement \in \mathcal{A}$.

(III) Für jede Folge $A_1, A_2, A_3, \dots \in \mathcal{A}$ gilt auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

(b) Ein **Messraum** ist ein Paar $(\Omega, \mathcal{A})$, wobei Ω eine Menge und $\mathcal{A}$ eine σ-Algebra auf Ω ist.

Ist $(\Omega, \mathcal{A})$ ein Messraum, dann verwendet man oft die folgende Sprechweise: Man nennt eine Menge $A \subseteq \Omega$ **messbar**, falls $A \in \mathcal{A}$ gilt.

Beachten Sie: Welche Mengen $A \subseteq \Omega$ messbar sind, hängt also von der σ-Algebra ab, die man auf Ω betrachtet.

In dieser Veranstaltung werden wir sehr oft Teilmengen $\mathcal{A} \subseteq 2^\Omega$ betrachten. Beachten Sie, dass jedes Element von $\mathcal{A}$ selbst eine Menge (genauer: eine Teilmenge von Ω) ist. Deshalb bezeichnet man ein $\mathcal{A} \subseteq 2^\Omega$ oft als eine **Menge von Mengen** oder – etwas griffiger – als ein **Mengensystem**.

Beispiele 1.1.3 (Einige σ-Algebren)**.**

(a) Sei Ω eine Menge. Dann ist die Potenzmenge 2^Ω selbst eine σ-Algebra auf Ω.

(b) Sei Ω eine Menge. Dann ist $\{\emptyset, \Omega\}$ eine σ-Algebra auf Ω.

(c) Sei Ω eine Menge und seien $k, \ell \in \mathbb{N}$ mit $k \leq \ell$. Dann ist

$$\mathcal{A} := \left\{ B \times \Omega^{\ell-k} \mid B \subseteq \Omega^k \right\}$$

eine σ-Algebra auf Ω^ℓ.

Beweis. (a) und (b) Die Axiome einer σ-Algebra lassen sich in diesen beiden Beispielen leicht überprüfen.

(c) Wir überprüfen, dass $\mathcal{A}$ die Axiome einer σ-Algebra (Definition 1.1.2(a)) erfüllt:

Axiom (I)*:* Es gilt $\emptyset = \emptyset \times \Omega^{\ell-k} \in \mathcal{A}$, da $\emptyset \subseteq \Omega^k$ ist.

Axiom (II)*:* Sei $A \in \mathcal{A}$. Dann gibt es eine Menge $B \subseteq \Omega^k$ mit $A = B \times \Omega^{\ell-k}$. Es gilt

$$\Omega^\ell \setminus A = (\Omega^k \setminus B) \times \Omega^{\ell-k} \in \mathcal{A};$$

hier haben wir die Schreibweise $\Omega^\ell \setminus A$ und $\Omega^k \setminus B$ anstelle von $A^\complement$ und $B^\complement$ verwendet um Verwechslungen auszuschließen, da die Komplemente von A und B in verschiedenen Mengen genommen werden.

Axiom (III)*:* Seien $A_1, A_2, \dots \in \mathcal{A}$. Dann gibt es eine Folge von Mengen $B_1, B_2, \dots \subseteq \Omega^k$ mit $A_n = B_n \times \Omega^{\ell-k}$ für alle $n \in \mathbb{N}$. Somit ist

$$\begin{aligned} \bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}} (B_n \times \Omega^{\ell-k}) &= \bigcup_{n\in\mathbb{N}} \left\{ (b,c) \mid b \in B_n,\ c \in \Omega^{k-\ell} \right\} \\ &= \left\{ (b,c) \mid b \in \bigcup_{n\in\mathbb{N}} B_n,\ c \in \Omega^{k-\ell} \right\} = \Big(\bigcup_{n\in\mathbb{N}} B_n \Big) \times \Omega^{\ell-k} \in \mathcal{A}, \end{aligned}$$

womit die Behauptung gezeigt ist. □

Beispiele 1.1.4 (Würfeln und Münzwurf)**.** Lassen Sie uns Beispiel 1.1.3(c) in zwei konkreten Situationen illustrieren.

(a) Sei Ω die Menge aller möglichen Ergebnisse eines Münzwurfs, also $\Omega = \{K, Z\}$, wobei wir K und Z als Abkürzungen für Kopf und Zahl verwenden. Wir werfen die Münze drei mal hintereinander; die Menge aller möglichen Ergebnisse ist somit gleich

$$\Omega^3 = \Omega \times \Omega \times \Omega = \Big\{(K,K,K),\ (K,K,Z),\ (K,Z,K),\ (K,Z,Z),\\ (Z,K,K),\ (Z,K,Z),\ (Z,Z,K),\ (Z,Z,Z)\Big\}.$$

Bezeichne $\omega \in \Omega^3$ das Gesamtergebnis Ihrer Dreimünzwürfe. Stellen Sie sich nun vor, dass Sie die Münze erst einmal geworfen haben. Dann kennen Sie ω_1 – das heißt, Sie wissen, ob $\omega_1 = K$ oder $\omega_1 = Z$ ist –, aber Sie wissen noch nicht, ob ω_2 und ω_3 jeweils Kopf oder Zahl sind.

Anders gesprochen: Sie können also beispielsweise bereits entscheiden, ob $\omega \in \{K\} \times \{K, Z\} \times \{K, Z\}$ gilt, aber Sie können noch nicht entscheiden, ob $\omega \in \{K, Z\} \times \{K\} \times \{K, Z\}$ gilt. Lassen Sie uns nun all diejenigen Mengen $A \subseteq \Omega^3$ auflisten, für die sich bereits nach dem ersten Wurf entscheiden lässt, ob $\omega \in A$ gilt. Dies sind die Mengen

$$\begin{aligned}\{K\} \times \{K, Z\} \times \{K, Z\},\\ \{Z\} \times \{K, Z\} \times \{K, Z\},\\ \{K, Z\} \times \{K, Z\} \times \{K, Z\} = \{K, Z\}^3,\\ \emptyset \times \{K, Z\} \times \{K, Z\} = \emptyset.\end{aligned}$$

Wenn wir diese vier Mengen als Elemente in einer Menge $\mathcal{A}$ packen, ist also

$$\mathcal{A} = \Big\{\{K\} \times \{K, Z\} \times \{K, Z\},\ \{Z\} \times \{K, Z\} \times \{K, Z\},\\ \{K, Z\} \times \{K, Z\} \times \{K, Z\},\ \emptyset \times \{K, Z\} \times \{K, Z\}\Big\} \subseteq 2^{\{K,Z\}^3}.$$

Es ist $\mathcal{A}$ genau die σ-Algebra aus Beispiel 1.1.3(c), für $\Omega = \{K, Z\}$, $k = 1$ und $\ell = 3$.

(b) Nun betrachten wir Beispiel 1.1.3(c) für $\Omega = \{1, \dots, 6\}$, $k = 3$ und $\ell = 10$. Dann ist

$$\mathcal{A} := \Big\{B \times \{1, \dots, 6\}^7 \mid B \subseteq \{1, \dots, 6\}^3\Big\}.$$

Wenn Sie einen Würfel mit den Augenzahlen 1 bis 6 zehn mal nacheinander werfen und sich alle Augenzahlen in der Reihenfolge notieren, in der Sie sie geworfen haben, dann erhalten Sie eine Liste $\omega \in \Omega^{10}$. Für jede Menge $A \in \mathcal{A}$ können Sie nach dem dritten Wurf bereits entscheiden, ob $\omega \in A$ gilt. Für Mengen $A \subseteq \Omega^{10}$ mit $A \notin \mathcal{A}$ können Sie hingegen nach dem dritten Wurf noch nicht entscheiden, ob $\omega \in A$ gilt.

In der Wahrscheinlichkeitsrechnung bezeichnet man Mengen $A \subseteq \Omega^{10}$ als *Ereignisse*, die beim zehnmaligen Werfen eines Würfels eintreten können (oder eben

auch nicht). Die Elemente der σ-Algebra $\mathcal{A}$ sind also genau diejenigen Ereignisse, von denen sich bereits nach dem dritten Wurf entscheiden lässt, ob sie eingetreten sind.

Bemerkung 1.1.5 (**σ-Algebren modellieren zugängliche Information**)**.** Die beiden Beispiele 1.1.4 zeigen, dass σ-Algebren verwendet werden können um zu modellieren, welche Information in einer bestimmten Situation oder zu einem bestimmen Zeitpunkt zugänglich oder verfügbar ist.

Vor diesem Hintergrund lassen sich auch die drei Axiome einer σ-Algebra (Definition 1.1.2(I)) gut interpretieren. Stellen Sie sich erneut den zehnfachen Wurf eines Würfels aus Beispiel 1.1.4(b) vor und sei $\omega \in \{1,\dots,6\}^{10}$ die Liste der dabei geworfenen Augenzahlen. Wie in diesem Beispiel bezeichne $\mathcal{A} \subseteq 2^{\{1,\dots,6\}^{10}}$ die Menge all derjenigen Ereignisse $A \in 2^{\{1,\dots,6\}}$, von denen sich bereits nach dem dritten Wurf entscheiden lässt, ob $\omega \in A$ gilt.

Axiom (I)*:* Natürlich kann man nach dem dritten Wurf entscheiden, ob $\omega \in \emptyset$ gilt: die Antwort ist Nein. Also ist $\emptyset$ ein Element von $\mathcal{A}$.

Axiom (II)*:* Sei $A \in \mathcal{A}$. Dann können Sie nach dem dritten Wurfen bereits entscheiden, ob $\omega \in A$ ist. Indem Sie diese Antwort negieren, können Sie nach dem dritten Wurf natürlich auch entscheiden, ob $\omega \in A^{\complement}$ gilt. Somit muss auch $A^{\complement} \in \mathcal{A}$ sein.

Axiom (III)*:* Wenn Sie Mengen $A_1, A_2, \dots \in \mathcal{A}$ betrachten, können Sie für jede dieser Mengen schon nach dem dritten Wurf entscheiden, ob ω in ihr liegt. Also können Sie auch nach dem dritten Wurf entscheiden, ob ω in mindestens einer dieser Mengen liegt, das heißt, ob $\omega \in \bigcup_{n\in\mathbb{N}} A_n$ gilt. Somit ist auch $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Dass man hier unendlich viele Mengen betrachtet, spielt in Beispiel 1.1.4(b) übrigens keine Rolle, die Menge $\{1,\dots,6\}^{10}$ ist ja endlich und somit besitzt sie auch nur endlich viele Teilmengen,[1] also tauchen in der Folge $A_1, A_2, \dots$ nur endlich viele verschiedene Mengen auf (von denen sich manche dann natürlich unendlich oft wiederholen).

Dass man auch unendlich viele Mengen vereinigen darf und wieder in $\mathcal{A}$ landet, wird dann wichtig, wenn Ω unendlich ist – was in den meisten Beispielen, die wir später betrachten, der Fall ist. Dass man nur Vereinigungen von abzählbar vielen Mengen zulässt, kann man anschaulich zum Beispiel folgendermaßen interpretieren: Wenn man überabzählbar viele Mengen $A_j \in \mathcal{A}$ betrachtet, mit j aus einer überabzählbaren Indexmenge J, und für jede einzelne j entscheiden kann, ob das Ereignis A_j eintritt, dann kann man trotzdem nicht unbedingt entscheiden, ob $\bigcup_{j\in J} A_j$ eintritt, weil man keine Möglichkeit hat, überabzählbar viele Ereignisse "durchzusehen" um zu gucken, ob sie alle eintreten.

[1] Genauer besitzt $\{1,\dots,6\}^{10}$ genau 6^{10} Elemente und somit $2^{6^{10}}$ Teilmengen. Übrigens ist $6^{10} \geq 60 \cdot 10^6$ (fragen Sie einen Taschenrechner oder Computer). Wegen $2^{10} = 1024 \geq 10^3$ folgt $2^{60} \geq 10^{18}$ und somit

$$2^{6^{10}} \geq 2^{60\cdot 10^6} \geq \left(10^{18}\right)^{10^6} = 10^{18\cdot 10^6}.$$

Also hat $2^{6^{10}}$ mehr als 18 Millionen Stellen – das heißt, $\{1,\dots,6\}^{10}$ hat wirklich extrem viele Teilmengen…

Proposition 1.1.6 (Einige Eigenschaften von $\boldsymbol{\sigma}$-Algebren)**.** *Sei $(\Omega, \mathcal{A})$ ein Messraum.*

(a) *Es gilt $\Omega \in \mathcal{A}$.*

(b) *Für alle $A_1, A_2, \dots \in \mathcal{A}$ gilt $\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$.*

(c) *Seien $m \in \mathbb{N}$ und $A_1, \dots, A_m \in \mathcal{A}$. Dann ist $A_1 \cup \dots \cup A_m \in \mathcal{A}$ und $A_1 \cap \dots \cap A_m \in \mathcal{A}$.*

Beweis. (a) Wegen $\emptyset \in \mathcal{A}$ gilt $\Omega = \emptyset^{\complement} \in \mathcal{A}$.

(b) Sei $A_1, A_2, \dots \in \mathcal{A}$. Für jedes $n \in \mathbb{N}$ gilt dann auch $A_n^{\complement} \in \mathcal{A}$ und somit ist

$$\Big(\bigcap_{n\in\mathbb{N}} A_n\Big)^{\complement} = \bigcup_{n\in\mathbb{N}} A_n^{\complement} \in \mathcal{A}, \qquad \text{also} \qquad \bigcap_{n\in\mathbb{N}} A_n = \Big(\Big(\bigcap_{n\in\mathbb{N}} A_n\Big)^{\complement}\Big)^{\complement} \in \mathcal{A}.$$

(c) Um die erste Aussage zu erhalten setzen wir $A_n := \emptyset \in \mathcal{A}$ für alle $n \in \mathbb{N}$ mit $n \geq m+1$. Dann ist $A_1 \cup \dots \cup A_m = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

Um die zweite Aussage zu zeigen setzen wir stattdessen $A_n := \Omega$ für alle $n \in \mathbb{N}$ mit $n \geq m+1$. Wegen (a) sind dann alle A_n messbar, also ist $A_1 \cap \dots \cap A_m = \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{A}$ wegen (b). □

Proposition 1.1.7 (Durchschnitt von $\boldsymbol{\sigma}$-Algebren)**.** *Sei Ω eine Menge.*

(a) *Sei J eine nichtleere Menge und für jedes $j \in J$ sei $\mathcal{A}_j$ eine σ-Algebra auf Ω. Dann ist auch $\bigcap_{j\in J} \mathcal{A}_j$ eine σ-Algebra auf Ω.*

(b) *Sei $\mathcal{E} \subseteq 2^{\Omega}$ und sei $\sigma(\mathcal{E}) \subseteq 2^{\Omega}$ der Durchschnitt aller σ-Algebren auf Ω, die $\mathcal{E}$ enthalten, das heißt es sei*

$$\sigma(\mathcal{E}) := \bigcap_{\substack{\mathcal{A}\supseteq\mathcal{E} \\ \mathcal{A} \text{ ist } \sigma\text{-Algebra auf } \Omega}} \mathcal{A}.$$

Dann ist $\sigma(\mathcal{E})$ die kleinste σ-Algebra auf Ω, die $\mathcal{E}$ enthält, das heißt:

(1) *Es ist $\sigma(\mathcal{E})$ eine σ-Algebra auf Ω und es gilt $\mathcal{E} \subseteq \sigma(\mathcal{E})$.*

(2) *Falls $\mathcal{A}$ eine σ-Algebra auf Ω ist und $\mathcal{E} \subseteq \mathcal{A}$ gilt, dann ist $\sigma(\mathcal{E}) \subseteq \mathcal{A}$.*

Beweis. (a) Wir zeigen, dass das Mengensystem $\bigcap_{j\in J} \mathcal{A}_j \subseteq 2^{\Omega}$ die drei Axiome einer σ-Algebra aus Definition 1.1.2(a) erfüllt:

Axiom (I)*:* Für jedes $j \in J$ gilt $\emptyset \in \mathcal{A}_j$, da $\mathcal{A}_j$ eine σ-Algebra ist. Also ist $\emptyset \in \bigcap_{j\in J} \mathcal{A}_j$.

Axiom (II)*:* Sei $A \in \bigcap_{j\in J} \mathcal{A}_j$. Für jedes $j \in J$ gilt dann $A \in \mathcal{A}_j$ und somit $A^{\complement} \in \mathcal{A}_j$, da $\mathcal{A}_j$ eine σ-Algebra ist. Also ist $A^{\complement} \in \bigcap_{j\in J} \mathcal{A}_j$.

Axiom (III)*:* Seien $A_1, A_2, \dots \in \bigcap_{j\in J} \mathcal{A}_j$. Für jedes $j \in J$ gilt dann $A_1, A_2, \dots \in \mathcal{A}_j$ und und somit $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}_j$, weil $\mathcal{A}_j$ eine σ-Algebra ist. Folglich ist $\bigcup_{n\in\mathbb{N}} A_n \in \bigcap_{j\in J} \mathcal{A}_j$.

(b) Wir zeigen die beiden Eigenschaften (1) und (2) einzeln:

(1) Als ein Durchschnitt von σ-Algebren auf Ω ist $\sigma(\mathcal{E})$ laut (a) ebenfalls eine σ-Algebra. Die Inklusion $\mathcal{E} \subseteq \sigma(\mathcal{E})$ gilt, weil alle Mengensysteme $\mathcal{A}$, die man schneidet um $\sigma(\mathcal{E})$ zu erhalten, Obermengen von $\mathcal{E}$ sind.

(2) Sei $\mathcal{A}$ eine σ-Algebra auf Ω und gelte $\mathcal{E} \subseteq \mathcal{A}$. Dann ist $\mathcal{A}$ eines der Mengensysteme, über die man schneidet, um $\sigma(\mathcal{E})$ zu erhalten, also gilt $\sigma(\mathcal{E}) \subseteq \mathcal{A}$. □

Definition 1.1.8 (Erzeugte $\boldsymbol{\sigma}$-Algebra)**.** Sei Ω eine Menge.

(a) Sei $\mathcal{E} \subseteq 2^\Omega$. Die σ-Algebra $\sigma(\mathcal{E})$ aus Proposition 1.1.7(b) nennt man die **von $\boldsymbol{\mathcal{E}}$ erzeugte $\boldsymbol{\sigma}$-Algebra** oder genauer die **von $\boldsymbol{\mathcal{E}}$ erzeugte $\boldsymbol{\sigma}$-Algebra auf $\boldsymbol{\Omega}$**.

(b) Sei $\mathcal{A}$ eine σ-Algebra auf Ω. Eine Menge $\mathcal{E} \subseteq 2^\Omega$ nennt man einen **Erzeuger** von $\mathcal{A}$, falls $\mathcal{A} = \sigma(\mathcal{E})$ gilt.

Vorlesung 2 (Fr, 17.10.) beginnt mit Beispiel 1.1.9

Beispiel 1.1.9 (Borel $\boldsymbol{\sigma}$-Algebra auf einem metrischen Raum)**.** Sei (M, d) ein metrischer Raum. Dann nennt man

$$\mathcal{B}(M) := \sigma\Big(\big\{U \subseteq M \mid U \text{ ist offen}\big\}\Big)$$

die **Borel-$\boldsymbol{\sigma}$-Algebra** auf M. Es gilt

$$\mathcal{B}(M) = \sigma\Big(\big\{C \subseteq M \mid C \text{ ist abgeschlossen}\big\}\Big).$$

Beweis. Der kürzeren Notation halber bezeichnen wir in diesem Beweis die Menge aller offenen Teilmengen von M mit $\mathcal{U} \subseteq 2^M$ und die Menge aller abgeschlossenen Teilmengen von M mit $\mathcal{C} \subseteq 2^M$. Per Definition ist $\mathcal{B}(M) = \sigma(\mathcal{U})$, wir müssen also $\sigma(\mathcal{U}) = \sigma(\mathcal{C})$ zeigen.

"$\subseteq$" Zuerst bemerken wir, dass $\mathcal{U} \subseteq \sigma(\mathcal{C})$ gilt: Für jedes $U \in \mathcal{U}$ ist nämlich $U^\complement \in \mathcal{C} \subseteq \sigma(\mathcal{C})$ und somit $U = (U^\complement)^\complement \in \sigma(\mathcal{C})$, da σ-Algebren abgeschlossen unter Komplementbildungen sind.

Somit ist $\sigma(\mathcal{C})$ eine σ-Algebra auf M, die $\mathcal{U}$ enthält. Weil $\sigma(\mathcal{U})$ die kleinste σ-Algebra auf M ist, die $\mathcal{U}$ enthält, folgt $\sigma(\mathcal{U}) \subseteq \sigma(\mathcal{C})$.

"$\supseteq$" Diese Inklusion zeigt man analog zur soeben bewiesenen Inklusion. □

Beachten Sie, dass die Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^d)$, die wir in der Analysis 2 besprochen haben (siehe [4, Definition 3.2.1]) ein Spezialfall von Beispiel 1.1.9 ist: Man erhält sie, wenn man in Beispiel 1.1.9 als metrischen Raum den $\mathbb{R}^d$ zusammen mit der Euklidischen Metrik nimmt.[2]

In der Analysis 2 hatten wir bewiesen, dass eine Abbildung zwischen zwei metrischen Räumen genau dann stetig ist, wenn Urbilder offener Mengen stets offen sind. Analog zu dieser Eigenschaft definiert man nun die Messbarkeit von Abbildungen zwischen Messräumen.

[2] Oder den $\mathbb{R}^d$ zusammen mit irgendeiner anderen Metrik, die von einer Norm auf $\mathbb{R}^d$ induziert wird, denn für all diese Metriken ist der Konvergenzbegriff auf $\mathbb{R}^d$ derselbe [4, Theorem 1.3.13] – somit hängt es nicht von der Wahl der Norm ab, welche Mengen in $\mathbb{R}^d$ offen sind und folglich hängt es ebenfalls nicht von der Wahl der Norm ab, welche Mengen in $\mathbb{R}^d$ offen sind.

Definition 1.1.10 (Messbare Funktionen)**.** Seien $(\Omega, \mathcal{A})$ und $(\hat{\Omega}, \hat{\mathcal{A}})$ Messräume. Eine Abbildung $f : \Omega \to \hat{\Omega}$ heißt **messbar**, falls für jede messbare Menge in $\hat{\Omega}$ auch das Urbild unter f messbar ist, das heißt, falls gilt:

$$\forall \hat{A} \in \hat{\mathcal{A}} : \quad f^{-1}(\hat{A}) \in \mathcal{A}$$

Proposition 1.1.11 (Messbarkeit via Erzeuger)**.** *Seien $(\Omega, \mathcal{A})$ und $(\hat{\Omega}, \hat{\mathcal{A}})$ Messräume, und sei $\hat{\mathcal{E}} \subseteq \hat{\mathcal{A}}$ ein Erzeuger von $\hat{\mathcal{A}}$. Eine Funktion $f\colon \Omega \to \hat{\Omega}$ ist genau dann messbar, wenn $f^{-1}(\hat{E}) \in \mathcal{A}$ für alle $\hat{E} \in \hat{\mathcal{E}}$ gilt.*

Beweis. *"⇒"* Sei f messbar und sei $\hat{E} \in \hat{\mathcal{E}}$. Dann ist auch $\hat{E} \in \hat{\mathcal{A}}$, also gilt wegen der Messbarkeit von f die Eigenschaft $f^{-1}(\hat{E}) \in \mathcal{A}$.

"⇐" Dies ist die interessante Implikation. Es gelte $f^{-1}(\hat{E}) \in \mathcal{A}$ für alle $\hat{E} \in \hat{\mathcal{E}}$. Dazu definieren wir das Mengensystem

$$\hat{\mathcal{B}} := \{\hat{B} \subseteq \hat{\Omega} \mid f^{-1}(\hat{B}) \in \mathcal{A}\} \subseteq 2^{\hat{\Omega}}.$$

Wegen der Prämisse der Implikation, die wir gerade zeigen, gilt $\hat{\mathcal{E}} \subseteq \hat{\mathcal{B}}$. Wir zeigen nun, dass $\hat{\mathcal{B}}$ eine σ-Algebra auf $\hat{\Omega}$ ist:

Axiom (I) Es gilt $f^{-1}(\emptyset) = \emptyset \in \mathcal{A}$, weil $\mathcal{A}$ eine σ-Algebra ist. Also ist $\emptyset \in \hat{\mathcal{B}}$.

Axiom (II) Sei $\hat{B} \in \hat{\mathcal{B}}$. Dann gilt $f^{-1}(\hat{B}) \in \mathcal{A}$. Also folgt $f^{-1}(\hat{\Omega} \setminus \hat{B}) = \Omega \setminus f^{-1}(\hat{B})$, und letztere Menge ist ebenfalls ein Elemente von $\mathcal{A}$, da $\mathcal{A}$ eine σ-Algebra ist. Somit gilt $\hat{\Omega} \setminus \hat{B} \in \hat{\mathcal{B}}$.

Axiom (III) Seien $\hat{B}_1, \hat{B}_2, \dots \in \hat{\mathcal{B}}$. Für jedes $n \in \mathbb{N}$ gilt dann $f^{-1}(\hat{B}_n) \in \mathcal{A}$ und weil $\mathcal{A}$ eine σ-Algebra ist, folgt somit

$$f^{-1}\Big(\bigcup_{n\in\mathbb{N}} \hat{B}_n\Big) = \bigcup_{n\in\mathbb{N}} f^{-1}(\hat{B}_n) \in \mathcal{A},$$

also ist $\bigcup_{n\in\mathbb{N}} \hat{B}_n \in \hat{\mathcal{B}}$.

Also ist $\hat{\mathcal{B}}$ eine σ-Algebra auf $\hat{\Omega}$, die $\hat{\mathcal{E}}$ enthält. Somit folgt $\hat{\mathcal{B}} \supseteq \sigma(\hat{\mathcal{E}}) = \hat{\mathcal{A}}$. Also folgt für jedes $\hat{A} \in \hat{\mathcal{A}}$, dass $\hat{A} \in \hat{\mathcal{B}}$ ist und somit $f^{-1}(\hat{A}) \in \mathcal{A}$ gilt. □

Die Beweistechnik, die wir verwendet haben um Proposition 1.1.11 zu zeigen, bezeichnet man häufig als das **Prinzip der guten Mengen**. Allgemeiner kann man das Prinzip in folgendem Lemma zusammenfassen:

Bemerkung 1.1.12 (Prinzip der guten Mengen)**.** Sei $(\Omega, \mathcal{A})$ ein Messraum und sei $\mathcal{E} \subseteq 2^{\Omega}$ ein Erzeuger von $\mathcal{A}$. Für jedes $A \in \mathcal{A}$ sei $P(A)$ eine Aussage. Wenn

$$\mathcal{B} := \{A \in \mathcal{A} \mid P(A) \text{ ist wahr}\}$$

eine σ-Algebra auf Ω ist und für jedes $E \in \mathcal{E}$ die Aussage $P(E)$ wahr ist, dann ist für jedes $A \in \mathcal{A}$ die Aussage $P(A)$ wahr.

Die Bezeichnung *Prinzip der guten Mengen* rührt daher, dass man sich diejenigen Mengen $A \in \mathcal{A}$, für die $P(A)$ wahr ist, als die "guten" Teilmengen von Ω vorstellt (eben weil sie die gewünschte Eigenschaft haben). Etwas umgangssprachlicher ausgedrückt besagt das Prinzip der guten Mengen also: Wenn jede Menge aus einem Erzeuger $\mathcal{E}$ einer σ-Algebra $\mathcal{A}$ gut ist und die guten Mengen ebenfalls eine σ-Algebra bilden, dann ist jede Menge in $\mathcal{A}$ gut.

Beweis von Bemerkung 1.1.12. Laut Voraussetzung ist $\mathcal{B}$ eine σ-Algebra, die $\mathcal{E}$ enthält, also gilt $\mathcal{B} \supseteq \sigma(\mathcal{E}) = \mathcal{A}$. Für jede Menge $A \in \mathcal{A}$ gilt also $A \in \mathcal{B}$, das heißt für jedes $A \in \mathcal{A}$ ist $P(A)$ wahr. □

Sie sehen im Beweis, dass Bemerkung 1.1.12 eigentlich gar kein eigenständiges Resultat ist – es handelt sich nur um eine etwas andere Formulierung davon, dass eine σ-Algebra, die von $\mathcal{E}$ erzeugt wird, die kleinste σ-Algebra ist, die $\mathcal{E}$ enthält. Der Witz an Bemerkung 1.1.12 ist, dass wir es auf eine Weise aufgeschrieben haben, in der diese Beobachtung sehr häufig und intuitiv eingesetzt werden kann.

Sie werden sehen, dass das Prinzip der guten Mengen und verwandte Beweistechniken in der Maß- und Integrationstheorie eine sehr große Rolle spielen. Das liegt daran, dass es oft nicht möglich ist, direkt anzugeben, welche Mengen konkret in einer σ-Algebra liegen (Ausnahme bestätigen die Regel); deshalb arbeitet man häufig mit Erzeugern von σ-Algebren. Man will also oft Eigenschaften für alle Mengen in einer σ-Algebra $\mathcal{A}$ zeigen, obwohl man nur einen Erzeuger $\mathcal{E}$ von $\mathcal{A}$ kennt und obwohl man kein konkretes Verfahren hat, um alle Elemente von $\mathcal{A}$ aus den Elementen von $\mathcal{E}$ zu konstruieren. Hierfür ist Bemerkung 1.1.12 äußerst nützlich.

1.2 Maße

Einstiegsfragen.

(a) Betrachten wir einen physikalischen Körper im $\mathbb{R}^3$. Was für ein mathematisches Objekt ist die Masseverteilung des Körpers?

(b) Was ist eigentlich eine Wahrscheinlichkeit?

Der Begriff eines **Maßes**, den wir in der folgenden Definition einführen, ist trotz (oder wegen?) seiner Einfachkeit zentral für die moderne Analysis und für die Wahrscheinlichkeitstheorie.

Definition 1.2.1 (Maße und Maßräume)**.**

(a) Sei $\mathcal{A} \subseteq 2^\Omega$ eine σ-Algebra auf einer Menge Ω. Eine Abbildung $\mu\colon \mathcal{A} \to [0,\infty]$ heißt **Maß**, falls sie folgende Eigenschaften erfüllt:

(I) Es gilt $\mu(\emptyset) = 0$.

(II) σ-*Additivität*: Für jede Folge von paarweise disjunkten[3] Mengen $A_1, A_2, \cdots \in \mathcal{A}$ gilt

$$\mu\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

(b) Ein **Maßraum** ist ein Tripel $(\Omega, \mathcal{A}, \mu)$, wobei Ω eine Menge ist, $\mathcal{A} \subseteq 2^\Omega$ eine σ-Algebra und $\mu\colon \mathcal{A} \to [0,\infty]$ ein Maß.

Im Laufe der Analysis 3 werden Sie zahlreiche verschiedene Beispiele für Maße sehen und Möglichkeiten kennenlernen um Maße mit interessanten Eigenschaften zu konstruieren. Ein paar erste Beispiele werden im Folgenden besprochen. Bereits an diesen Beispielen deutet sich an, dass Maße auch in der Wahrscheinlichkeitstheorie auftauchen.

Beispiele 1.2.2 (Einige Maßräume)**.**

(a) Sei $\Omega = \{1,\dots,6\}$ und $\mathcal{A} = 2^\Omega$. Sei $\mu\colon \mathcal{A} \to [0,\infty]$, $\mu(A) := \frac{\#A}{6}$ für alle $A \in \mathcal{A}$. Dann sieht man leicht, dass μ ein Maß ist.

Anschaulich beschreibt $\mu(A)$ die Wahrscheinlichkeit beim Werfen eines Würfels eine Augenzahl zu erhalten, die in der Menge A liegt.

(b) Sei $\Omega = \{K, Z\}$, wobei K und Z wieder für *Kopf* und *Zahl* stehen. Ein Element ω der Menge $\Omega^3 = \{K,Z\}^3$ interpretieren wir als das Ergebnis eines dreimaligen Münzwurfs. Die Abbildung $\mu\colon 2^{\Omega^3} \to [0,\infty]$, $\mu(A) := \frac{\#A}{\#\Omega^3} = \frac{\#A}{8}$ für $A \in 2^{\Omega^3}$ ist, wie man leicht überprüft, ein Maß auf der σ-Algebra 2^{Ω^3}. Anschaulich beschreibt $\mu(A)$ die Wahrscheinlichkeit, dass das Ergebnis ω des dreifachen Münzwurfs in der Menge A liegt.

Lassen Sie uns nun wie wie in Beispiel 1.1.4(a) eine kleinere σ-Algebra auf Ω^3 betrachten, nämlich die σ-Algebra

$$\mathcal{A} := \Big\{B \times \Omega^2 \mid B \subseteq \Omega\Big\}.$$

Wie in Beispiel 1.1.4(a) besprochen, besteht $\mathcal{A}$ aus genau denjenigen Mengen $A \subseteq \Omega^3$, von denen wir bereits nach dem ersten Wurf entscheiden können, ob $\omega \in A$ gilt.

Da Mengen in $\mathcal{A}$ nur Information codieren, die bereits nach dem ersten Münzwurf zur Verfügung stehen, erwartet man anschaulich, dass sich für $A \in \mathcal{A}$ die Wahrscheinlichkeit $\mu(A)$ nur davon abhängt, was A über den ersten der drei Würfe aussagt. Das ist tatsächlich so: Jedes $A \in \mathcal{A}$ lässt sich laut Definition von $\mathcal{A}$ schreiben als $A = B \times \Omega^2$ für ein $B \subseteq \Omega$. Somit gilt $\omega \in A$ genau dann, wenn $\omega_1 \in B$ ist, das heißt die komplette Information, die in der Aussage $\omega \in A$ steckt, steckt bereits in der Aussage $\omega_1 \in B$, welche nur die Menge B und das Ergebnis ω_1 des ersten

[3] Das heißt $A_j \cap A_k = \emptyset$ für alle Indizes j, k mit $j \neq k$.

Münzwurfs verwendet. Die Wahrscheinlichkeit $\mu(A)$ kann man ebenfalls mit Hilfe von B ausdrücken: Es ist

$$\mu(A) = \frac{\#A}{8} = \frac{\#(B \times \Omega^2)}{8} = \frac{(\#B) \cdot (\#(\Omega^2))}{8} = \frac{(\#B) \cdot 4}{8} = \frac{\#B}{2}.$$

(c) Sei $\mathbb{R}^d$ mit der Euklidischen Metrik[4] ausgestattet. Wir werden im Laufe der nächsten Vorlesungen zeigen, dass es genau ein Maß $\lambda_d \colon \mathcal{B}(\mathbb{R}^d) \to [0,\infty]$ gibt, welches die Eigenschaft

$$\lambda_d\big([a_1,b_1] \times \cdots \times [a_d,b_d]\big) = (b_1 - a_1) \cdots (b_d - a_d) \tag{1.2.1}$$

für alle $a_1,\dots,a_d,\, b_1,\dots,b_d \in \mathbb{R}$ mit $a_1 \le b_1,\, \dots,\, a_d \le b_d$ erfüllt.[5] Man nennt λ_d das ***d*-dimensionale Lebesgue-Maß**.

Proposition 1.2.3 (Grundlegende Eigenschaften von Maßen)**.** *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum.*

(a) *Ist $n \in \mathbb{N}$ und sind $A_1,\dots,A_n \in \mathcal{A}$ paarweise disjunkt, dann gilt*

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \sum_{k=1}^{n} \mu(A_k).$$

(b) *Für alle $A,B \in \mathcal{A}$ gilt*

$$\begin{aligned} &\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B) \\ \text{und}\quad &\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B). \end{aligned}$$

Vorlesung 3 (Mi, 22.10.) beginnt mit dem Beweis von Proposition 1.2.3(c)

(c) *Für alle $A,B \in \mathcal{A}$ mit $A \subseteq B$ ist $\mu(A) \le \mu(B)$. Gilt zusätzlich $\mu(B) < \infty$, so ist*

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

(d) *Sind $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$ Elemente von $\mathcal{A}$, dann gilt*

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \mu(A_n).$$

(e) *Sind $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$ Elemente von $\mathcal{A}$ und ist $\mu(A_1) < \infty$, dann gilt*

$$\mu\Big(\bigcap_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \mu(A_n).$$

[4] Also mit derjenigen Metrik, die von der Euklidischen Norm induziert wird.

[5] Falls Sie im Sommersemester 2025 die Analysis 2 gehört haben, haben Sie diese Behauptung dort bereits in Theorem 3.2.3 gesehen, zusammen mit dem Verweis, dass wir dieses Resultat in der Analysis 3 beweisen werden.

(f) *Für alle $A_1, A_2, A_3, \dots \in \mathcal{A}$ gilt*

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) \le \sum_{n=1}^{\infty} \mu(A_n).$$

Beweis. (a) Setze $A_k := \emptyset$ für alle $k \ge n+1$. Dann sind die Mengen $A_1, A_2, \dots$ paarweise disjunkt und es folgt

$$\mu\Big(\bigcup_{k=1}^{n} A_k\Big) = \mu\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \mu(A_k) = \sum_{k=1}^{n} \mu(A_k);$$

für die zweite Gleichheit haben wir die σ-Additivität von μ (Eigenschaft (II) in Definition 1.2.1(a)) verwendet und für die dritte Gleichheit die Eigenschaft $\mu(\emptyset) = 0$ (Eigenschaft (I) in Definition 1.2.1(a)).

(b) Seien $A, B \in \mathcal{A}$. Dann ist $A \cup B$ die Vereinigung der drei paarweise disjunkten Mengen $A \setminus B$, $B \setminus A$ und $A \cap B$, also folgt wegen (a) die erste Gleichung. Indem man zur ersten Gleichung $\mu(A \cap B)$ hinzu addiert, erhält man außerdem

$$\mu(A \cup B) + \mu(A \cap B) = \underbrace{\mu(A \setminus B) + \mu(A \cap B)}_{=\mu(A)} + \underbrace{\mu(B \setminus A) + \mu(A \cap B)}_{=\mu(B)} = \mu(A) + \mu(B);$$

hierbei haben wir verwendet, dass A die disjunkte Vereinigung von $A \setminus B$ und $A \cap B$ ist und dass B die disjunkte Vereinigung von $B \setminus A$ und $A \cap B$ ist.

(c) Seien $A, B \in \mathcal{A}$ mit $A \subseteq B$. Dann ist B die disjunkte Vereinigung von $B \setminus A$ und A, also gilt $\mu(B) = \mu(B \setminus A) + \mu(A) \ge \mu(A)$.

Wenn zusätzlich $\mu(B) < \infty$ ist, dann folgt aus gerade gezeigten Ungleichung, dass auch $\mu(B \setminus A)$ und $\mu(A)$ endlich sind. Wir können dann in der Gleichung $\mu(B) = \mu(B \setminus A) + \mu(A)$ die Zahl $\mu(A)$ subtrahieren und erhalten somit wie behautet $\mu(B \setminus A) = \mu(B) - \mu(A)$.

(d) Definiere eine Folge $B_1, B_2, \dots$ in $\mathcal{A}$ durch $B_1 := A_1$ und $B_n := A_n \setminus A_{n-1}$ für alle $n \ge 2$. Dann sind die Mengen $B_1, B_2, \dots$ paarweise disjunkt und es gilt $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$. Somit gilt

$$\begin{aligned}\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} B_n\Big) &= \sum_{n=1}^{\infty} \mu(B_n) \\ &= \lim_{N\to\infty} \sum_{n=1}^{N} \mu(B_n) = \lim_{N\to\infty} \mu\Big(\bigcup_{n=1}^{N} B_n\Big) = \lim_{N\to\infty} \mu(A_N);\end{aligned}$$

für die zweite Gleichheit haben wir die σ-Additivität von μ benutzt und für die vierte Gleichheit die bereits bewiesene Eigenschaft (a).

(e) Es gilt

$$\mu(A_1) - \mu\Big(\bigcap_{n=1}^{\infty} A_n\Big) = \mu\Big(A_1 \setminus \bigcap_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} (A_1 \setminus A_n)\Big)$$

$$\overset{(d)}{=} \lim_{N\to\infty} \mu(A_1 \setminus A_N) = \lim_{N\to\infty} (\mu(A_1) - \mu(A_N));$$

für die erste und die vierte Gleichheit haben wir die Annahme $\mu(A_1) < \infty$ und die Aussage (c) verwendet. Also gilt $\mu(A_1) - \mu(A_N) \to \mu(A_1) - \mu\big(\bigcap_{n=1}^{\infty} A_n\big)$ für $N \to \infty$. Subtraktion von $\mu(A_1)$ (hier verwenden wir erneut $\mu(A_1) < \infty$) und anschließende Multiplikation mit -1 liefert $\mu(A_N) \to \mu\big(\bigcap_{n=1}^{\infty} A_n\big)$ für $N \to \infty$.

(f) Wir definieren die Mengen $C_n := A_n \setminus \bigcup_{k=1}^{n-1} A_k \in \mathcal{A}$ für alle $n \in \mathbb{N}$. Diese Mengen sind paarweise disjunkt und erfüllen $\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n$, also folgt aus der σ-Additivität von μ

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \mu\Big(\bigcup_{n=1}^{\infty} C_n\Big) = \sum_{n=1}^{\infty} \mu(C_n) \le \sum_{n=1}^{\infty} \mu(A_n),$$

wobei die Ungleichung am Ende wegen $C_n \subseteq A_n$ aus (c) folgt. □

Definition 1.2.4 (Spezielle Eigenschaften von Maßen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

(a) Man nennt das Maß μ **endlich**, falls $\mu(\Omega) < \infty$ gilt. In diesem Fall nennt man $(\Omega, \mathcal{A}, \mu)$ einen **endlichen Maßraum**.[6]

(b) Man nennt das Maß μ ein **Wahrscheinlichkeitsmaß**, falls $\mu(\Omega) = 1$ gilt. In diesem Fall nennt man den Maßraum $(\Omega, \mathcal{A}, \mu)$ einen **Wahrscheinlichkeitsraum**.

(c) Man nennt das Maß μ ***σ*-endlich**, falls es eine Folge von Mengen $A_1, A_2, \ldots$ in $\mathcal{A}$ gibt, die $\bigcup_{n\in\mathbb{N}} A_n = \Omega$ und $\mu(A_n) < \infty$ für alle $n \in \mathbb{N}$ erfüllen. In diesem Fall nennt $(\Omega, \mathcal{A}, \mu)$ einen ***σ*-endlichen Maßraum**.

Beispiele 1.2.5 (Eigenschaften der Maße aus den Beispielen 1.2.2)**.**

(a) Das Maß μ auf $\{1, \ldots, 6\}$ aus Beispiel 1.2.2(a) war durch $\mu(A) := \frac{\#A}{6}$ für alle $A \in 2^{\{1,\ldots,6\}}$ definiert. Da $\mu(\{1, \ldots, 6\}) = \frac{6}{6} = 1$ gilt, ist μ ein Wahrscheinlichkeitsmaß und somit insbesondere endlich.

(b) Das Maß μ auf $2^{\{K,Z\}^3}$ aus Beispiel 1.2.2(b) war durch $\mu(A) := \frac{\#A}{8}$ für alle $A \in 2^{\{K,Z\}^3}$ definiert. Wegen $\mu\big(\{K, Z\}^3\big) = \frac{8}{8} = 1$ ist μ ein Wahrscheinlichkeitsmaß und somit endlich.

(c) Betrachten wir das Lebesgue-Maß λ_d auf $\mathcal{B}(\mathbb{R}^d)$, dass wir in Beispiel 1.2.2(c) angesprochen haben. Es ist σ-endlich, denn es gilt

$$\mathbb{R}^d = \bigcup_{n\in\mathbb{N}} [-n, n]^d \qquad \text{und} \qquad \lambda_d([-n, n]^d) = (2n)^d < \infty \quad \text{für alle } n \in \mathbb{N}.$$

Allerdings ist λ_d nicht endlich, denn aus Proposition 1.2.3(d) folgt

$$\lambda_d(\mathbb{R}^d) = \lim_{n\to\infty} \lambda_d([-n, n]^d) = \lim_{n\to\infty} (2n)^d = \infty.$$

[6] Beachten Sie unbedingt, dass sich das Wort *endlich* im Begriff *endlicher Maßraum* laut dieser Definition auf die Endlichkeit des Maßes μ bezieht. Der Begriff bedeutet nicht, dass die Menge Ω endlich sein muss!

Im Rest dieses Abschnitts besprechen wir eine Technik, mit der man zeigen kann, dass zwei Maße gleich sind. Das ist zum Beispiel nützlich um die Eindeutigkeitsaussage über das Lebesgue-Maß in Beispiel 1.2.2(c) zu beweisen. Die Technik basiert auf dem folgenden Begriff.

Definition 1.2.6 (Dynkin-Systeme)**.** Sei Ω eine Menge. Eine Teilmenge $\mathcal{D} \subseteq 2^\Omega$ nennt man ein **Dynkin-System** oder genauer ein **Dynkin-System auf Ω**, falls es die folgenden Axiome erfüllt:

(I) Es ist $\emptyset \in \mathcal{D}$.

(II) Für jedes $A \in \mathcal{D}$ gilt auch $A^\complement \in \mathcal{D}$.

(III) Für jede Folge von paarweise disjunkten Mengen $A_1, A_2, A_3, \dots \in \mathcal{D}$ ist $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{D}$.

Beachten Sie, dass die Definition eines Dynkin-Systems sich von der Definition einer σ-Algebra (Definition 1.1.2) nur dadurch unterscheidet, dass die Bedingung (III) für Dynkin-Systeme nur für paarweise disjunkte Mengen gefordert wird. Ein Dynkin-System zu sein ist also eine etwas schwächere Eigenschaft als eine σ-Algebra zu sein. Es sollte nicht allzu überraschend sein, dass Dynkin-Systeme auftauchen, wenn man mit Maßen arbeitet: Dass man in Bedingung (III) der Definition von Dykin-Systemen nur paarweise disjunkte Mengen betrachtet, passt sehr gut zur σ-Additivität von Maßen (Definition 1.2.1(a)(II)), in der ebenfalls nur paarweise disjunkte Mengen auftauchen.

Analog zur erzeugten σ-Algebra in Definition 1.1.8(a) definieren wir auch das von einem Mengensystem erzeugte Dynkin-System:

Definition 1.2.7 (Erzeugtes Dynkin-System)**.** Sei Ω eine Menge und $\mathcal{E} \subseteq 2^\Omega$. Das Mengensystem

$$\delta(\mathcal{E}) := \bigcap_{\substack{\mathcal{D}\supseteq\mathcal{E} \\ \mathcal{D} \text{ Dynkin-System auf } \Omega}} \mathcal{D} \subseteq 2^\Omega$$

nennt man das von **von $\mathcal{E}$ erzeugte Dynkin-System** oder genauer das **von $\mathcal{E}$ erzeugte Dynkin-System auf Ω**.

Die Definition der von einem Mengensystem erzeugten σ-Algebra war durch die Eigenschaften in Proposition 1.1.7 motiviert gewesen. Dieselben Aussagen wie in dieser Proposition kann man völlig analog auch für Dynkin-System beweisen. Inbesondere erhält man somit:

Proposition 1.2.8 (Kleinstes Dynkin-System)**.** *Sei Ω eine Menge und $\mathcal{E} \subseteq 2^\Omega$. Dann ist $\delta(\mathcal{E})$ das kleinste Dynkin-System auf Ω, das $\mathcal{E}$ enthält.*

Der Witz an Dynkin-Systemen ist, dass sie manchmal einfacher handhabbar sind als σ-Algebren, dass in vielen Fällen aber die von einem Mengensystem $\mathcal{E}$ erzeugte σ-Algebra mit dem von $\mathcal{E}$ erzeugten Dynkin-System übereinstimmt. Um das zu präzisieren benötigen wir den folgenden Begriff.

Definition 1.2.9 (Stabilität bezüglich endlicher Durschnitte)**.** Sei Ω eine Menge. Eine Mengensystem $\mathcal{E}$ heißt **stabil bezüglich endlicher Durchschnitte**, falls für alle $A, B \in \mathcal{E}$ auch $A \cap B \in \mathcal{E}$ gilt.

Wenn $\mathcal{E} \subseteq 2^\Omega$ stabil bezüglich endlicher Durchschnitte ist, dann sieht man induktiv sofort, dass $A_1 \cap \cdots \cap A_n \in \mathcal{E}$ für alle $n \in \mathbb{N}$ und alle $A_1, \ldots, A_n \in \mathcal{E}$ gilt. Das erklärt, weshalb man den Begriff *stabil bezüglich endlicher Durchschnitte* verwendet.

Proposition 1.2.10 ($\boldsymbol{\sigma}$-Algebren vs. Dynkin-Systeme)**.** *Sei Ω eine Menge und $\mathcal{D} \subseteq 2^\Omega$ ein Dynkin-System. Es ist $\mathcal{D}$ genau dann eine σ-Algebra, wenn $\mathcal{D}$ stabil bezüglich endlicher Durchschnitte ist.*

Beweis. Der Beweis ist nicht schwer, aber eine gute Möglichkeit um die Axiome und Eigenschaften von σ-Algebren einzuüben. Wir lagern ihn deshalb in die Übungen aus. □

Wenn man zeigen will, dass ein gegebenes Dynkin-System stabil bezüglich endlicher Durchschnitte (und somit eine σ-Algebra) ist, kann das folgende Lemma manchmal hilfreich sein.

Lemma 1.2.11 (Durchschnitte mit Komplementen in Dynkin-Systemen)**.** *Sei Ω eine Menge, sei $\mathcal{D} \subseteq 2^\Omega$ ein Dynkin-System und seien $C, D \in \mathcal{D}$. Wenn $C \cap D \in \mathcal{D}$ ist, dann ist auch $C^\complement \cap D \in \mathcal{D}$.*

Beweis. Es gilt $C^\complement \cap D = (C \cap D)^\complement \cap D$ und somit $\left(C^\complement \cap D\right)^\complement = (C \cap D) \cup D^\complement$. Also ist $\left(C^\complement \cap D\right)^\complement$ die disjunkte Vereinigung von zwei Elementen von $\mathcal{D}$ und somit selbst ein Element von $\mathcal{D}$. Somit ist auch $C^\complement \cap D \in \mathcal{D}$. □

Theorem 1.2.12 (Erzeugte $\boldsymbol{\sigma}$-Algebra vs. erzeugtes Dynkin-System)**.** *Sei Ω eine Menge und sei $\mathcal{E} \subseteq 2^\Omega$ stabil bezüglich endlicher Durchschnitte. Dann gilt $\sigma(\mathcal{E}) = \delta(\mathcal{E})$.*

Beweis. Wir zeigen beide Inklusionen einzeln.

$\supseteq$ Da jede σ-Algebra auch ein Dynkin-System ist, gilt: Mindestens über jedes Mengensystem, über das bei der Definition von $\sigma(\mathcal{E})$ geschnitten wird, wird auch bei der Definition von $\delta(\mathcal{E})$ geschnitten. Somit ist $\sigma(\mathcal{E}) \supseteq \delta(\mathcal{E})$.

$\subseteq$ Es genügt zu zeigen, dass $\delta(\mathcal{E})$ eine σ-Algebra ist; da $\sigma(\mathcal{E})$ die kleinste σ-Algebra ist, die $\mathcal{E}$ enthält, folgt dann $\sigma(\mathcal{E}) \subseteq \delta(\mathcal{E})$.

Um zu zeigen, dass $\delta(\mathcal{E})$ eine σ-Algebra ist, genügt es laut Proposition 1.2.10 zu zeigen, dass $\delta(\mathcal{E})$ stabil bezüglich endlicher Durchschnitte ist. Hier fixieren gehen wir in zwei Schritten vor:

Schritt 1: Wir zeigen für alle $D \in \delta(\mathcal{E})$ und alle $E \in \mathcal{E}$, dass $D \cap E \in \delta(\mathcal{E})$ gilt. Hierzu fixieren wir ein beliebiges, festes $E \in \mathcal{E}$ und betrachten das Mengensystem

$$\mathcal{D} \coloneqq \left\{D \in \delta(\mathcal{E}) \mid D \cap E \in \delta(\mathcal{E})\right\}.$$

Man kann direkt nachrechnen, dass $\mathcal{D}$ ein Dynkin-System ist; die Abgeschlossenheit bezüglich Komplementbildung (Axiom 1.2.6 in Definition (II)) folgt hierbei aus Lemma 1.2.11. Da $\mathcal{E}$ laut Voraussetzung stabil bezüglich endlicher Durchschnitte ist, gilt außerdem $\mathcal{E} \subseteq \mathcal{D}$.

Weil $\delta(\mathcal{E})$ das kleinsten Dynkin-System auf Ω ist, dass $\mathcal{E}$ enthält (Proposition 1.2.8), folgt $\delta(\mathcal{E}) \subseteq \mathcal{D}$. Also gilt für jedes $D \in \delta(\mathcal{E})$, dass $D \in \mathcal{D}$ und somit $D \cap E \in \delta(E)$ ist.

Schritt 2: Nun zeigen wir, dass $\delta(\mathcal{E})$ tatsächlich stabil bezüglich endlicher Durchschnitte ist, womit dann die Behauptung gezeigt ist. Hierzu gehen wir ähnlich vor wie in Schritt 1: Wir fixieren ein beliebiges, festes $D \in \delta(\mathcal{E})$ und betrachten das Mengensystem

$$\mathcal{C} \coloneqq \big\{C \in \delta(\mathcal{E}) \mid C \cap D \in \delta(\mathcal{E})\big\}.$$

Erneut kann man direkt nachrechnen, dass $\mathcal{C}$ ein Dynkin-System auf Ω ist, wobei die Abgeschlossenheit bezüglich der Komplementbildung wieder aus Lemma 1.2.11 folgt. Wegen Schritt 1 ist $\mathcal{E} \subseteq \mathcal{C}$. Somit ist $\delta(\mathcal{E}) \subseteq \mathcal{C}$, das heißt für alle $C \in \delta(\mathcal{E})$ gilt wie behauptet $C \cap D \in \delta(\mathcal{E})$. □

Vorlesung 4 (Fr, 24.10.) beginnt mit Theorem 1.2.13

Theorem 1.2.13 (Eindeutigkeitssatz für Maße)**.** *Sei $\mathcal{A} \subseteq 2^\Omega$ eine σ-Algebra auf einer Menge Ω, sei $\mathcal{E} \subseteq \mathcal{A}$ ein Erzeuger von $\mathcal{A}$ und seien $\mu, \nu \colon \mathcal{A} \to [0,\infty]$ zwei Maße. Es seien folgende Eigenschaften erfüllt:*

(1) *Der Erzeuger $\mathcal{E}$ ist stabil bezüglich endlicher Durchschnitte.*

(2) *Für alle $E \in \mathcal{E}$ ist $\mu(E) = \nu(E)$.*

(3) *Es gibt eine Folge $E_1 \subseteq E_2 \subseteq \dots$ in $\mathcal{E}$ mit der Eigenschaft $\bigcup_{n\in\mathbb{N}} E_n = \Omega$ und $\mu(E_n) < \infty$ für alle $n \in \mathbb{N}$.*

Dann gilt $\mu = \nu$.

Man kann zeigen, dass das Theorem auch dann noch gültig bleibt, wenn man in Annahme (3) nicht voraussetzt, dass die Mengen $E_1, E_2, \dots$, sukzessive ineinander enthalten sind, weglässt. Das ist nicht viel schwieriger, macht den Beweis aber am Ende ein wenig technischer. Weil die Aussage so, wie wir Sie oben formuliert haben, für viele Zwecke bereits ausreicht, verzichten wir darauf die allgemeinere Variante in der Vorlesung zu besprechen. Wenn Sie sich für den Beweis der allgemeineren Variante interessieren, finden Sie ihn zum Beispiel in [1, Theorem II.5.6 auf Seite 64].

Vor dem Beweis von Theorem 1.2.13 ist es sinnvoll etwas Intuition für seine Aussage zu gewinnen, indem wir damit die Eindeutigkeit das Lebesgue-Maßes zeigen.

Beispiel 1.2.14 (Eindeutigkeit des Lebesgue-Maßes)**.** Sie $\mathbb{R}^d$ mit der Euklidischen Metrik versehen. Es gibt höchstens ein Maß $\lambda_d \colon \mathcal{B}(\mathbb{R}^d) \to [0,\infty]$, welches auf abgeschlossenen Quadern die Formel (1.2.1) aus Beispiel 1.2.2(c) erfüllt.

Beweis. Seien $\lambda_d, \tilde{\lambda}_d \colon \mathcal{B}(\mathbb{R}^d) \to [0,\infty]$ zwei Maße, die die Formel (1.2.1) erfüllen. Sei $\mathcal{E} \subseteq 2^{\mathbb{R}^d}$ die Menge aller Quader der Form

$$[a_1, b_1] \times \cdots \times [a_d, b_d] \subseteq \mathbb{R}^d$$

für $a_1, \ldots, a_d, b_1, \ldots, b_d \in \mathbb{R}$ mit $a_1 \leq b_1, \ldots, a_d \leq b_d$. In Aufgabe 2.4(a) auf Hausaufgabeblatt 2 haben Sie gezeigt, dass sich jede nichtleere offene Menge in $\mathbb{R}^d$ als Vereinigung von abzählbaren vielen solcher Quader schreiben lässt und erhalten somit $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{E})$. Außerdem ist $\mathcal{E}$ offenbar stabil bezüglich endlicher Durchschnitte. Die Maße λ_d und $\tilde{\lambda}_d$ erfüllen die beiden Annahmen (2) und (3) in Theorem 1.2.13:

(2) Für jedes $E \in \mathcal{E}$ sind $\lambda_d(E)$ und $\tilde{\lambda}_d(E)$ laut Voraussetzung durch die Formel (1.2.1) gegeben, also ist $\lambda_d(E) = \tilde{\lambda}_d(E)$.

(3) Setze $E_n \coloneqq [-n, n]^d$ für jedes $n \in \mathbb{N}$. Dann gilt $E_1 \subseteq E_2 \subseteq \ldots$ und $\bigcup_{n\in\mathbb{N}} E_n = \mathbb{R}^d$. Für jedes $n \in \mathbb{N}$ ist $E_n \in \mathcal{E}$ und laut Formel (1.2.1) $\lambda_d(E_n) = (2n)^d < \infty$.

Also sind alle Annahmen aus Theorem 1.2.13 erfüllt und das Theorem impliziert somit $\lambda_d = \tilde{\lambda}_d$. □

Beweis von Theorem 1.2.13. Betrachten wir zunächst eine feste Menge $E \in \mathcal{E}$ mit $\mu(E) < \infty$. Wir setzen

$$\mathcal{D}_E \coloneqq \big\{A \in \mathcal{A} \mid \mu(A \cap E) = \nu(A \cap E)\big\}.$$

Lassen Sie uns zeigen, dass $\mathcal{D}_E$ ein Dynkin-System auf Ω ist.

Axiom (I)*:* Es gilt $\emptyset \in \mathcal{A}$ und $\mu(\emptyset \cap E) = 0 = \nu(\emptyset \cap E)$, also ist $\emptyset \in \mathcal{D}_E$.

Axiom (II)*:* Sei $A \in \mathcal{D}_E$. Dann ist wegen $A \in \mathcal{A}$ auch $A^\complement \in \mathcal{A}$. Wir verwenden nun die Mengengleichheit $E \cap A^\complement = E \setminus (E \cap A)$. Wegen $\mu(E) = \nu(E) < \infty$ folgt aus dieser Mengengleichheit laut Proposition 1.2.3(c)

$$\mu(E \cap A^\complement) = \mu(E) - \mu(E \cap A) = \nu(E) - \nu(E \cap A) = \nu(E \cap A^\complement),$$

wobei wir für die zweite Gleichheit verwendet haben, dass $E \cap A \in \mathcal{D}_E$ gilt. Also ist tatsächlich $A^\complement \in \mathcal{D}_E$.

Axiom (III)*:* Seien $A_1, A_2, \cdots \in \mathcal{D}_E$ paarweise disjunkt. Dann sind auch die Mengen $A_1 \cap E, A_2 \cap E, \ldots$ paarweise disjunkt, also folgt aus der σ-Additivität von Maßen

$$\begin{aligned}\mu\big(E \cap \bigcup_{k\in\mathbb{N}} A_k\big) = \mu\big(\bigcup_{k\in\mathbb{N}} E \cap A_k\big) &= \sum_{k=1}^{\infty} \mu(E \cap A_k) \\ &= \sum_{k=1}^{\infty} \nu(E \cap A_k) = \nu\big(\bigcup_{k\in\mathbb{N}} E \cap A_k\big) = \nu\big(E \cap \bigcup_{k\in\mathbb{N}} A_k\big);\end{aligned}$$

für die Gleichheit zwischen den beiden Zeilen haben wir verwendet, dass jede der Mengen A_k ein Element von $\mathcal{D}_E$ ist. Wir haben $\bigcup_{k\in\mathbb{N}} A_k \in \mathcal{D}_E$ gezeigt.

Also ist $\mathcal{D}_E$ tatsächlich ein Dynkin-System. Weil $\mathcal{E}$ laut Annahme (1) stabil bezüglich endlicher Durchschnitte ist, folgt aus Annahme (2), dass $\mathcal{E} \subseteq \mathcal{D}_E$ ist und somit gilt $\delta(\mathcal{E}) \subseteq \mathcal{D}_E$. Also haben wir $\mu(A \cap E) = \nu(A \cap E)$ für alle $A \in \delta(\mathcal{E})$ gezeigt.

Da $\mathcal{E}$ stabil bezüglich endlicher Durchschnitte ist (Annahme (1)), folgt aus Theorem 1.2.12 $\delta(\mathcal{E}) = \sigma(\mathcal{E}) = \mathcal{A}$; das heißt, wir wissen nun, dass $\mu(A \cap E) = \nu(A \cap E)$ für alle $A \in \delta(\mathcal{E})$ für alle $A \in \mathcal{A}$ ist. Da $E \in \mathcal{E}$ mit $\mu(E) < \infty$ beliebig war, gilt dies für all solche E.

Wir zeigen nun mit Hilfe von Annahme (3), dass hieraus die Behauptung folgt. Für jedes $A \in \mathcal{A}$ gilt $A = \bigcup_{n\in\mathbb{N}}(A \cap E_n)$. Wegen $A \cap E_1 \subseteq A \cap E_2 \subseteq \dots$ folgt somit aus Proposition 1.2.3(d)

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap E_n) = \lim_{n\to\infty} \nu(A \cap E_n) = \nu(A),$$

womit die Behauptung gezeigt ist. □

1.3 Konstruktion von Maßen

Einstiegsfragen.

(a) Woher weiß man, dass das Lebesgue-Maß (Beispiel 1.2.2(c)) existiert?

(b) Wie würden Sie versuchen die "Länge" einer Menge $M \subseteq \mathbb{R}$ "von außen" zu approximieren?

Für viele interessante Maße, zum Beispiel für das Lebesgue-Maß, kann man die Existenz nicht einfach direkt zeigen, sondern man benötigt eine Reihe vorbereitender Schritte. Der erste Schritt ist es oft, zuerst eine Funktion zu bauen, die auf der kompletten Potenzmenge 2^Ω der gegebenen Menge Ω definiert ist, dafür aber etwas schwächere Eigenschaften hat als ein Maß:

Definition 1.3.1 (Äußere Maße)**.** Sei Ω eine Menge. Eine Funktion $\mu^* : 2^\Omega \to [0,\infty]$ nennt man ein **äußeres Maß**, falls sie die folgenden drei Axiome erfüllt:

(I) Es gilt $\mu^*(\emptyset) = 0$.

(II) σ*-Subadditivität*: Für jede Folge von Mengen $A_1, A_2, \dots \in 2^\Omega$ gilt

$$\mu^*\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

(III) Für alle $A, B \in 2^\Omega$ mit $A \subseteq B$ gilt $\mu^*(A) \leq \mu^*(B)$.

Der Vorteil von äußeren Maßen ist, dass man relativ einfach welche konstruieren kann. Hier ist ein wichtiges Beispiel, dass wir anschließend verwenden werden um daraus das Lebesgue-Maß auf $\mathbb{R}$ zu konstruieren.

Beispiel 1.3.2 (Das Lebesguesche äußere Maß auf $\mathbb{R}$)**.** Für jedes $A \in 2^{\mathbb{R}}$ sei $\lambda_1^*(A) \in [0,\infty]$ definiert durch

$$\lambda_1^*(A) := \inf\Big\{ \sum_{n=1}^{\infty} (b_n - a_n) \mid a_n, b_n \in \mathbb{R},\ a_n \le b_n,\ \bigcup_{n\in\mathbb{N}} [a_n, b_n) \supseteq A \Big\}.$$

Dann ist $\lambda_1^* \colon 2^{\mathbb{R}} \to [0,\infty]$ ein äußeres Maß auf $\mathbb{R}$, das sogenannte **ein-dimensionale Lebesguesche äußere Maß**. Für alle $a, b \in \mathbb{R}$ mit $a \le b$ gilt $\lambda_1^*([a,b)) = b - a$.

Beweis. Wir prüfen zuerst für λ_1^* die drei Axiome aus Definition 1.3.1 nach.

(I) Wir wählen $a_n := b_n := 0$ für alle $n \in \mathbb{N}$. Dann gilt $\emptyset = [0,0) = \bigcup_{n\in\mathbb{N}} [0,0) = \bigcup_{n\in\mathbb{N}} [a_n, b_n)$ und $\sum_{n=1}^{\infty}(b_n - a_n) = 0$. Also ist $\lambda_1^*(\emptyset) \le 0$ und somit $\lambda_1^*(\emptyset) = 0$.

(II) Seien $A_1, A_2, \dots \in 2^{\mathbb{R}}$. Wenn $\lambda_1^*(A_n) = \infty$ für mindestens einen Index n gilt, dann ist die behauptete Ungleichung klar. Also sei nun $\lambda_1^*(A_n) < \infty$ für alle $n \in \mathbb{N}$.

Sei $\varepsilon > 0$ beliebig, fest. Für jedes $n \in \mathbb{N}$ gibt es Zahlen $a_1^{(n)} \le b_1^{(n)}$, $a_2^{(n)} \le b_2^{(n)}$, …, mit $A_n \subseteq \bigcup_{k=1}^{\infty} [a_k^{(n)}, b_k^{(n)})$ und $\sum_{k=1}^{\infty}(b_k^{(n)} - a_k^{(n)}) \le \lambda_1^*(A_n) + \frac{\varepsilon}{2^n}$. Wenn wir alle diese halboffenen Intervalle $[a_k^{(n)}, b_k^{(n)})$ für alle $n, k \in \mathbb{N}$ zusammennehmen, dann sind es noch immer abzählbar viele (da $\mathbb{N}\times\mathbb{N}$ abzählbar ist) und sie überdecken gemeinsam die Menge $\bigcup_{n\in\mathbb{N}} A_n$. Außerdem ist

$$\lambda_1^*\Big(\bigcup_{n\in\mathbb{N}} A_n\Big) \le \sum_{k,n\in\mathbb{N}} (b_k^{(n)} - a_k^{(n)}) \le \sum_{n=1}^{\infty} \Big(\lambda_1^*(A_n) + \frac{\varepsilon}{2^n}\Big) = \sum_{n=1}^{\infty} \lambda_1^*(A_n) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt $\lambda_1^*\big(\bigcup_{n\in\mathbb{N}} A_n\big) \le \sum_{n\in\mathbb{N}} \lambda_1^*(A_n)$.

(III) Seien $AB \in \mathbb{R}$ mit $A \subseteq B$. Dann ist jede Überdeckung von B durch Intervalle der Form $[a_n, b_n)$ auch eine Überdeckung von A. Also ist die Menge, deren Infimum in der Definition von $\lambda_1^*(B)$ gebildet wird, eine Teilmenge der Menge, die in der Definition von $\lambda_1^*(A)$ gebildet wird. Weil kleinere Mengen größere Infima haben, folgt $\lambda_1^*(A) \le \lambda_1^*(B)$.

Vorlesung 5 (Mi, 29.10.) beginnt mit dem Beweis von $\lambda_1^*([a,b)) = b - a$ in Beispiel 1.3.2

Seien nun $a, b \in \mathbb{R}$ mit $a \le b$ fest. Wir müssen noch die Formel $\lambda_1^*([a,b)) = b - a$ zeigen.

"$\le$" Wegen $[a,b) \subseteq [a,b) \cup \bigcup_{n=2}^{\infty} [b,b)$ gilt $\lambda_1^*([a,b)) \le (b-a) + \sum_{n=2}^{\infty}(b-b) = b - a$.

"$\ge$" Seien $a_1 \le b_1$, $a_2 \le b_2$, … reelle Zahlen $[a,b) \subseteq \bigcup_{n=1}^{\infty} [a_n, b_n)$. Wir müssen $b - a \le \sum_{n=1}^{\infty}(b_n - a_n)$ zeigen.

Sei dazu $\varepsilon > 0$ beliebig, fest. Wir fügen noch ein weiteres halboffenes Intervall hinzu, indem wir $a_0 := b$ and $b_0 := b + \varepsilon$ setzen. Dann gilt

$$[a,b] \subseteq \bigcup_{n=0}^{\infty} [a_n, b_n) \subseteq \bigcup_{n=0}^{\infty} (a_n - \frac{\varepsilon}{2^n}, b_n).$$

Weil $[a,b]$ kompakt ist und die Intervalle $(a_n - \frac{\varepsilon}{2^n}, b_n)$ offen sind, finden wir endlich viele verschiedene Indizes $n_1, \dots, n_\ell \in \mathbb{N}_0$ mit

$$[a,b] \subseteq \bigcup_{k=1}^{\ell} (a_{n_k} - \frac{\varepsilon}{2^{n_k}}, b_{n_k})$$

Laut Aufgabe 3.3 auf Hausaufgabenblatt 3 folgt daraus

$$b-a \leq \sum_{k=1}^{\ell} \left(b_{n_k} - \left(a_{n_k} - \frac{\varepsilon}{2^{n_k}}\right)\right) \leq \sum_{n=0}^{\infty} (b_n - a_n + \frac{\varepsilon}{2^n})$$
$$\sum_{n=0}^{\infty} (b_n - a_n) + 2\varepsilon = \sum_{n=1}^{\infty} (b_n - a_n) + 3\varepsilon,$$

wobei wir im vorletzten Schritt die geometrische Summenformel $\sum_{n=0}^{\infty} \frac{1}{2^n} = 2$ und im letzten Schritt die Eigenschaft $b_0 - a_0 = \varepsilon$ verwendet haben. Da $\varepsilon > 0$ beliebig war, folgt die behauptete Ungleichung. □

Äußere Maße sind zwar im Allgemeinen keine Maße, aber man kann aus einem äußeren Maß μ^* stets ein Maß μ konstruieren, indem man μ^* auf eine geeignete σ-Algebra $\mathcal{A}_{\mu^*} \subseteq 2^\Omega$ einschränkt. Das ist der Inhalt des folgenden sehr nützlichen Resultats.

Theorem 1.3.3 (Von äußeren Maßen zu Maßen)**.** *Sei Ω eine Menge und $\mu^* : 2^\Omega \to [0,\infty]$ ein äußeres Maß. Wir setzen*

$$\mathcal{A}_{\mu^*} := \left\{A \in 2^\Omega \mid \mu^*(C \cap A) + \mu^*(C \cap A^\complement) = \mu^*(C) \text{ für alle } C \in 2^\Omega\right\}.$$

Dann ist $\mathcal{A}_{\mu^*}$ eine σ-Algebra auf Ω und die Einschränkung $\mu^*|_{\mathcal{A}_{\mu^*}} : \mathcal{A}_{\mu^*} \to [0,\infty]$ ein Maß.

Beweis. Der einfacheren Notation halber setzen wir für den Beweis $\mathcal{A} := \mathcal{A}_{\mu^*}$.

Beweisschritt 1: Zunächst beobachten wir, dass eine Menge $A \in 2^\Omega$ genau dann ein Element von $\mathcal{A}$ ist, wenn für alle $C \in 2^\Omega$ die Ungleichung

$$\mu^*(C \cap A) + \mu^*(C \cap A^\complement) \leq \mu^*(C) \tag{1.3.1}$$

gilt, denn die umgekehrte Ungleichung gilt wegen der σ-Subadditivität von äußeren Maßen (Definition 1.3.1(II)) für alle $A, C \in 2^\Omega$.

Beweisschritt 2: Als nächstes zeigen wir, dass $\mathcal{A}$ stabil bezüglich endlicher Durchschnitte ist. Seien dazu $A, B \in \mathcal{A}$ und $C \in 2^\Omega$. Wegen

$$C \cap (A \cap B)^\complement = (C \cap A^\complement) \cup (C \cap B^\complement) = (C \cap A^\complement \cap B) \cup (C \cap B^\complement)$$

und wegen der Subadditivität von äußeren Maßen gilt

$$\mu^*(C \cap (A \cap B)) + \mu^*(C \cap (A \cap B)^\complement) \leq \mu^*(C \cap A \cap B) + \mu^*(C \cap A^\complement \cap B) + \mu^*(C \cap B^\complement)$$
$$= \mu^*(C \cap B) + \mu^*(C \cap B^\complement) = \mu^*(C);$$

Dabei haben wir im ersten Schritt die Subadditivität von μ^* verwendet, im zweiten Schritt die Eigenschaft $A \in \mathcal{A}$ (allerdings nicht mit der Menge $C \in 2^\Omega$, sondern mit der Menge $C \cap B \in 2^\Omega$), und im dritten Schritt die Eigenschaft $B \in \mathcal{A}$. Also ist die Ungleichung (1.3.1) (für die Menge $A \cap B$ anstelle der Menge A) erfüllt, das heißt, es gilt tatsächlich $A \cap B \in \mathcal{A}$.

Beweisschritt 3: Nun zeigen wir, dass $\mathcal{A}$ ein Dynkin-System ist. Da wir aus Beweisschritt 2 wissen, dass $\mathcal{A}$ stabil bezüglich endlicher Durchschnitte ist, folgt dann aus Proposition 1.2.10, dass $\mathcal{A}$ sogar eine σ-Algebra ist. Wir prüfen für $\mathcal{A}$ die drei Axiome aus der Definition von Dynkin-Systemen (Definition 1.2.6) nach:

(I) Sie $C \in 2^\Omega$. Dann gilt

$$\mu^*(C \cap \emptyset) + \mu^*(C \cap \emptyset^\complement) = \mu^*(\emptyset) + \mu^*(C) = \mu^*(C),$$

wobei wir im letzten Schritt verwendet haben, dass äußere Maße die leere Menge auf 0 abbilden. Also ist $\emptyset \in \mathcal{A}$.

(II) Sei $A \in \mathcal{A}$ und sei $C \in 2^\Omega$. Dann gilt

$$\mu^*\left(C \cap A^\complement\right) + \mu^*\left(C \cap (A^\complement)^\complement\right) = \mu^*\left(C \cap A^\complement\right) + \mu^*(C \cap A) = \mu^*(C),$$

wobei wir im letzten Schritt die Eigenschaft $A \in \mathcal{A}$ benutzt haben. Es folgt $A^\complement \in \mathcal{A}$.

(III) Seien $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt. Sei außerdem $C \in 2^\Omega$. Da A_1 zu allen anderen A_n disjunkt ist, gilt $A_1 \subseteq \bigcap_{n=2}^\infty A_n^\complement$. Daraus erhalten wir

$$\begin{aligned}\mu^*\left(C \cap A_1\right) + \mu^*\left(C \cap \bigcap_{n=1}^\infty A_n^\complement\right) &= \mu^*\left(C \cap \bigcap_{n=2}^\infty A_n^\complement \cap A_1\right) + \mu^*\left(C \cap \bigcap_{n=2}^\infty A_n^\complement \cap A_1^\complement\right) \\ &= \mu^*\left(C \cap \bigcap_{n=2}^\infty A_n^\complement\right),\end{aligned}$$

weil $A_1 \in \mathcal{A}$ ist. Wenn wir genau dasselbe Argument für A_2 wiederholen, erhalten wir

$$\mu^*\left(C \cap A_2\right) + \mu^*\left(C \cap \bigcap_{n=2}^\infty A_n^\complement\right) = \mu^*\left(C \cap \bigcap_{n=3}^\infty A_n^\complement\right)$$

und somit insgesamt

$$\sum_{n=1}^2 \mu^*\left(C \cap A_n\right) + \mu^*\left(C \cap \bigcap_{n=1}^\infty A_n^\complement\right) = \mu^*\left(C \cap \bigcap_{n=3}^\infty A_n^\complement\right).$$

Indem wir dieses Argument fortführen, erhalten wir induktiv

$$\sum_{n=1}^N \mu^*\left(C \cap A_n\right) + \mu^*\left(C \cap \bigcap_{n=1}^\infty A_n^\complement\right) = \mu^*\left(C \cap \bigcap_{n=N+1}^\infty A_n^\complement\right) \leq \mu^*(C)$$

für alle $N \in \mathbb{N}$. Wegen der Subadditivität von μ^* folgt also

$$\mu^*(C) \leq \mu^*\Big(C \cap \bigcup_{n=1}^{\infty} A_n\Big) + \mu^*\Big(C \cap \big(\bigcup_{n=1}^{\infty} A_n\big)^{\complement}\Big) \tag{1.3.2}$$

$$\leq \sum_{n=1}^{\infty} \mu^*(C \cap A_n) + \mu^*\Big(C \cap \bigcap_{n=1}^{\infty} A_n^{\complement}\Big) \leq \mu^*(C). \tag{1.3.3}$$

Somit ist $\bigcup_{n=1} A_n \in \mathcal{A}$, wie behauptet.

Beweisschritt 4: Zuletzt beweisen wir, dass die Einschränkung $\mu^*|_{\mathcal{A}}$ ein Maß ist. Da μ^* ein äußeres Maß ist, wissen wir bereits, dass $\mu^*(\emptyset) = 0$ ist. Wenn $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt sind und wir Formel (1.3.2) die Menge $C := \bigcup_{n=1}^{\infty} A_n$ einsetzen, dann erhalten wir

$$\mu^*\big(\bigcup_{n=1}^{\infty} A_n\big) = \sum_{n=1}^{\infty} \mu^*(A_n),$$

was die Behauptung zeigt. □

Manchmal ist es bequem, den Elementen der σ-Algebra $\mathcal{A}_{\mu^*}$ aus Theorem 1.3.3 folgende Bezeichnung zu geben:

Definition 1.3.4 (Messbarkeit bezüglich äußerer Maße)**.** Man nennt in der Situation von Theorem 1.3.3 diejenigen Mengen $A \subseteq \Omega$, die Elemente von $\mathcal{A}_{\mu^*}$ sind, $\boldsymbol{\mu^*}$**-messbar**.

Mit Hilfe von Theorem 1.3.3 können wir die Existenz des Lebesgue-Maßes zeigen. Wir geben die Details nur im eindimensionalen Fall an (Beispiel 1.3.5), aber im mehrdimensionalen Fall kann man ganz analog vorgehen (Beispiel 1.3.6).

Vorlesung 6 (Fr, 31.10.) beginnt mit Beispiel 1.3.5

Beispiel 1.3.5 (Das Lebesgue-Maß auf $\mathbb{R}$)**.** Für das Lebesguesche äußere Maß $\lambda_1^*\colon 2^{\mathbb{R}} \to [0,\infty]$ gilt $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_{\lambda_1^*}$ und $\lambda_1^*([a,b]) = b - a$ für alle $a, b \in \mathbb{R}$ mit $a \leq b$. Somit ist

$$\lambda_1 := \lambda_1^*|_{\mathcal{B}(\mathbb{R})}\colon \mathcal{B}(\mathbb{R}) \to [0,\infty]$$

ein Maß mit der Eigenschaft $\lambda_1([a,b]) = b - a$ für alle $a, b \in \mathbb{R}$ mit $a \leq b$ (und laut Beispiel 1.2.14 das einzige Maß $\mathcal{B}(\mathbb{R}) \to [0,\infty]$ mit dieser Eigenschaft).

Damit ist die Existenz (und die Eindeutigkeit) des Lebesgue-Maßes auf $\mathbb{R}$ gezeigt.

Beweis. Die folgende Notation ist hilfreich, um den Beweis effizient aufzuschreiben. Wir setzen $\mathcal{I} := \big\{[a,b) \mid\ a, b \in \mathbb{R},\ a \leq b\big\} \subseteq 2^{\mathbb{R}}$ und für jedes $[a,b) \in \mathcal{I}$ mit $a \leq b$ bezeichne die Länge von $[a,b)$ mit mit $\ell([a,b)) := b - a$.

Man überlegt sich schnell, dass $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$ gilt. Es genügt also $\mathcal{I} \subseteq \mathcal{A}_{\lambda_1^*}$ zu zeigen um die behauptete Inklusion $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_{\lambda_1^*}$ zu erhalten. Sei dazu also $[a,b) \in \mathcal{I}$ mit $a \leq b$ und sei $C \in 2^{\mathbb{R}}$. Wir wollen $[a,b) \in \mathcal{A}_{\lambda_1^*}$ zeigen. Weil λ_1^* σ-subadditiv und wegen $\lambda_1^*(\emptyset) = 0$) somit auch subadditiv ist, genügt es

$$\lambda_1^*\big(C \cap [a,b)\big) + \lambda_1^*\big(C \cap [a,b)^{\complement}\big) \leq \lambda_1^*(C)$$

zu zeigen. Sei dazu $\varepsilon > 0$ beliebig, fest. Dann gibt es reelle Zahlen $a_1 \le b_1$, $a_2 \le b_2$, ... mit der Eigenschaft $\bigcup_{n=1}^{\infty} [a_n, b_n) \supseteq C$ und mit $\sum_{n=1}^{\infty} \ell([a_b, b_n)) \le \lambda_1^*(C) + \varepsilon$. Wegen

$$C \cap [a,b) \subseteq \bigcup_{n=1}^{\infty} \underbrace{[a_n, b_n) \cap [a,b)}_{\in \mathcal{I}}$$

$$\text{und} \qquad C \cap [a,b)^{\complement} \subseteq \bigcup_{n=1}^{\infty} \underbrace{[a_n, b_n) \cap (-\infty, a)}_{\in \mathcal{I}} \cup \bigcup_{n=1}^{\infty} \underbrace{[a_n, b_n) \cap [b, \infty)}_{\in \mathcal{I}}$$

erhalten wir

$$\begin{aligned}
&\lambda_1^*\big(C \cap [a,b)\big) + \lambda_1^*\big(C \cap [a,b)^{\complement}\big) \\
&\le \sum_{n=1}^{\infty} \ell\big([a_n, b_n) \cap [a,b)\big) + \sum_{n=1}^{\infty} \ell\big([a_n, b_n) \cap (-\infty, a)\big) + \sum_{n=1}^{\infty} \ell\big([a_n, b_n) \cap [b, \infty)\big) \\
&= \sum_{n=1}^{\infty} \ell\big([a_n, b_n)\big) \le \lambda_1^*(C) + \varepsilon;
\end{aligned}$$

Für die Gleichheit zwischen der zweiten und dritten Zeile haben wir verwendet, dass $\ell\big([a_n, b_n)\big) = \ell\big([a_n, b_n) \cap [a,b)\big) + \ell\big([a_n, b_n) \cap (-\infty, a)\big) + \ell\big([a_n, b_n) \cap [b, \infty)\big)$ gilt – das heißt, dass die Länge eines Intervalls, dass aus endlich vielen aneinander angrenzenden Teilintervallen zusammengesetzt ist, gleich der Summe der Länge der Teilintervalle ist. Da $\varepsilon > 0$ beliebig war, gilt tatsächlich $[a,b) \in \mathcal{A}_{\lambda_1^*}$. Also ist wie behauptet $\mathcal{B}(\mathbb{R}) \subseteq \mathcal{A}_{\lambda_1^*}$.

Wir müssen noch zeigen, dass für $a, b \in \mathbb{R}$ mit $[a,b]$ die Gleichheit $\lambda_1^*\big([a,b]\big) = b - a$ gilt. Aus Beispiel 1.3.5 wissen wir bereits, dass $\lambda_1^*\big([a,b)\big) = b - a$ gilt. Wir jedes $\varepsilon > 0$ gilt $\{b\} \subseteq [b, b+\varepsilon)$ und somit $\lambda_1^*\big(\{b\}\big) \le \lambda_1^*\big([b, b+\varepsilon)\big) = \varepsilon$, wobei wir für die Ungleichung die Monotonie äußerer Maße verwendet haben. Also ist $\lambda_1^*\big(\{b\}\big) = 0$.

Weil $[a,b]$ und $\{b\}$ Elemente von $\mathcal{B}(\mathbb{R})$ sind und λ_1^* eingeschränkt auf $\mathcal{B}(\mathbb{R})$ ein Maß ist, folgt $\lambda_1^*\big([a,b]\big) = \lambda_1^*\big([a,b)\big) + \lambda_1^*\big(\{b\}\big) = (b-a) + 0 = b - a$. □

Beispiel 1.3.6 (Das Lebesgue-Maß auf $\mathbb{R}^{\boldsymbol{d}}$)**.** Die Existenz des Lebesgue-Maßes $\lambda_d \colon \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$ kann man im Prinzip genauso zeigen wie im eindimensionalen Fall.

Anstelle der halboffenen Intervalle $[a,b)$ arbeitet man nun mit halboffenen Quadern der Form

$$[a_1, b_1) \times \cdots \times [a_d, b_d).$$

Dadurch werden viele der Beweise, die Sie in diesem Abschnitt 1.3 im eindimensionalen Fall gesehen haben, noch etwas technischer im Aufschrieb. Die Quintessenz der Beweise bleibt aber dieselbe, weshalb wir den mehrdimensionalen Fall hier nicht im Detail aufschreiben.

Sie werden im nachfolgenden Abschnitt 1.4 übrigens eine Technik kennenlernen, die man unter anderem verwenden kann, um das mehrdimensionale Lebesgue-Maß aus dem eindimensionalen Lebesgue-Maß zu konstruieren: sogenannte **Produktmaße**. Hierzu benötigt man aber das Existenzresultat im unten stehen Theorem 1.4.5, dessen Beweis wir erst im Laufe von Kapitel 2 geben werden.

1.4 Produktmaße

Einstiegsfragen.

(a) Wenn Sie das die Fläche einer Menge $M \subseteq \mathbb{R}^2$ und die Länge einer Menge $I \subseteq \mathbb{R}$ kennen, kennen Sie dann auch das Volumen von $I \times M \subseteq \mathbb{R}^3$?

(b) Nehmen Sie an, Sie werfen einen Würfel zweimal in Folge. Wie hoch ist die Wahrscheinlichkeit, dass der erste Wurf eine ungerade Zahl ergibt und der zweite Wurf eine der beiden Zahlen 1 oder 2?

Wenn Sie ein Rechteckt im $\mathbb{R}^2$ betrachten, dann ist seine Fläche gleich dem Produkt seiner beiden Seitenlängen. Wenn Sie einen Quader im $\mathbb{R}^3$ betrachten, ist sein Volumen gleich dem Produkt seiner drei Seitenlängen;[7] außerdem kann man das Volumen auch als Produkt einer Seitenlänge und der dazu orthogonalen Seitenfläche schreiben. Diese Idee – einem kartesischen Produkt aus zwei (oder endliche vielen) Mengen – als Maß das Produkt der Maße der einzelnen Mengen zuzuordnen – kann durch die unten stehende Definition 1.4.3 axiomatisieren. Hierfür benötigen wir noch die folgende Konvention über das Produkt von 0 und ∞:

Definition 1.4.1 (Null mal unendlich)**.** In der Maß- und Integrationstheorie setzen wir

$$0 \cdot \infty := \infty \cdot 0 := 0.$$

Das finden Sie vermutlich ungewohnt und auf den ersten Blick etwas merkwürdig. Die Motivation für Definition 1.4.1 ist geometrisch: Betrachten Sie zum Beispiel im $\mathbb{R}^2$ eine unendlich Lange gerade, zum Beispiel die x_1-Achse $\mathbb{R} \times \{0\}$. Diese Linie besitzt die Fläche $\lambda_2(\mathbb{R} \times \{0\}) = 0$ (übrigens: warum eigentlich?). Zugleich erwartet man aber anschaulich, dass die Fläche der Menge $\mathbb{R} \times \{0\}$ dort die Formel "Länge mal Breite" gegeben ist. Die "Länge" von $\mathbb{R} \times \{0\}$ ist ∞ und die "Breite" gleich 0, also drängt es sich auf das Produkt von 0 und ∞ also 0 zu definieren. Aber Vorsicht:

Bemerkung 1.4.2 (Unstetigkeit der Multiplikation bei $(\mathbf{0},\infty)$)**.** Aus der Analysis 1 und 2 sind Sie es gewohnt, dass Multiplikation sich stetig verhält: Für zwei konvergente Zahlenfolgen $a_n \to a$ und $b_n \to b$ gilt $a_n b_n \to ab$ für $n \to \infty$.

Wenn aber zum Beispiel $a = \infty$ ist (das heißt,(a_n) konvergiert nicht gegen eine reelle Zahl, sondern konvergiert uneigentlich gegen ∞) und $b = 0$ ist, dann stimmt das im Allgemeinen nicht mehr (manchmal aber schon):

(a) Sei $a_n = n^2$ und $b_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$. Dann gilt für $n \to \infty$

$$a_n \to \infty, \qquad b_n \to 0, \qquad a_n b_n = n \to \infty \neq \infty \cdot 0.$$

(b) Sei nun stattdessen Sei $a_n = n$ und $b_n = \frac{1}{n^2}$ für alle $n \in \mathbb{N}$. Dann gilt für $n \to \infty$

$$a_n \to \infty, \qquad b_n \to 0, \qquad a_n b_n = \frac{1}{n} \to 0 = \infty \cdot 0.$$

[7] Das hatten wir übrigens – in beliebiger Dimension $d \in \mathbb{N}$ – als definierende Eigenschaft des Lebesgue-Maßes verwendet, siehe Beispiel 1.2.2(c).

Man muss also aufpassen, unter welchen genauen Bedingungen Resultate über Konvergenz und Stetigkeit tatsächlich gelten. Nun kommen wir zur bereits angekündigten Definition, um die sich dieser Abschnitt dreht.

Definition 1.4.3 (Produkt-σ-Algebren und Produktmaße)**.**

(a) Sei $m \in \mathbb{N}$ und seien $(\Omega_1, \mathcal{A}_1), \ldots, (\Omega_m, \mathcal{A}_m)$ Messräume. Dann nennt man die σ-Algebra

$$\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m := \sigma\Big(\{A_1 \times \cdots \times A_m \mid A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m\}\Big)$$

auf $\Omega_1 \times \cdots \times \Omega_m$ die **Produkt-σ-Algebra** der σ-Algebren $\mathcal{A}_1, \ldots, \mathcal{A}_m$.

(b) Sei $m \in \mathbb{N}$ und seien $(\Omega_1, \mathcal{A}_1, \mu_1), \ldots, (\Omega_m, \mathcal{A}_m, \mu_m)$ Maßräume. Ein Maß

$$\mu \colon \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m \to [0, \infty]$$

nennt man ein **Produktmaß** der Maße $\mu_1, \ldots, \mu_m$, falls es

$$\mu(A_1 \times \cdots \times A_m) = \mu_1(A_1) \cdot \cdots \cdot \mu_m(A_m)$$

für alle $A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m$ erfüllt.

Hier ist ein einfaches Beispiel für ein Produktmaß.

Beispiel 1.4.4 (Würfeln)**.** Auf $\Omega_1 := \{1, \ldots, 6\}$ betrachten wir das Wahrscheinlichkeitsmaß $\mu_1 \colon 2^{\Omega_1} \to [0, \infty]$, das durch $\mu_1(A) = \frac{\#A}{6}$ für alle $A \in 2^{\Omega_1}$ gegeben ist. Außerdem setzen wir $\Omega_2 := \Omega_1$ und $\mu_2 := \mu_1$.

Lassen Sie uns auf dem Produktraum $\Omega_1 \times \Omega_2 = \{1, \ldots, 6\}^2$ das Wahrscheinlichkeitsmaß $\mu \colon 2^{\Omega_1 \times \Omega_2} \to [0, \infty]$ betrachten, dass durch $\mu(C) := \frac{\#C}{36}$ für alle $C \in 2^{\Omega_1 \times \Omega_2}$ definiert ist; es beschreibt anschaulich die Wahrscheinlichkeit, beim zweifachen Werfen eines Würfels ein Paar von Augenzahlen zu erhalten, dass in C liegt. Es gilt folgendes:

(a) Es ist $2^{\Omega_1 \times \Omega_2} = 2^{\Omega_1} \otimes 2^{\Omega_2}$.

(b) Das Maß μ ist ein Produktmaß von μ_1 und μ_2.

Beweis. (a) In den Übungen werden Sie beweisen, dass das sogar für alle endlichen Mengen Ω_1 und Ω_2 gilt.

(b) Seien $A \in 2^{\Omega_1}$ und $B \in 2^{\Omega_2} = 2^{\Omega_1}$. Dann gilt

$$\mu(A \times B) = \frac{\#(A \times B)}{36} = \frac{\#A \cdot \#B}{36} = \frac{\#A}{6}\frac{\#B}{6} = \mu_1(A)\mu_2(B). \qquad \square$$

Falls Sie übrigens finden, dass die bisherigen Beispiele zu Wahrscheinlichkeitsmaßen eher einfach gestrickt waren, keine Sorge: In Kapitel 2 werden wir auch interessantere Beispiele von Wahrscheinlichkeitsmaßen treffen.

Es liegt die Frage nahe, ob zwei Maße (oder allgemeiner endlich viele Maße) immer ein Produktmaß besitzen. Die Antwort ist ja:

Theorem 1.4.5 (Existenz von Produktmaßen)**.** *Es sei $m \in \mathbb{N}$ und es seien $(\Omega_1, \mathcal{A}_1, \mu_1), \ldots, (\Omega_m, \mathcal{A}_m, \mu_m)$ Maßräume. Dann gibt es ein Produktmaß μ der Maße $\mu_1, \ldots, \mu_m$.*

Wir werden das Theorem an dieser Stelle noch nicht beweisen, kommen darauf aber in Kapitel 2 noch einmal zurück. Zusätzlich zur Existenz drängt sich die Frage auf, ob Maße $\mu_1, \ldots, \mu_m$ mehrere verschiedene Produktmaße besitzen können oder ob es stets nur eines gibt. Für σ-endliche Maße ist das Produktmaß tatsächlich eindeutig, wie wir nun zeigen.

Korollar 1.4.6 (Existenz und Eindeutigkeit von Produktmaßen im $\boldsymbol{\sigma}$-endlichen Fall)**.** *Sei $m \in \mathbb{N}$ und seien $(\Omega_1, \mathcal{A}_1, \mu_1), \ldots, (\Omega_m, \mathcal{A}_m, \mu_m)$ σ-endliche Maßräume. Dann gibt es genau ein Produktmaß der Maße $\mu_1, \ldots, \mu_m$.*[8]

Vorsicht: Wenn man die σ-Endlichkeit als Voraussetzung weglässt, dann kann es im Allgemeinen mehr als ein Produktmaß geben. Darauf gehen wir an dieser Stelle aber nicht weiter ein.[9]

Beweis von Korollar 1.4.6. Theorem 1.4.5 besagt die Existenz eines Produktmaßes, wir müssen also nur noch die Eindeutigkeit zeigen. Seien $\mu, \nu \colon \mathcal{A}_1 \otimes \times \otimes \mathcal{A}_m \to [0, \infty]$ zwei Produktmaße von $\mu_1, \ldots, \mu_m$. Sei

$$\mathcal{E} \coloneqq \{A_1 \times \cdots \times A_m \mid A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m\}.$$

Wir verwenden den Eindeutigkeitssatz für Maße, Theorem 1.2.13, für das Mengensystem $\mathcal{E}$. Lassen Sie uns nachprüfen, dass die Voraussetzungen des Theorems erfüllt sind:

(1) Seien $A, B \in \mathcal{E}$. Dann gibt es Mengen $A_1, B_1 \in \mathcal{A}_1, \ldots, A_m, B_m \in \mathcal{A}_m$ mit $A = A_1 \times \cdots \times A_m$ und $B = B_1 \times \cdots \times B_m$ und somit ist

$$A \cap B = (A_1 \cap B_1) \times \cdots \times (A_m \cap B_m) \in \mathcal{E}.$$

Also ist $\mathcal{E}$ tatsächlich stabil bezüglich endlicher Durchschnitte.

(2) Sei $A \in \mathcal{E}$. Dann gibt es $A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m$ mit $A = A_1 \times \cdots \times A_m$. Somit folgt aus der Definition von Produktmaßen

$$\mu(A) = \mu_1(A_1) \cdot \cdots \cdot \mu_m(A_m) = \nu(A).$$

Also stimmen die Maße μ und ν auf dem Erzeuger $\mathcal{E}$ überein.

[8] Wir werden weiter unten in diesem Abschnitt zeigen, dass es immer mindestens ein Produktmaß gibt. Zusammen mit dieser Proposition wissen wir dann also: Für σ-endliche Maße $\mu_1, \ldots, \mu_m$ gibt es genau ein Produktmaß.

[9] In einer früheren Version des Manuskripts wurde hier behauptet, dass wir ein Gegenbeispiel in den Übungen besprechen. Das ist aber offenbar doch etwas technischer, als ich zunächst gedacht hätte, deshalb ist die Zeit vielleicht besser genutzt, wenn wir stattdessen andere Dinge in den Übungen besprechen.

(3) Für jedes $k \in \{1,\dots,m\}$ gibt es wegen der σ-Endlichkeit von $(\Omega_k, \mathcal{A}_k, \mu_k)$ eine Folge von Mengen $E_1^{(k)} \subseteq E_2^{(k)} \subseteq \dots$ in $\mathcal{A}_k$, die Ω_k überdecken und $\mu_k(E_n^{(k)}) < \infty$ für alle $n \in \mathbb{N}$ erfüllen. Wir setzen

$$E_n := E_n^{(1)} \times \cdots \times E_n^{(m)} \in \mathcal{E}$$

für jedes $n \in \mathbb{N}$. Dann gilt $E_1 \subseteq E_2 \subseteq \dots$ und $\mu(E_n) = \mu_1(E_n^{(1)}) \times \cdots \times \mu_m(E_n^{(m)}) < \infty$ und

$$\bigcup_{n=1}^{\infty} E_n = \Big(\bigcup_{n=1}^{\infty} E_n^{(1)}\Big) \times \cdots \times \Big(\bigcup_{n=1}^{\infty} E_n^{(m)}\Big) = \Omega_1 \times \cdots \times \Omega_m.$$

Also sind alle Voraussetzungen von Theorem 1.2.13 erfüllt und das Theorem impliziert, dass $\mu = \nu$ gilt. □

Vorlesung 7 (Mi, 05.11.) beginnt voraussichtlich mit Notation 1.3.5

Notation 1.4.7 (Produktmaß)**.** Sei $m \in \mathbb{N}$ und seien $(\Omega_1, \mathcal{A}_1, \mu_1)$, …, $(\Omega_m, \mathcal{A}_m, \mu_m)$ σ-endliche Maßräume. Dann notiert man das (laut Korollar 1.4.6 existente und eindeutig bestimmte) Produktmaß der Maße $\mu_1, \dots, \mu_m$ als

$$\mu_1 \otimes \cdots \otimes \mu_m \colon \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m \to [0,\infty].$$

Wie Sie wissen, kann man für "interessante" σ-Algebren wie beispielsweise die Borel-σ-Algebra auf $\mathbb{R}^d$ nicht alle Elemente explizit angeben, und arbeitet stattdessen meist mit Erzeugern. Um in solche einem Fall mit Produkt-σ-Algebren arbeiten zu können, ist die folgende Proposition nützlich.

Proposition 1.4.8 (Produkt-$\boldsymbol{\sigma}$-Algebra via Erzeugern)**.** *Sei $m \in \mathbb{N}$ und seien $(\Omega_1, \mathcal{A}_1)$, …, $(\Omega_m, \mathcal{A}_m)$ Messräume. Für jedes $k \in \{1,\dots,m\}$ sei außerdem $\mathcal{E}_k \subseteq \mathcal{A}_k$ ein Erzeuger von $\mathcal{A}_k$ und es gebe eine Folge von Mengen $E_{k,1}, E_{k,2}, \dots \in \mathcal{E}_k$ mit $\bigcup_{n\in\mathbb{N}} E_{k,n} = \Omega_k$. Dann ist*

$$\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m = \sigma\Big(\{E_1 \times \cdots \times E_m \mid E_1 \in \mathcal{E}_1, \dots, E_m \in \mathcal{E}_m\}\Big).$$

Beweis. Wir zeigen beide Inklusionen der behaupteten Mengengleichheit:

"$\supseteq$" Diese Inklusion folgt aus

$$\{A_1 \times \cdots \times A_m \mid A_1 \in \mathcal{A}_1, \dots, A_m \in \mathcal{A}_m\} \supseteq \{E_1 \times \cdots \times E_m \mid E_1 \in \mathcal{E}_1, \dots, \mathcal{E}_m\}.$$

"$\subseteq$" Wir werden mehrfach das Prinzip der guten Mengen verwenden. Zur einfacheren Notation setzen wir

$$\mathcal{C} := \sigma\Big(\{E_1 \times \cdots \times E_m \mid E_1 \in \mathcal{E}_1, \dots, E_m \in \mathcal{E}_m\}\Big).$$

Sei

$$\mathcal{B}_1 := \{B_1 \in \mathcal{A}_1 \mid B_1 \times E_2 \times \cdots \times E_m \in \mathcal{C} \text{ für alle } E_1 \in \mathcal{E}_1, \dots, E_m \in \mathcal{E}_m\}.$$

Dann gilt $\mathcal{E}_1 \subseteq \mathcal{B}_1$. Außerdem kann man nachrechnen, dass $\mathcal{B}_1$ eine σ-Algebra auf Ω_1 ist.[10] Hieraus folgt $\mathcal{A}_1 = \sigma(\mathcal{E}_1) \subseteq \mathcal{B}_1$. Für jedes $A_1 \in \mathcal{A}_1$ gilt deshalb $A_1 \in \mathcal{B}_1$ und somit $A_1 \times E_2 \times \cdots \times E_m \in \mathcal{C}$ für alle $E_1 \in \mathcal{E}_1, \ldots, E_m \in \mathcal{E}_m$.

Als nächstes definiert man

$$\mathcal{B}_2 := \{B_2 \in \mathcal{A}_2 \mid A_1 \times B_2 \times E_3 \times \cdots \times E_m \in \mathcal{C} \text{ für alle } A_1 \in \mathcal{A}_1 \text{ und } E_3 \in \mathcal{E}_3, \ldots, E_m \in \mathcal{E}_m\}.$$

Aufgrund des vorangehenden Beweisschritts gilt $\mathcal{E}_2 \subseteq \mathcal{B}_2$. Außerdem kann man direkt nachrechnen, dass $\mathcal{B}_2$ eine σ-Algebra auf Ω_2 ist. Folglich gilt $\mathcal{A}_2 = \sigma(\mathcal{E}_2) \subseteq \mathcal{B}_2$. Also gilt $A_1 \times A_2 \times E_3 \times \cdots \times E_m \in \mathcal{C}$ für alle $A_1 \in \mathcal{A}_2$ und $A_2 \in \mathcal{A}_2$ und alle $E_3 \in \mathcal{E}_3, \ldots, E_m \in \mathcal{E}_m$.

Indem man das entsprechende Argument noch $m-2$-mal wiederholt, erhält man $A_1 \times \cdots \times A_m \in \mathcal{C}$ für alle $A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m$. Also ist

$$\{A_1 \times \cdots \times A_m \mid A_1 \in \mathcal{A}_1, \ldots, A_m \in \mathcal{A}_m\} \subseteq \mathcal{C}$$

und somit $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_m \subseteq \mathcal{C}$. □

Das Paradebeispiel für eine Anwendung von Proposition 1.4.8 ist das folgende Resultat über Borel-σ-Algebren.

Beispiel 1.4.9 (Das Produkt von Borel-$\boldsymbol{\sigma}$-Algebren)**.** Sei $m \in \mathbb{N}$ und seien $d_1, \ldots, d_m \in \mathbb{N}$. Dann gilt

$$\mathcal{B}(\mathbb{R}^{d_1}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}^{d_m}) = \mathcal{B}(\mathbb{R}^{d_1 + \cdots + d_m}).$$

Beweis. Für jedes $d \in \mathbb{N}$ sei $\mathcal{Q}_d \subseteq 2^{\mathbb{R}^d}$ die Menge der kompakten Quader in $\mathbb{R}^d$, das heißt, das Mengensystem, dessen Elemente alle Quader der Form

$$[a_1, b_1] \times \cdots \times [a_d, b_d]$$

mit reellen Zahlen $a_1 \leq b_1, \ldots a_d \leq b_d$ sind. Dann ist $\mathcal{Q}_d$ stabil bezüglicher endlicher Durchschnitte und wie bereits im Beweis von Beispiel 1.2.14 bemerkt, gilt $\sigma(\mathcal{Q}_d) = \mathcal{B}(\mathbb{R}^d)$. Als nächstes beobachten wir, dass

$$\{Q_1 \times \cdots \times Q_m \mid Q_1 \in \mathcal{Q}_{d_1}, \ldots, Q_m \in \mathcal{Q}_{d_m}\} = \mathcal{Q}_{d_1 + \cdots + d_m}$$

gilt. Somit folgt

$$\begin{aligned}\mathcal{B}(\mathbb{R}^{d_1 + \cdots + d_m}) &= \sigma(\mathcal{Q}_{d_1 + \cdots + d_m}) = \sigma\Big(\{Q_1 \times \cdots \times Q_m \mid Q_1 \in \mathcal{Q}_{d_1}, \ldots, Q_m \in \mathcal{Q}_{d_m}\}\Big) \\ &= \sigma(\mathcal{Q}_{d_1}) \otimes \cdots \otimes \sigma(\mathcal{Q}_{d_m}) = \mathcal{B}(\mathbb{R}^{d_1}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R}^{d_m}),\end{aligned}$$

wobei die Gleichheit zwischen den beiden Zeilen aus Proposition 1.4.8 folgt. □

[10] Für die Abgeschlossenheit bezüglich Komplementbildung benötigt man, dass $\Omega_1 \times E_2 \times \cdots \times E_m \in \mathcal{C}$ für alle $E_2 \in \mathcal{E}_2, \ldots, E_m \in \mathcal{E}_m$ gilt. Dies folgt aus der Annahme, dass man Ω_1 als abzählbare Vereinigung von Elementen aus $\mathcal{E}_1$ schreiben kann.

Nach dem vorangehenden Beispiel drängt sich das folgende Beispiel geradezu auf.

Beispiel 1.4.10 (Produkt von Lebesgue-Maßen)**.** Sei $m \in \mathbb{N}$ und seien $d_1, \dots, d_m \in \mathbb{N}$. Dann gilt

$$\lambda_{d_1} \otimes \cdots \otimes \lambda_{d_m} = \lambda_{d_1 + \cdots + d_m}.$$

Beweis. Das folgt aus der definierenden Eigenschaft des Lebesgue-Maßes (Definition (c)), der Definition von Produktmaßen (Definition 1.4.3) und der Eindeutigkeit von Produktmaßen σ-endlicher Maße (Korollar 1.4.6). Die Details besprechen wir in einer Übungsaufgabe. □

1.5 Toplogische im Vergleich zu maßtheoretischen Eigenschaften von Mengen

Einstiegsfragen.

(a) Stellen Sie sich $[0,1] \cap \mathbb{Q}$ im Vergleich zu $[0,1]$ eher als kleine oder eher als große Menge vor? Warum?

(b) Was ist Ihre Anschauung, wenn Sie an eine "große" Teilmenge von $\mathbb{R}$ denken sollen? Und wenn Sie an eine "große" Teilmenge von $[0,1]$ denken sollen?

In der Analysis 2 hatten wir für eine Teilmenge $C \subseteq M$ eines metrischen Raumes (M, d) den **Abschluss** $\overline{C}$ definiert als den Durchschnitt aller abgeschlossenen Teilmengen von M, die C enthalten. Da der Durchschnitt beliebig vieler abgeschlossener Mengen ebenfalls abgeschlossen ist, ist auch $\overline{C}$ abgeschlossen – und man sieht leicht, dass $\overline{C}$ die kleinste abgeschlossene Teilmenge von M ist, die C enthält.[11]

Wir führen jetzt noch zwei weitere Begriffe ein, die zu Teilmengen eines metrischen Raumes gehören. Einen von beiden, den **Rand**, hatten wir in der Analysis 2 ab und zu umgangssprachlich verwendet, ihn aber bisher nicht präzisiert.

Definition 1.5.1 (Inneres und Rand einer Menge)**.** Sei (M, d) ein metrischer Raum und $C \subseteq M$.

(a) Das **Innere von C**[12,13] ist die Menge

$$C^\circ := \bigcup_{\substack{U \subseteq C, \\ U \text{ offen}}} U.$$

[11] Vergleichen Sie das am besten mit der Definition der von einem Mengensystem $\mathcal{E} \subseteq 2^\Omega$ auf einer Menge Ω erzeugten σ-Algebra $\sigma(\mathcal{E}) \subseteq 2^\Omega$. Sie werden feststellen, dass $\overline{C}$ und $\sigma(\mathcal{E})$ ganz analog definiert sind.

[12] In manchen Quellen wird alternativ auch der Begriff **offener Kern von C** benutzt. Das ist aber vielleicht ein bisschen aus der Mode gekommen, und vor allem ist der Begriff "Kern" in der Analysis ohnehin schon mit mehreren Bedeutungen überladen. Ich würde Ihnen deshalb empfehlen, einfach vom Inneren von C zu sprechen.

[13] Manchmal setzt man in der Notation C° den Kreis auch direkt über C. Ein Nachteil von dieser alternativen Notation ist, dass sie sich nicht gut für Mengen verwenden lässt, die aus mehreren Mengen zusammengesetzt sind: Zum Beispiel ist die Bedeutung von $(C_1 \cap C_2)^\circ$ klar; wenn man den Kreis stattdessen über die Menge $C_1 \cap C_2$ setzen möchte, wird die Notation etwas merkwürdig und im schlimmsten Fall uneindeutig.

Die Elemente von C° nennt man die **inneren Punkt von C**.

(b) Die Menge $\partial C := \overline{C} \setminus (C^\circ)$ nennt man den **Rand von C**.

In der folgenden Proposition fassen wir einige elementare, aber wichtige Eigenschaften von Abschluss, Innerem und Rand zusammen. Für den Abschluss haben wir manche dieser Eigenschaften bereits in der Analysis 2 besprochen, aber der Vollständigkeit halber führen wir sie trotzdem noch einmal mit auf.

Proposition 1.5.2 (Eigenschaften von Innerem, Rand und Abschluss)**.** *Sei (M,d) ein metrischer Raum und sei $C \subseteq M$.*

(a) *Es gilt $C^\circ \subseteq C \subseteq \overline{D}$.*

(b) *Ist $D \subseteq M$ mit $C \subseteq D$, so gilt $\overline{C} \subseteq \overline{D}$ und $C^\circ \subseteq D^\circ$.*

(c) *Die Menge C ist genau dann abgeschlossen, wenn $\overline{C} = C$ gilt. Sie ist genau dann offen, wenn $C^\circ = C$ gilt.*

(d) *Die Operationen Abschluss und Inneres sind **idempotent**, das heißt, es gilt $\overline{\overline{C}} = \overline{C}$ und $(C^\circ)^\circ = C^\circ$.*

(e) *Es gilt $\overline{C}^{\complement} = (C^{\complement})^\circ$ und $\overline{C^{\complement}} = (C^\circ)^{\complement}$.*

(f) *Der Rand ∂C ist abgeschlossen und erfüllt $\partial C = \overline{C} \cap \overline{C^{\complement}}$.*

(g) *Es gilt $\partial C = \partial(C^{\complement})$.*

(h) *Für alle $C_1, \ldots, C_n$ gilt*

$$\overline{C_1 \cup \cdots \cup C_n} = \overline{C_1} \cup \cdots \cup \overline{C_n} \quad \textit{und} \quad \overline{C_1 \cap \cdots \cap C_n} \subseteq \overline{C_1} \cap \cdots \cap \overline{C_n}$$

sowie

$$(C_1 \cup \cdots \cup C_n)^\circ \supseteq C_1^\circ \cup \cdots \cup C_n^\circ \quad \textit{und} \quad (C_1 \cap \cdots \cap C_n)^\circ = C_1^\circ \cap \cdots \cap C_n^\circ.$$

Beweis.

(a), (b) *und* (c) Das folgt leicht aus den Definitionen von Abschluss und Innerem.

(d) Das folgt unmittelbar aus (c).

(e) Das erhält man aus der Definition des Randes und des Inneren und daraus, dass eine Menge genau dann abgeschlossen ist, wenn ihr Komplememt offen ist (und umgekehrt).

(f) Es gilt $\partial C = \overline{C} \setminus (C^\circ) = \overline{C} \cap (C^\circ)^{\complement} = \overline{C} \cap \overline{C^{\complement}}$, wobei die dritte Gleichheit aus (e) folgt. Insbesondere ist ∂C als Durchschnitt von zwei abgeschlossenen Mengen abgeschlossen.

(g) Wegen $\left(C^{\complement}\right)^{\complement} = C$ folgt dies direkt aus der Formel in (f).

(h) Wir zeigen die Aussagen über den Abschluss; die Aussagen über das Innere kann man analog beweisen oder, alternativ, mit Hilfe von (e) aus den Aussagen über den Abschluss herleiten.

Erste Inklusion: Wir zeigen $\overline{C_1 \cup \cdots \cup C_n} \subseteq \overline{C_1} \cup \cdots \cup \overline{C_n}$.

Wegen (a) gilt $C_1 \cup \cdots \cup C_n \subseteq \overline{C_1} \cup \cdots \cup \overline{C_n}$. Außerdem ist die rechtsstehende Menge als Vereinigung endlich vieler abgeschlossener Mengen abgeschlossen und somit folgt

$$\overline{C_1 \cup \cdots \cup C_n} \overset{\text{(b)}}{\subseteq} \overline{\overline{C_1} \cup \cdots \cup \overline{C_n}} \overset{\text{(c)}}{=} \overline{C_1} \cup \cdots \cup \overline{C_n}.$$

Zweite Inklusion: Wir zeigen $\overline{C_1 \cup \cdots \cup C_n} \supseteq \overline{C_1} \cup \cdots \cup \overline{C_n}$.

Für jedes $k \in \{1, \ldots, n\}$ gilt $C_1 \cup \cdots \cup C_n \supseteq C_k$ und somit wegen (b) auch $\overline{C_1 \cup \cdots \cup C_n} \supseteq \overline{C_k}$. Hieraus folgt die behauptete Inklusion.

Dritte Inklusion: Wir zeigen $\overline{C_1 \cap \cdots \cap C_n} \subseteq \overline{C_1} \cap \cdots \cap \overline{C_n}$.

Wegen (a) gilt $C_1 \cap \cdots \cap C_n \subseteq \overline{C_1} \cap \cdots \cap \overline{C_n}$. Die rechtsstehende Menge ist als Durchschnitt abgeschlossener Mengen wieder abgeschlossen, also folgt

$$\overline{C_1 \cap \cdots \cap C_n} \overset{\text{(b)}}{\subseteq} \overline{\overline{C_1} \cap \cdots \cap \overline{C_n}} \overset{\text{(c)}}{=} \overline{C_1} \cap \cdots \cap \overline{C_n},$$

womit alle behaupteten Inklusionen gezeigt sind. □

Beispiele 1.5.3 (Einige Beispiele zu Innerem, Rand und Abschluss)**.** Wir statten $\mathbb{R}$ mit der Euklidischen Metrik aus.

(a) Es gilt $\overline{\mathbb{Q}} = \mathbb{R}$, $\mathbb{Q}^\circ = \emptyset$ and $\partial\mathbb{Q} = \mathbb{R}$.

(b) Für die Menge $C := \mathbb{Q} \cap [0,1]$ gilt $\overline{C} = [0,1]$, $C^\circ = \emptyset$ und $\partial C = [0,1]$.

(c) Seien $a, b \in \mathbb{R}$ mit $a < b$. Dann ist $\overline{[a,b)} = [a,b]$, $[a,b)^\circ = (a,b)$ und $\partial[a,b) = \{a,b\}$.

Beweis. (a) In der Analysis 1 haben wir gezeigt, dass man jede reelle Zahl durch eine Folge von rationalen Zahlen approximieren kann, also gilt $\overline{\mathbb{Q}} = \mathbb{R}$. Weil zwischen zwei verschiedenen rationalen Zahlen stets eine irrationale liegt, enthält $\mathbb{Q}$ kein Intervall mit mehr als einem Punkt, also besitzt $\mathbb{Q}$ keine nichtleere offene Teilmenge. Somit gilt $\mathbb{Q}^\circ = \emptyset$. Damit folgt $\partial\mathbb{Q} = \overline{\mathbb{Q}} \setminus (\mathbb{Q}^\circ) = \mathbb{R} \setminus \emptyset = \mathbb{R}$.

(b) Das beweist man analog zu (a).

(c) Den Beweis stellen wir als Übungsaufgabe. □

Wir erinnern daran, dass man eine Teilmenge $C \subseteq M$ eines metrischen Raumes (M, d) **dicht** – oder genauer, **dicht in** $\boldsymbol{M}$ – nennt, falls $\overline{C} = M$ gilt. Die folgende Charakterisierung von dichten Teilmengen ist einfach, aber oft nützlich.

Vorlesung 8 (Fr, 07.11.) beginnt mit Proposition 1.5.4

Proposition 1.5.4 (Charakterisierung dichter Mengen)**.** *Sei* (M,d) *ein metrischer Raum. Eine Menge* $C \subseteq M$ *ist genau dann dicht, wenn* $V \cap C \neq \emptyset$ *für jede nichtleere offene Menge* $V \subseteq M$ *gilt.*

Beweis. "$\Rightarrow$" Sei C dicht in M und sei $V \subseteq M$ nichtleer und offen. Wir wählen einen Punkt $v \in V$. Da V offen ist, gibt es $\varepsilon > 0$ mit $B_{\leq\varepsilon}(v) \subseteq V$. Wegen $\overline{C} = M$ gibt es eine Folge (c_n) in C mit $c_n \to v$.[14] Insbesondere gibt es ein $n_0 \in \mathbb{N}$ mit $\mathrm{d}(c_{n_0}, v) \leq \varepsilon$, also ist $c_{n_0} \in B_{\leq\varepsilon}(v) \subseteq V$ und somit gilt $c_{n_0} \in C \cap V$.

"$\Leftarrow$" Sei $V \cap C \neq \emptyset$ für jede nichtleere offene Menge $V \subseteq M$. Sei $x \in M$. Wir müssen zeigen, dass $x \in \overline{C}$ gilt; hjerzu werden wir eine Folge in C bauen, die gegen x konvergiert.

Für jedes $n \in \mathbb{N}$ ist $B_{<\frac{1}{n}}(x)$ offen und nichtleer und schneidet somit laut Annahme die Menge C, das heißt, wir finden einen Punkt $c_n \in C \cap B_{<\frac{1}{n}}(x)$. Somit ist $(c_n)_{n\in\mathbb{N}}$ eine Folge in C und es gilt $\mathrm{d}(c_n, x) < \frac{1}{n} \to 0$ und folglich $c_n \to c$ für $n \to \infty$. □

Es gibt dichte Teilmengen von $\mathbb{R}$, die Lebesgue-Maß 0 haben, wie Teil (c) des folgenden Beispiels zeigt.

Beispiel 1.5.5 (Einige Mengen mit Lebesgue-Maß **0**)**.**

(a) Alle Einpunkt-Mengen in $\mathbb{R}$ haben eindimensionales Lebesgue-Maß 0, das heißt, für jedes $x \in \mathbb{R}$ gilt $\lambda_1\big(\{x\}\big) = 0$.

(b) Sei $A \subseteq \mathbb{R}$ abzählbar.[15] Dann ist A Borel-messbar (das heißt $A \in \mathcal{B}(\mathbb{R})$) und es gilt $\lambda_1(A) = 0$.

(c) Die Menge $\mathbb{Q} \subseteq \mathbb{R}$ ist dicht in $\mathbb{R}$, ist Borel-messbar und erfüllt $\lambda_1(\mathbb{Q}) = 0$.

Beweis. (a) Laut Definition des ein-dimensionalen Lebesgue-Maßes (Beispiel (c)) gilt $\lambda_1\big(\{x\}\big) = \lambda_1\big([x,x]\big) = x - x = 0$ für jedes $x \in \mathbb{R}$.

(b) Es ist $A = \bigcup_{a\in A}\{a\}$, also ist A als abzählbare Vereinigung von Borel-messbaren Mengen ebenfalls Borel-messbar. Außerdem folgt aus der σ-Additivität von λ_1 und aus (a)

$$\lambda_1(A) = \sum_{a\in A} \lambda_1\big(\{a\}\big) = \sum_{a\in A} 0 = 0.$$

(c) Die Borel-Messbarkeit von $\mathbb{Q}$ und die Eigenschaft $\lambda_1(\mathbb{Q}) = 0$ folgen aus (b), weil $\mathbb{Q}$ abzählbar ist. Das $\mathbb{Q}$ dicht in $\mathbb{R}$ ist, steht in Beispiel 1.5.3(a). □

Definition 1.5.6 (Nirgends dichte Mengen)**.** Sei (M,d) ein metrischer Raum.

[14] Wir verwenden wir übrigens eine kleine notationelle Ungenauigkeit, die man sehr häufig sieht: Wir schreiben nicht $(c_n)_{n\in\mathbb{N}}$ oder $(c_n)_{n\in\mathbb{N}_0}$, sondern einfach nur (c_n) und lassen die genaue Indexmenge der Folge somit im Unklaren. Das kann man immer dann tun, wenn die genaue Indexmenge für das Argument keine Rolle spielt.

[15] Also endlich oder abzählbar unendlich.

(a) Eine Menge $C \subseteq M$ heißt **nirgends dicht**, falls ihr Abschluss keine inneren Punkte besitzt, das heißt, falls $(\overline{C})^\circ = \emptyset$ gilt.

(b) Eine Menge $C \subseteq M$ heißt **mager**, wenn man sie als abzählbare Vereinigung von nirgends dichten Mengen schreiben kann.

Lassen Sie uns die Begriffe "nirgends dicht", "dicht" und "mager" an einem einfachen Beispiel veranschaulichen:

Beispiel 1.5.7 (Nirgends dichte, dichte und magere Mengen in $\mathbb{R}$)**.** Sei $\mathbb{R}$ mit der Euklidischen Metrik ausgestattet.

(a) Jede endliche Teilmenge von $\mathbb{R}$ ist nirgends dicht.[16]

(b) Die Menge $\mathbb{Q}$ ist dicht und mager in $\mathbb{R}$.

Beweis. (a) Sei $E \subseteq \mathbb{R}$ endlich. Dann ist E abgeschlossen und hat außerdem leeres Inneres, weil in E kein Intervall mit mehr als einem Punkt liegt. Also gilt $\overline{E}^\circ = E^\circ = \emptyset$, das heißt, wie behauptet ist E nirgends dicht.

(b) Laut Beispiel 1.5.3(a) ist $\overline{\mathbb{Q}} = \mathbb{R}$, das heißt, $\mathbb{Q}$ ist in der Tat dicht in $\mathbb{R}$. Andererseits gilt $\mathbb{Q} = \bigcup_{q\in\mathbb{Q}} \{q\}$; weil $\mathbb{Q}$ abzählbar ist, ist $\mathbb{Q}$ also eine abzählbare Vereinigung von Ein-Punkt-Mengen. Laut (a) sind Ein-Punkt-Mengen in $\mathbb{R}$ nirgends dicht, also ist $\mathbb{Q}$ eine abzählbare Vereinigung von nirgends dichten Mengen und somit mager. □

Bitte beachten Sie unbedingt: Wie viele topologische Eigenschaften hängen die Eigenschaften dicht, nirgends dicht und mager nicht nur von der Metrik, sondern auch von der Wahl das umgebenden Raums ab. Zum Beispiel kann man sich leicht überlegen, dass im metrischen Raum $(\mathbb{N}, \mathrm{d}_{\mathrm{Eukl}})$ jede Ein-Punkt-Menge offen ist und somit nicht nirgends dicht ist.

Proposition 1.5.8 (Eigenschaften von mageren Mengen)**.** *Sei* (M, d) *ein metrischer Raum.*

(a) *Seien* $C_1, C_2, \dots \subseteq M$ *mager. Dann ist auch* $\bigcup_{n\in\mathbb{N}} C_n$ *mager.*

(b) *Jede Teilmenge einer nirgends dichten Menge ist nirgends dicht.*

(c) *Jede Teilmenge einer mageren Menge ist mager.*

Beweis. (a) Für jedes $n \in \mathbb{N}$ gibt es, da C_n mager ist, nirgends dichte Mengen $D_n^{(1)}, D_n^{(2)}, \dots \subseteq M$ mit $C_n = \bigcup_{k\in\mathbb{N}} D_n^{(k)}$. Somit ist

$$\bigcup_{n\in\mathbb{N}} C_n = \bigcup_{n\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} D_n^{(k)} = \bigcup_{(n,k)\in\mathbb{N}\times\mathbb{N}} D_n^{(k)}.$$

Weil $\mathbb{N}\times\mathbb{N}$ abzählbar ist, ist $\bigcup_{n\in\mathbb{N}} C_n$ also eine abzählbare Vereinigung von nirgends dichten Teilmengen und somit mager.

[16] Vorsicht: Wenn wir von einer endlichen Mengen sprechen, dann meinen wir eine Menge mit nur endlich vielen Elementen, nicht etwas eine Menge mit endlichem Maß.

(b) Sei $C \subseteq M$ nirgends dicht und sei $B \subseteq C$. Weil Abschluss und Inneres laut Proposition 1.5.2(b) monoton sind, folgt $\emptyset \subseteq \overline{B}^\circ \subseteq \overline{C}^\circ = \emptyset$, also ist B tatsächlich nirgends dicht.

(c) Sei $C \subseteq M$ mager und sei $B \subseteq C$. Dann gibt es nirgends dichts Mengen $D_1, D_2, \cdots \subseteq M$ mit $C = \bigcup_{n\in\mathbb{N}} C_n$. Somit ist

$$B = B \cap C = \bigcup_{n\in\mathbb{N}} (B \cap D_n).$$

Wegen (b) ist $B \cap C_n$ für jedes n nirgends dicht, also ist B tatsächlich eine abzählbare Vereinigung von nirgens dichten Mengen und somit mager. □

Magere Mengen sind vor allem in sogenannten **Baire-Räumen** interessant. Diese definieren wir in Definition 1.5.10 als diejenigen metrischen Räume, die die äquivalenten Bedingungen in der folgenden Proposition 1.5.9 erfüllen.

Proposition 1.5.9 (Charakterisierung von Baire-Räumen)**.** *Sei* (M, d) *ein metrischer Raum. Die folgenden Aussagen sind äquivalent.*

(i) *Jede magere Teilmenge von M hat leeres Inneres.*

(ii) *Das Komplement jeder mageren Teilmenge von M ist dicht.*

(iii) *Wenn $C_1, C_2, \cdots \subseteq M$ abgeschlossene Mengen mit leerem Inneren sind, dann hat auch $\bigcup_{n\in\mathbb{N}} C_n$ leeres Inneres.*

(iv) *Wenn $U_1, U_2, \cdots \subseteq M$ offen und dicht sind, dann ist auch $\bigcap_{n\in\mathbb{N}} U_n$ dicht.*

Beweis. *"*(i) ⇔ (ii)*" und "*(iii) ⇔ (iv)*"* Aus Proposition 1.5.2(e) folgt, dass eine Teilmenge C eines metrischen Raumes genau dann leeres Inneres hat, wenn ihr Komplement dicht ist. Hieraus kann man leicht die beiden behaupteten Äquivalenzen herleiten.

"(i) ⇒ (iii)*"* Wenn $C_1, C_2, \cdots \subseteq M$ abgeschlossene Mengen mit leerem Inneren sind, dann ist jede dieser Mengen nirgends dicht. Also ist $\bigcup_{n\in\mathbb{N}} C_n$ mager und hat somit laut (i) leeres Inneres.

"(iii) ⇒ (i)*"* Sei $C \subseteq M$ mager. Dann gibt es nirgends dichte Mengen $D_1, D_2, \cdots \subseteq M$ mit $C = \bigcup_{n\in\mathbb{N}} D_n$. Die Mengen $\overline{D_1}, \overline{D_2}, \ldots$ sind abgeschlossen und besitzen leeres Inneres, also besitzt laut (iii) auch $\bigcup_{n\in\mathbb{N}} \overline{D_n}$ leeres Inneres. Somit ist

$$C^\circ = \Big(\bigcup_{n\in\mathbb{N}} D_n\Big)^\circ \subseteq \Big(\bigcup_{n\in\mathbb{N}} \overline{D_n}\Big)^\circ = \emptyset,$$

das heißt C besitzt wie behauptet leeres Inneres. □

Definition 1.5.10 (Baire-Räume)**.** Ein metrischer Raum (M, d) heißt ein **Baire-Raum**, wenn er eine (und somit alle) der äquivalenten Eigenschaften aus Proposition 1.5.9 erfüllt.

Folgende Einfache Konsequenz aus Proposition 1.5.9 und der Definition von Baire-Räumen ist oft sehr nützlich.

Korollar 1.5.11 (Überdeckung durch abzählbare viele abgeschlossene Mengen)**.** *Sei der metrische Raum (M,d) ein Baire-Raum und $M \neq \emptyset$. Seien $C_1, C_2, \cdots \subseteq M$ abgeschlossene Mengen mit $\bigcup_{n\in\mathbb{N}} C_n = M$. Dann hat mindestens eine der Mengen C_n nichtleeres Inneres.*

Beweis. Wenn jeder der abgeschlossenen Mengen C_n leeres Inneres hat, dann sind alle diese Mengen mager und somit ist auch $M = \bigcup_{n\in\mathbb{N}} C_n$ mager. Da M ein Baire-Raum ist, folgt hieraus, dass M selbst leeres Inneres hat. Also ist $\emptyset = M^\circ = M$, was der Annahme $M \neq \emptyset$ widerspricht. □

Damit Proposition 1.5.9 und Korollar 1.5.11 uns irgendetwas nützen, brauchen wir ein Kriterium, mit dem wir beweisen können, dass zumindest manche interessante metrische Räume Baire-Räume sind. Solch ein Kriterium folgt im nächsten Theorem.

Theorem 1.5.12 (Satz von Baire)**.** *Jeder vollständige metrischer Raum ist ein Baire-Raum.*[17]

Beweis. Sei (M,d) ein vollständiger metrischer Raum. Wir zeigen Eigenschaft (iv) aus Proposition 1.5.9, also dass der Durchschnitt abzählbar vieler offener dichter Teilmengen von M wieder dicht ist. Seien also $U_1, U_2, \cdots \subseteq M$ offen und dicht. Laut Proposition 1.5.4 genügt es hierfür zu zeigen, dass $V \cap \bigcap_{n=1}^{\infty} U_n \neq \emptyset$ für jedes offene $\emptyset \neq V \subseteq M$ gilt. Fixieren wir also solch ein V.

Wir bauen nun rekursiv eine Folge $(x_n)_{n\in\mathbb{N}_0}$ in V und eine Folge von Zahlen $(r_n)_{n\in\mathbb{N}_0}$ in $(0,\infty)$ derart, dass $r_n \leq \frac{1}{2} r_{n-1}$ und $B_{\leq r_n}(x_n) \subseteq V \cap B_{<r_{n-1}}(x_{n-1}) \cap U_n$ für alle $n \in \mathbb{N}$ gilt. Anschließend werden wir zeigen, dass die Folge $(x_n)_{n\in\mathbb{N}_0}$ konvergiert und ihr Grenzwert von $V \cap \bigcap_{n\in\mathbb{N}} U_n$ liegt, womit diese Menge nicht leer sein kann. Wir konstruieren $(x_n)_{n\in\mathbb{N}_0}$ und $(r_n)_{n\in\mathbb{N}_0}$ auf folgende Weise:

(1) Wir setzen $r_0 := 1$ und wählen x_0 als ein beliebiges Element von V; solch ein Element existiert wegen $V \neq \emptyset$.

(2) Seien nun für ein $n \in \mathbb{N}$ die Elemente $x_{n-1} \in V$ und $r_{n-1} \in (0,\infty)$ bereits gewählt. Dann ist die Menge $V \cap B_{<r_{n-1}}(x_{n-1})$ als Durchschnitt zweier offener Mengen offen und sie ist nichtleer, da sie x_{n-1} enthält. Da U_n laut Annahme dicht in M ist, ist somit auch die offene Menge $V \cap B_{<r_{n-1}}(x_{n-1}) \cap U_n$ nichtleer. Also gibt es einen Punkt x_n in dieser Menge und wegen der Offenheit der Menge finden wir ein $r_n \in (0,\infty)$ mit $B_{\leq r_n}(x_n) \subseteq V \cap B_{<r_{n-1}}(x_{n-1}) \cap U_n$. Indem wir r_n, falls nötig, noch etwas verkleinern, erreichen wir zudem, dass $r_n \leq \frac{1}{2} r_{n-1}$ gilt.

[17] Übrigens kann man das Konzept metrischer Räume verallgemeinern und sogenannte **topologische Räume** studieren. Für topologische Räume gilt Proposition 1.5.9 ebenfalls und auch Baire-Räume definiert man genauso. In diesem Rahmen kann man dann noch weitere Klassen von Räumen identifizieren, die Baire-Räume sind (beispielsweise alle **kompakten Räume** und allgemeiner alle **lokal kompakten Räume**).

Eine Vorlesung *Einführung in die Topologie* findet in Wuppertal üblicherweise jedes Wintersemester statt. Außerdem geben wir im kommenden Sommersemester innerhalb der Vorlesung *Grundlagen der Funktionalanalysis* eine sehr knappe Einführung in topologische Räume (aber mit etwas anderem Fokus als in der *Einführung in die Topologie*).

Für alle $n, m \in \mathbb{N}_0$ mit $n \geq m$ gilt

$$x_n \in B_{\leq r_n}(x_n) \subseteq B_{\leq r_{n-1}}(x_{n-1}) \subseteq \cdots \subseteq B_{\leq r_m}(x_m) \tag{1.5.1}$$

und somit $\mathrm{d}(x_n, x_m) \leq r_m \leq \frac{1}{2^m}$. Also ist $(x_n)_{n\in\mathbb{N}_0}$ eine Cauchy-Folge und somit wegen der Vollständigkeit von (M, d) konvergent gegen ein $x \in M$.

Für jedes feste $m \in \mathbb{N}_0$ liegt die Folge $(x_n)_{n\in\mathbb{N}_0}$ wegen (1.5.1) schließlich[18] in der abgeschlossenen Menge $B_{\leq r_m}(x_m)$. Also gilt für jedes $m \in \mathbb{N}$ auch $x \in B_{\leq r_m}(x_m)$ und somit $x \in V \cap U_m$, das heißt, es ist $x \in V \cap \bigcap_{m\in\mathbb{N}} U_m$. □

Vorlesung 9 (Mi, 12.11.) beginnt mit Beispiel 1.5.13

Man interpretiert magere Mengen – zumindest in Baire-Räumen – als "topologisch klein". Überraschenderweise ist die Frage, ob eine Teilmenge von $\mathbb{R}$ topologisch klein ist, unabhängig davon, ob auch ihr Lebesguemaß klein ist. Das zeigen wir in den folgenden Beispielen mit Hilfe von sogenannten **Cantormengen**.

Beispiel 1.5.13 (Cantormengen)**.** Wir geben eine Verfahren an um eine Menge sukzessive "auszudünnen":

Ausdünnen eines Intervalls: Dazu sei $\eta \in (0,1)$. Für ein kompaktes Intervall $[a,b] \subseteq \mathbb{R}$ mit $a < b$ bezeichne $S_\eta([a,b]) \subseteq \mathbb{R}$ diejenige Menge, die entsteht, wenn man genau in der Mitte dasjenige offene Intervall von $[a,b]$ wegnimmt, dessen Länge den Anteil $1-\eta$ an der Länge von $[a,b]$ hatte. Expliziter ausgedrückt ist also

$$S_\eta([a,b]) := [a, a+\tfrac{\eta}{2}(b-a)] \cup [b-\tfrac{\eta}{2}(b-a), b].$$

Ausdünnen mehrerer Intervalle: Sei weiterhin $\eta \in (0,1)$. Wenn $A \subseteq \mathbb{R}$ eine Vereinigung von endlich vielen paarweise disjunkten kompakten Intervallen mit je mehr als einem Punkt ist – das heißt, wenn $A = [a_1, b_1] \cup \cdots \cup [a_n, b_n]$ für Punkte $a_1 < b_1 < a_2 < b_2 < \cdots < a_n < b_n$ ist –, setzen wir $S_\eta(A) := S_\eta([a_1, b_1]) \cup \cdots \cup S_\eta([a_n, b_n])$. Beachten Sie, dass $S_\eta(A)$ wieder von derselben Form ist, aber aus $2n$ Intervallen besteht.

Sukzessives Ausdünnen eines Intervalls: Nun sei $(\eta_n)_{n\in\mathbb{N}}$ eine Folge in $(0,1)$. Wir konstruieren eine absteigende Folge von Mengen $[0,1] \supseteq C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots$ durch

$$C_0 := [0,1] \qquad \text{und} \qquad C_n := S_{\eta_n}\big(C_{n-1}\big) \quad \text{für alle } n \in \mathbb{N},$$

und wir setzen

$$C := \bigcap_{n\in\mathbb{N}_0} C_n \subseteq [0,1].$$

Beachte Sie, dass sowohl die Mengen C_n also auch ihr Durschnitt C von der Wahl der Folge $(\eta_n)_{n\in\mathbb{N}}$ abhängen. Man nennt C eine **Cantor-Menge**. Sie hat folgende Eigenschaften:

[18] Wenn man sagt, dass eine Folge (x_n) **schließlich** in einer Menge C liegt, dann meint man damit, dass $x_n \in C$ für alle genügend großen Indizes n gilt – das heißt, dass es ein n_0 gibt mit $x_n \in C$ für alle $n \geq n_0$.

(a) Die Menge C ist kompakt.

(b) Die Menge C ist im metrischen Raum $(\mathbb{R}, \mathrm{d}_{\mathrm{Eukl}})$ nirgends dicht (also insbesondere mager).

(c) Die Menge C besitzt keine isolierten Punkte, das heißt, für jedes $c \in C$ und jedes $\varepsilon > 0$ enthält $B_{<\varepsilon}(c) \cap C$ noch weitere Punkte außer c.[19]

(d) Die Menge C ist überabzählbar.

(e) Die Menge C hat das Lebesguemaß $\lambda_1(C) = \lim_{n\to\infty} \prod_{k=1}^{n} \eta_k$.

Beweis. (a) Es ist C als Durchschnitt abgeschlossener Mengen ebenfalls abgeschlossen und als Teilmenge von $[0,1]$ beschränkt, somit kompakt.[20]

(b) Wegen (a) ist C abgeschlossen, also müssen wir nur $C^\circ = \emptyset$ zeigen. Für jedes $n \in \mathbb{N}_0$ besteht C_n aus endlich vielen paarweise disjunkten kompakten Intervallen;[21] beim Übergang von C_n zu C_{n+1} wird jedes dieser Intervalle in der Mittel aufgeteilt und ein Teil in der Mitte weggenommen. Somit wird die Länge er einzelnen Intervalle in jedem Schritt mindestens halbiert, das heißt, jedes der Intervalle C_n, aus denen C_n hat höchstens die Länge $\frac{1}{2^n}$. Somit enthält C_n keine Intervall als Teilmenge, welches länger ist als $\frac{1}{2^n}$.

Wenn $\varepsilon > 0$ gegeben ist, können wir ein $n \in \mathbb{N}_0$ mit $\frac{1}{2^n} < \varepsilon$ betrachten. Dann enthält C_n und somit auch seine Teilmenge C keine Intervalle der Länge ε. Also enthält C überhaupt kein Intervall mit mehr als einem Punkt und somit keine nichtleere offene Menge. Das zeigt $C^\circ = \emptyset$.

(c) Sei $c \in C$ und sei $\varepsilon > 0$. Wähle wie im Beweis der vorangehenden Teilaussage ein $n \in \mathbb{N}_0$ mit $\frac{1}{2^n} < \varepsilon$. Es gilt $c \in C_n$, also liegt c in einem der Intervalle, aus denen C_n besteht – nennen wir es I_n. Sei $d \in I_n$ einer der beiden Randpunkte von I_n und $d \neq c$. Weil die Länge von I_n höchstens $\frac{1}{2^n}$ und somit kleiner als ε ist, gilt $d \in B_{<\varepsilon}(c)$. Weil d ein Randpunkt von C_n ist, folgt aus der Konstruktion der Mengen C_m, dass $d \in C_m$ für alle $m \geq n$ gilt, denn in keinem Schritt wird von einem Intervall ein Randpunkt entfernt. Zugleich gilt $d \in C_n \subseteq C_{n-1} \subseteq \cdots \subseteq C_0$, also $d \in C$.

(d) Aus der Konstruktion der Mengen C_n folgt, dass $0 \in C_n$ für alle $n \in \mathbb{N}_0$ und somit $0 \in c$ gilt; also ist $C \neq \emptyset$.[22] Aus (c) folgt, dass der metrische Raum $(C, \mathrm{d}_{\mathrm{Eukl}})$ keine

[19] Man kann leicht überprüfen, dass das gleichbedeutend damit ist, dass im metrischen Raum $(C, \mathrm{d}_{\mathrm{Eukl}})$ kein Punkt isoliert im Sinne von Aufgabe 5.1 auf Hausaufgabenblatt 5 ist.

[20] Hier haben wir verwendet, dass Teilmengen in $(\mathbb{R}^d, \mathrm{d}_{\mathrm{Eukl}})$ genau dann kompakt sind, wenn sie beschränkt und abgeschlossen sind. Man kann aber auch ohne diese Charakterisierung argumentieren: Mit Hilfe der Definition von Kompaktheit lässt sich nämlich direkt zeigen, dass in einem metrischen Raum der Durchschnitt von beliebig vielen kompakten Mengen wieder kompakt ist.

[21] Es sind übrigens, wie man leicht sieht, genau 2^n Stück, aber das ist für den Rest des Arguments nicht wichtig.

[22] Allgemeiner folgt übrigens aus der Heine–Borel-Eigenschaft kompakter Mengen, dass der Durchschnitt einer absteigenden Folge kompakter nichtleerer Mengen immer nichtleer ist. Das impliziert ebenfalls $C \neq 0$, allerdings ist dieses allgemeine Argument in der vorliegenden Situation nicht nötig.

isolierten Punkte hat. Weil C laut (a) kompakt ist,[23] ist $(C, \mathrm{d}_{\mathrm{Eukl}})$ vollständig, also ist C laut Aufgabe 5.1(b) auf Hausaufgabenblatt 5 überabzählbar.

(e) Das Anwenden von S_{η_n} auf eine Menge verkleinert das Maß der Menge auf das η_n-fache. Also gilt für jedes $n \in \mathbb{N}$

$$\lambda_1(C_n) = \eta_n \lambda_1(C_{n-1}) = \eta_n \eta_{n-1} \lambda_1(C_{n-2}) = \cdots = \prod_{k=1}^{n} \eta_k \cdot \lambda_1(C_0) = \prod_{k=1}^{n} \eta_k.$$

Weil λ_1 von oben stetig ist (Proposition 1.2.3(e)) und $\lambda_1([0,1]) = 1 < \infty$ ist, folgt die Behauptung. □

Wenn als $(\eta_n)_{n\in\mathbb{N}}$ im vorangehenden Beispiel geeignete konkrete Folgen einsetzen, sieht man mehrere bemerkenswerte Phänomene:

Beispiele 1.5.14 (Topologische Größe – Maßtheoretische Größe – Kardinalität)**.** Alle topologischen Eigenschaften im Folgenden sind im metrischen Raum $(\mathbb{R}, \mathrm{d}_{\mathrm{Eukl}})$ gemeint.

(a) Es gibt eine überabzählbare Menge $C \subseteq [0,1]$, die nirgends dicht ist.

(b) Es gibt eine Borel-messbare Menge $C \subseteq [0,1]$, die überabzählbar ist, aber Lebesguemaß 0 besitzt.

(c) Es gibt eine Borel-messbare Menge $C \subseteq [0,1]$, die nirgends dicht ist, aber $\lambda_1(C) > 0$ erfüllt.

Mehr noch: Für jedes $\varepsilon > 0$ gibt es eine Borel-messbare Menge $C \subseteq [0,1]$, die nirgends dicht ist, aber $\lambda_1(C) \geq 1 - \varepsilon$ erfüllt.

(d) Es gibt eine Borel-messbare Menge $C \subseteq [0,1]$, die mager ist, aber $\lambda_1(C) = 1$ erfüllt.

Beweis. (a) Wir nehmen die Menge C aus Beispiel 1.5.13 für irgendeine beliebige Folge $(\eta_n)_{n\in\mathbb{N}}$. Diese Menge ist, wie in Beispiel 1.5.13 gezeigt, überabzählbar und nirgends dicht.

(b) Wir nehmen die Menge C aus Beispiel 1.5.13 mit $\eta_n := \frac{1}{2}$ für alle $n \in \mathbb{N}$. Dann ist C Borel-messbar (da kompakt) und überabzählbar und es gilt $\lambda_1(C) = \lim_{n\to\infty} \prod_{k=1}^{n} \frac{1}{2} = \lim_{n\to\infty} \frac{1}{2^n} = 0$.

(c) Sei $\varepsilon > 0$. Wir nehmen die Menge C aus Beispiel 1.5.13 für $\eta_n := 1 - \frac{\varepsilon}{2^n}$ für alle $n \in \mathbb{N}$. Wir zeigen per Induktion über n, dass $\prod_{k=1}^{n} \eta_k \geq 1 - \sum_{k=1}^{n} \frac{\varepsilon}{2^k}$ für alle $n \in \mathbb{N}$ gilt.

Für $n = 1$ ist die Behauptung klar. Ist die Behauptung für ein festes $n \in \mathbb{N}$ bereits bewiesen, dann folgt

$$\prod_{k=1}^{n+1} \eta_k \geq \big(1 - \sum_{k=1}^{n} \frac{\varepsilon}{2^k}\big)\big(1 - \frac{\varepsilon}{2^{n+1}}\big)$$

[23] Beachten Sie, dass Kompaktheit nicht von der umgebenenden Mengen abhängt – das heißt zu sagen, dass C im metrischen Raum $(\mathbb{R}, \mathrm{d}_{\mathrm{Eukl}})$ kompakt ist, ist äquivalent dazu zu sagen, dass C im metrischen Raum $(C, \mathrm{d}_{\mathrm{Eukl}})$ kompakt ist.

$$= 1 - \sum_{k=1}^{n} \frac{\varepsilon}{2^k} - \frac{\varepsilon}{2^{n+1}} + \Big(\sum_{k=1}^{n} \frac{\varepsilon}{2^k}\Big) \cdot \frac{\varepsilon}{2^{n+1}} \geq 1 - \sum_{k=1}^{n+1} \frac{\varepsilon}{2^k}.$$

Also ist der Induktionsbeweis erbracht. Insgesamt folgt

$$\lambda_1(C) = \lim_{n\to\infty} \prod_{k=1}^{n} \eta_k \geq \lim_{n\to\infty} \Big(1 - \sum_{k=1}^{n+1} \frac{\varepsilon}{2^k}\Big) = 1 - \sum_{k=1}^{\infty} \frac{\varepsilon}{2^k} = 1 - \varepsilon,$$

(d) Für jedes $n \in \mathbb{N}$ gibt es laut (c) eine Borel-messbare Menge $D_n \subseteq [0,1]$, die nirgends dicht ist und $\lambda_1(D_n) \geq 1 - \frac{1}{n}$ erfüllt. Wir setzen $C := \bigcup_{n\in\mathbb{N}} D_n$. Dann ist C Borel-messbar und mager und für jedes $n \in \mathbb{N}$ folgt aus $D_n \subseteq C \subseteq [0,1]$

$$1 - \frac{1}{n} \leq \lambda_1(D_n) \leq \lambda_1(C) \leq \lambda_1([0,1]) = 1,$$

also ist $\lambda_1(C) = 1$. □

Man kann sich fragen, ob man in Proposition 1.5.14(c) vielleicht sogar $\lambda_1(C) = 1$ erreichen kann. Das ist nicht der Fall, wie wir in einer Aufgabe auf Hausgabenblatt 6 zeigen werden.

Kapitel 2

Integration bezüglich Maßen

2.1 Messbarkeit skalarwertiger Funktionen

Einstiegsfragen. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

(a) Können Sie ein paar messbare Funktionen von Ω nach $\mathbb{R}$ angeben ohne mehr über Ω and $\mathcal{A}$ zu wissen?

(b) Wie kann man entscheiden, ob eine gegebene Funktion $f : \Omega \to \mathbb{R}$ messbar ist ohne sich jedes mal tot zu rechnen?

(c) Kann man messare Funktionen $f : \Omega \to \mathbb{R}$ durch besonders "einfache" Funktionen approximieren? Und was bedeutet überhaupt "approximieren"?

Wir sprechen in diesem Abschnitt über Messbarkeit von Funktionen; diesen Begriff hatten wir in Definition 1.1.10 eingeführt. Folgende Eigenschaft folgt leicht aus dieser Definition.

Proposition 2.1.1. *Seien $(\Omega_0, \mathcal{A}_0)$, $(\Omega_1, \mathcal{A}_1)$, $(\Omega_2, \mathcal{A}_2)$ Messräume und seien $f_1 : \Omega_0 \to \Omega_1$ und $f_2 : \Omega_1 \to \Omega_2$ messbar. Dann ist auch die Hintereinanderausführung $f_2 \circ f_1 : \Omega_0 \to \Omega_2$ messbar.*

Wenn zwei metrische Räume (M_1, d_1) und (M_2, d_2) jeweils mit der Borel-σ-Algebra ausgestattet werden, dann ist jede stetige Funktion $f : M_1 \to M_2$ messbar. Das haben Sie in Aufgabe 2.1(a) auf Hausaufgabenblatt 2 bewiesen.

Definition 2.1.2 (Zerlegung von skalarwertigen Funktionen)**.** Sei Ω eine Menge.

(a) Für zwei Zahlen $a, b \in \mathbb{R}$ verwendet man manchmal die Kurzschreibweise

$$a \vee b := \max\{a, b\} \qquad \text{und} \qquad a \wedge b := \min\{a, b\}.$$

(b) Für $f : \Omega \to \mathbb{R}$ definiert man die Funktionen $f^+, f^-, |f| : \Omega \to [0, \infty)$ durch

$$f^+(\omega) := f(\omega) \vee 0, \qquad f^-(\omega) := (-f(\omega)) \vee 0 = -(f(\omega) \wedge 0), \qquad |f|(\omega) := |f(\omega)|$$

für alle $\omega \in \Omega$.

(c) Für eine Funktion $g : \Omega \to \mathbb{C}$ definiert man die Funktionen $\operatorname{Re} g, \operatorname{Im} g : \Omega \to \mathbb{R}$ durch

$$(\operatorname{Re} g)(\omega) := \operatorname{Re}\big(g(\omega)\big) \qquad \text{und} \qquad (\operatorname{Im} g)(\omega) := \operatorname{Im}\big(g(\omega)\big)$$

für alle $\omega \in \Omega$.

Vorlesung 10 (Fr, 14.11.) beginnt mit Definition 2.1.2(d)

(d) Für $f_1, f_2 : \Omega \to \mathbb{R}$ definieren wir $f_1 \vee f_2 : \Omega \to \mathbb{R}$ und $f_1 \wedge f_2 : \Omega \to \mathbb{R}$ ebenfalls punktweise, also durch

$$(f_1 \vee f_2)(\omega) := f_1(\omega) \vee f_2(\omega) \qquad \text{und} \qquad (f_1 \wedge f_2)(\omega) := f_1(\omega) \wedge f_2(\omega)$$

für alle $\omega \in \Omega$.

Mit diesen Objekten kann man eine skalarwerige Funktion in positive Anteile zerlegen: Jedes $g : \Omega \to \mathbb{C}$ lässt sich schreiben als $g = \operatorname{Re} g + \mathrm{i} \operatorname{Im} g$, und jedes $f : \Omega \to \mathbb{R}$ lässt sich schreiben als $f = f^+ - f^-$.

Wir wollen nun über Messbarkeit skalarwertiger Funktionen sprechen.

Konvention 2.1.3. Wenn nichts anderes gesagt ist, statten wir $\mathbb{R}$ beziehungsweise $\mathbb{C}$ als Wertemengen von Funktionen stets mit der Euklidischen Metrik und der Borel-σ-Algebra $\mathcal{B}(\mathbb{R})$ beziehungsweise $\mathcal{B}(\mathbb{C})$ aus.

Für einen Messraum $(\Omega, \mathcal{A})$ ist eine Funktion $f : \Omega \to \mathbb{R}$ somit laut Definition 1.1.10 messbar, wenn $f^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{R})$ gilt. Analog dazu ist $g : \Omega \to \mathbb{C}$ genau dann messbar, wenn $g^{-1}(B) \in \mathcal{A}$ für alle $B \in \mathcal{B}(\mathbb{C})$ gilt.

Messbarkeit von Funktionen lässt sich laut Proposition 1.1.11 mit Hilfe eines Erzeuger der σ-Algebra im Wertebereich nachprüfen. Jedes der folgenden Mengensystem ist ein Erzeuger der Borel-σ-Algebra $\mathcal{B}(\mathbb{R})$:

$$\begin{aligned}
\mathcal{E}_0 &:= \{U \subseteq \mathbb{R} \mid U \text{ ist offen}\},\\
\mathcal{E}_1 &:= \{C \subseteq \mathbb{R} \mid C \text{ ist abgeschlossen}\},\\
\mathcal{E}_2 &:= \{K \subseteq \mathbb{R} \mid K \text{ ist kompakt}\},\\
\mathcal{E}_3 &:= \{(a,b) \mid a, b \in \mathbb{Q}\},\\
\mathcal{E}_4 &:= \{[a,b] \mid a, b \in \mathbb{Q}\},\\
\mathcal{E}_5 &:= \{(-\infty, b) \mid b \in \mathbb{Q}\},\\
\mathcal{E}_6 &:= \{(-\infty, b] \mid b \in \mathbb{Q}\},\\
\mathcal{E}_7 &:= \{(a, \infty) \mid a \in \mathbb{Q}\},\\
\mathcal{E}_8 &:= \{[a, \infty) \mid a \in \mathbb{Q}\}.
\end{aligned}$$

Für $\mathcal{E}_0$ ist das die Definition der Borel-σ-Algebra. Für manche der Mengen $\mathcal{E}_1, \dots, \mathcal{E}_8$ hatten wir die Eigenschaft $\sigma(\mathcal{E}_k) = \mathcal{B}(\mathbb{R})$ beispielsweise in den Übungen gezeigt, und für die verbleibenden funktioniert der Beweis ganz ähnlich.

Proposition 2.1.4 (Konsistenzeigenschaften skalarwertiger messbarer Funktionen)**.** *Sei* $(\Omega, \mathcal{A})$ *ein Messraum und seien* $f, g : \Omega \to \mathbb{R}$.

(a) *Sei $k \in \{0,1,\dots,8\}$ und sei $\mathcal{E}_k$ der Erzeuger von $\mathcal{B}(\mathbb{R})$ aus der Liste vor dieser Proposition. Die Funktion f is genau dann messbar, wenn $f^{-1}(E) \in \mathcal{A}$ für alle $E \in \mathcal{E}_k$ gilt.*

(b) *Sind f, g messbar, dann auch $\alpha f + \beta g : \Omega \to \mathbb{R}$ für alle $\alpha, \beta \in \mathbb{R}$.*

(c) *Sind f, g messbar, dann auch $fg : \Omega \to \mathbb{R}$.*

(d) *Sei (f_n) eine Folge von messbaren Funktionen $f_n : \Omega \to \mathbb{R}$, die punktweise gegen f konvergiert.*[1] *Dann ist auch f messbar.*

(e) *Ist f messbar, so sind auch f^+, f^- und $|f|$ messbar.*

Beweis. (a) Das ist ein Spezialfall von Proposition 1.1.11.

(b) Seien $\alpha, \beta \in \mathbb{R}$ und sei $\mathbb{R}^2$ mit der Borel-σ-Algebra ausgestattet. Dann ist

$$h := \begin{pmatrix} f \\ g \end{pmatrix} : \Omega \to \mathbb{R}^2, \qquad \omega \mapsto h(\omega) = \begin{pmatrix} f(\omega) \\ g(\omega) \end{pmatrix}$$

messbar; das werden Sie auf Hausaufgabenblat 6 zeigen.[2] Außerdem ist $\ell : \mathbb{R}^2 \to \mathbb{R}$, $x \mapsto \alpha x_1 + \beta x_2$ stetig und somit messbar. Die Funktion

$$\alpha f + \beta g = \ell \circ h$$

ist somit laut Proposition 2.1.1 ebenfalls messbar.

(c) Das zeigt man völlig analog zu (b). Man muss lediglich eine andere Funktion $\ell : \mathbb{R}^2 \to \mathbb{R}$ verwenden, nämlich $\ell(x) := x_1 x_2$.

(d) Sei $b \in \mathbb{R}$. Wir zeigen die Gleichheit

$$f^{-1}\big((-\infty, b]\big) = \underbrace{\bigcap_{k=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} f_m^{-1}\big((-\infty, b + \tfrac{1}{k}]\big)}_{=:A}.$$

Weil A ein Element von $\mathcal{A}$ ist, folgt dann die Behauptung, indem wir (a) für den Erzeuger $\mathcal{E}_6$ anwenden.

"$\subseteq$" Sei $\omega \in f^{-1}\big((-\infty, b]\big)$, das heißt, $f(\omega) \le b$. Sei $k \in \mathbb{N}$. Wegen $f_n(\omega) \to f(\omega)$ für $n \to \infty$ gibt es ein $n \in \mathbb{N}$ mit der Eigenschaft $f_m(\omega) \le b + \frac{1}{k}$ für alle $m \ge n$. Also haben wir für das gegebene k ein n gefunden mit $\omega \in \bigcap_{m=n}^{\infty} f_m^{-1}\big((-\infty, b + \frac{1}{k}]\big)$. Somit gilt $\omega \in A$.

"$\supseteq$" Sei $\omega \in A$. Wir betrachten ein beliebiges, festes $k \in \mathbb{N}$. Wegen $\omega \in A$ gilt $\omega \in \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} f_m^{-1}\big((-\infty, b + \frac{1}{k}]\big)$, also gibt es $n \in \mathbb{N}$ derart, dass für alle $m \ge n$ die Ungleichung $f_m(\omega) \le b + \frac{1}{k}$ gilt. Wegen $f_m(\omega) \to f(\omega)$ für $m \to \infty$ folgt $f(\omega) \le b + \frac{1}{k}$. Da $k \in \mathbb{N}$ beliebig war, muss somit $f(\omega) \le b$ gelten, also ist tatsächlich $\omega \in f^{-1}\big((-\infty, b]\big)$.

[1] Das heißt, für jedes $\omega \in \Omega$ gelte $f_n(\omega) \to f(\omega)$ für $n \to \infty$.

[2] Und dabei wird ganz wesentlich sein, dass laut Beispiel 1.4.9 $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ gilt.

(e) Den Beweis dieser Teilaussage stellen wir als Übungsaufgabe auf Präsenzblatt 6. □

Zur Entwicklung der Integrationstheorie ist es oft hilfreich Funktionen zuzulassen, die den Wert ∞ annehmen. Damit man beim Rechnen mit solchen Funktionen nicht zu Widersprüchen gelangt, tut man das üblicherweise nur für Funktionen, deren Werte ≥ 0 sind.[3] Damit wir auch für Funktionen, die nach $[0,\infty]$ abbilden, von Messbarkeit sprechen können, führen wir eine σ-Algebra auf dieser Menge ein.

Definition 2.1.5 (σ-Algebra auf $[0,\infty]$)**.** Auf der Menge $[0,\infty] := [0,\infty) \cup \{\infty\}$ definieren wir die σ-Algebra

$$\mathcal{B}\big([0,\infty]\big) := \sigma\Big(\big\{A \in 2^{[0,\infty]} \mid A \subseteq [0,\infty) \text{ und } A \in \mathcal{B}(\mathbb{R})\big\}\Big).$$

Beachten Sie unbedingt, dass hierbei die auf der Menge $[0,\infty]$ erzeugte σ-Algebra gemeint ist.

Von nun an werden wir $[0,\infty]$, sofern nichts anderes gesagt wird, stets mit der Borel-σ-Algebra $\mathcal{B}\big([0,\infty]\big)$ ausstatten.

Damit man beim Arbeiten mit messaren Funktionen nicht durcheinander kommt, sollte man die folgende Eigenschaft im Hinterkopf behalten: Sei $(\Omega,\mathcal{A})$ ein Messraum und sei $f : \Omega \to [0,\infty)$. Dann kann man sich überlegen, dass f genau dann messbar von $(\Omega,\mathcal{A})$ nach $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ ist, wenn f messbar von $(\Omega,\mathcal{A})$ nach $\big([0,\infty],\mathcal{B}([0,\infty])\big)$ ist. Es ist für Messbarkeitsfragen also egal, ob man ein $f : \Omega \to [0,\infty)$ als Funktion von Ω nach $\mathbb{R}$ oder als Funktion von Ω nach $[0,\infty]$ auffasst.

Konvention 2.1.6 (Operationen mit Funktionen, die nach $[0,\infty]$ abbilden)**.** Sei Ω eine Menge und seien $f,g : \Omega \to [0,\infty]$ und $\alpha \in [0,\infty]$. Dann definiert man die Funktionen

$$\alpha f, \qquad f+g, \qquad fg, \qquad f \vee g, \qquad f \wedge g$$

analog zum $\mathbb{R}$-wertigen Fall, also durch die entsprechenden punktweisen Operationen.

Bemerkung 2.1.7 (Konsistenzeigenschaften messbarer Funktionen mit Werten in $[0,\infty]$)**.** Völlig analoge Eigenschaften zu Proposition 2.1.4 kann man auch für Funktionen mit Werten in $[0,\infty]$ zeigen.

Definition 2.1.8 (Indikatorfunktionen)**.** Sei Ω eine Menge und $A \subseteq \Omega$. Dann nennt man die Funktion $\mathbb{1}_A : \Omega \to [0,\infty)$, die durch

$$\mathbb{1}_A(\omega) := \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \in A^{\complement} \end{cases}$$

gegeben ist, als die **Indikatorfunktion von** A.

[3] Wir verwenden dann die naheliegenden Konventionen $\infty + a := a + \infty := \infty$ für alle $a \in [0,\infty]$ und $\infty \cdot a := a \cdot \infty := \infty$ für alle $a \in (0,\infty)$. Außerdem haben wir bereits in Definition 1.4.1 $0 \cdot \infty := \infty \cdot 0 := 0$ vereinbart.

Wenn $(\Omega,\mathcal{A})$ ein Messraum und $A\subseteq\Omega$ ist, dann kann man sich leicht überlegen, dass $\mathbb{1}_A$ genau dann messbar ist, wenn $A\in\mathcal{A}$ gilt. Somit sind auch die Funkionen in der folgenden Definition messbar:

Definition 2.1.9 (Einfache Funktionen)**.** Sei $(\Omega,\mathcal{A})$ ein Messraum und seien $A_1,\dots,A_n\in\mathcal{A}$ sowie $\alpha_1,\dots,\alpha_n\in[0,\infty)$. Dann nennen wir die Funktion

$$\alpha_1\,\mathbb{1}_{A_1}+\cdots+\alpha_n\,\mathbb{1}_{A_n}:\Omega\to[0,\infty)$$

eine **einfache Funktion**.[4]

Beachte Sie, dass für zwei einfache Funktionen f,g auch $f\vee g$ und $f\wedge g$ einfache Funktionen sind.

Proposition 2.1.10 (Approximation messbarer Funktionen durch einfache Funktionen)**.** *Sei $(\Omega,\mathcal{A})$ ein Messraum und sei $f:\Omega\to[0,\infty]$ messbar. Dann gibt es eine monotone steigende[5] Folge von einfachen Funktionen $f_n:\Omega\to[0,\infty)$ mit der Eigenschaft $f_n(\omega)\to f(\omega)$ für $n\to\infty$ für jedes $\omega\in\Omega$.*

Beweis. Sei $n\in\mathbb{N}$ und sei $0=y_0<y_1<\cdots<y_{n2^n}=n$ eine äquidistante Unterteilung des Intervalls $[0,n]$ in $n2^n$ Teilintervalle der Länge $\frac{1}{2^n}$. Wir setzen $g_n:=f\wedge(n\,\mathbb{1}_\Omega)$ und

$$f_n:=\sum_{k=1}^{n2^n}y_{k-1}\,\mathbb{1}_{g_n^{-1}\left([y_{k-1},y_k)\right)}.$$

Dann kann man überprüfen, dass (f_n) monoton steigend ist und $f_n(\omega)\to f(\omega)$ für jedes $\omega\in\Omega$ gilt. □

2.2 Das Lebesgue-Integral für Funktionen nach $[0,\infty]$

Einstiegsfragen. (a) Sei # das Zählmaß auf $\mathbb{N}$ und sei $f:\mathbb{N}\to\mathbb{R}$. Was möchten Sie gerne unter dem Integral $\int_{\mathbb{N}}f\,\mathrm{d}\#$ verstehen?

(b) Bezeichne $\mu\in\mathcal{B}(\mathbb{R}^3)$ die Masseverteilung eines physikalischen Körpers. Was ist der Schwerpunkt des Körpers?

Für eine Menge Ω und zwei Funktionen $f,g:\Omega\to\mathbb{R}$ oder $f,g:\Omega\to[0,\infty]$ verwenden wir die Notation $f\le g$ um zu sagen, dass $f(\omega)\le g(\omega)$ für alle $\omega\in\Omega$ gilt.

Vorlesung 11 (Mi, 19.11.) beginnt mit Proposition 2.2.1

Proposition 2.2.1 (Das Integral einfacher Funktionen hängt nicht von ihrer Darstellung ab)**.** *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum und seien Mengen $A_1,\dots,A_n,B_1,\dots,B_m\in\mathcal{A}$ sowie Zahlen $\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_m\in[0,\infty)$.*

[4] Man kann auch einfache Funktionen definieren, bei denen die Koeffizienten $\alpha_1,\dots,\alpha_n$ negative Werte annehmen dürfen. Im Folgenden ist aber der Fall für uns wichtig, indem $\alpha_1,\dots,\alpha_n\ge 0$ gilt. Um diese Bedingung nicht jedes Mal dazu sagen zu müssen, haben wir den Begriff **einfache Funktion** hier nur für Koeffizienten in $[0,\infty)$ definiert.

[5] Damit meinen wir genauer gesprochen punktweise monoton steigend, das heißt, für jedes $\omega\in\Omega$ ist die Zahlenfolge $(f_n(\omega))$ monoton steigend.

(a) *Wenn* $\sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j} \le \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j}$ *gilt, dann ist*

$$\sum_{j=1}^{n} \alpha_j \mu(A_j) \le \sum_{j=1}^{m} \beta_j \mu(B_j).$$

(b) *Wenn* $\sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j} = \sum_{j=1}^{m} \beta_j \mathbb{1}_{B_j}$ *gilt, dann ist*

$$\sum_{j=1}^{n} \alpha_j \mu(A_j) = \sum_{j=1}^{m} \beta_j \mu(B_j).$$

Die Aussage Proposition von Proposition 2.2.1 ist anschaulich recht naheliegend. Sie formal zu beweisen ist aber ein wenig technisch; wir begnügen und deshalb damit, die Aussage in der Vorlesung mit einer Skizze zu motivieren. Proposition 2.2.1 zeigt, dass in der folgenden Definition das Integral einer einfachen Funktion wohldefiniert ist.

Definition 2.2.2 (Das Lebesgue-Integral positiver Funktionen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum.

(a) Sei $f : \Omega \to [0, \infty)$ eine einfache Funktion. Dann definieren wir

$$\int_\Omega f \, \mathrm{d}\mu = \sum_{j=1}^{n} \alpha_j \mu(A_j) \in [0, \infty],$$

wobei $A_1, \ldots, A_n \in \mathcal{A}$ und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ so gewählt sind, dass $f = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{A_j}$ gilt.

(b) Sei $g : \Omega \to [0, \infty]$ messbar. Dann nennt man

$$\int_\Omega g \, \mathrm{d}\mu := \sup \left\{ \int_\Omega f \, \mathrm{d}\mu \mid f : \Omega \to [0, \infty) \text{ ist eine einfache Funktion mit } f \le g \right\}.$$

das **Lebesgue-Integral von g bezüglich μ**.[6]

(c) Sei $g : \Omega \to [0, \infty]$. Man nennt g **integrierbar**, falls g messbar ist und $\int_\Omega g \, \mathrm{d}\mu < \infty$ gilt.

Bitte beachten Sie in Definition (b) folgendes: Wenn $f : \Omega \to [0, \infty)$ eine einfache Funktion ist, dann ist der Ausdruck $\int_\Omega f \, \mathrm{d}\mu$ auf zwei verschiedene Weisen definiert – einmal in Teil (a) und einmal in Teil (b) der Definition. Es folgt aber aus Proposition 2.2.1(a), dass man beides Mal denselben Wert erhält, also ist Teil (b) der Definition konsistent mit Teil (a).

Um Lebesgue-Integrale in einigen konkreten Beispielen ausrechnen zu können, benötigen wir zuerst ein paar nützliche Eigenschaften des Integrals.

Theorem 2.2.3 (Eigenschaft des Lebesgue-Integrals für Funktionen nach $[0, \infty]$)**.** *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Maßraum.*

[6] Man beachte, dass $\int_\Omega g \, \mathrm{d}\mu \in [0, \infty]$ gilt.

(a) *Sind $f, g : \Omega \to [0,\infty]$ messbar mit $f \le g$, dann gilt $\int_\Omega f \,\mathrm{d}\mu \le \int_\Omega g \,\mathrm{d}\mu$.*

(b) *Es gilt der **Satz von der monotonen Konvergenz**: Seien $g, f_n : \Omega \to [0,\infty]$ messbar für alle $n \in \mathbb{N}$. Wenn (f_n) monoton wachsend ist und punktweise gegen g konvergiert, dann gilt $\int_\Omega f_n \,\mathrm{d}\mu \to \int_\Omega g \,\mathrm{d}\mu$ für $n \to \infty$.*

(c) *Für alle messbaren Funktionen $f, g : [0,\infty] \to [0,\infty]$ und alle*[7] *$\alpha \in [0,\infty)$ gilt*

$$\int_\Omega \alpha f \,\mathrm{d}\mu = \alpha \int_\Omega f \,\mathrm{d}\mu \qquad \text{und} \qquad \int_\Omega f + g \,\mathrm{d}\mu = \int_\Omega f \,\mathrm{d}\mu + \int_\Omega g \,\mathrm{d}\mu.$$

Vor dem Beweis des Theorems besprechen wir ein paar erste Beispiele für Integrale bezüglich verschiedenen Maßen. Um die Integrale in den Beispielen zu berechnen, werden wir insbesondere den Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) verwenden.

Beispiele 2.2.4 (Einige Integrale positiver Funktionen)**.**

(a) Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a,b] \to [0,\infty)$ stetig. Dann

$$\int_{[a,b]} f \,\mathrm{d}\lambda_1 = \int_a^b f(x) \,\mathrm{d}x,$$

wobei $\int_a^b f(x) \,\mathrm{d}x$ das Riemann-Integral bezeichnet.

(b) Sei $n \in \mathbb{N}$, sei $\# : 2^{\{1,\dots,n\}} \to [0,\infty]$ das Zählmaß und $f : \{1,\dots,n\} \to [0,\infty)$. Dann gilt

$$\int_{\{1,\dots,n\}} f \,\mathrm{d}\# = \sum_{k=1}^{n} f(k).$$

(c) Bezeichne $\# : 2^{\mathbb{N}} \to [0,\infty]$ das Zählmaß. Für jedes $f : \mathbb{N} \to [0,\infty)$ gilt

$$\int_{\mathbb{N}} f \,\mathrm{d}\# = \sum_{k=1}^{\infty} f(k).$$

Beweis. (a) Für jedes $n \in \mathbb{N}_0$ sei $p^{(n)}$ eine Partition von $[a,b]$ in 2^n äquidistante Intervalle, das heißt,

$$p_k^{(n)} := a + \frac{k}{2^n}(b-a)$$

für alle $k \in \{0,1,\dots,2^n\}$. Außerdem sei für jedes $n \in \mathbb{N}$ und jedes $k \in \{1,\dots,2^n\}$ der Punkt $q_k^{(n)} \in [p_{k-1}^{(n)}, p_k^{(n)}]$ eine Stelle, an welcher $f|_{[p_{k-1}^{(n)}, p_k^{(n)}]}$ minimal ist; so eine Minimalstelle existiert, weil f stetig ist und das Intervall $[p_{k-1}^{(n)}, p_k^{(n)}]$ kompakt ist. Dann ist

[7] Man kann sich übrigens überlegen, dass dieselbe Aussage auch für $\alpha = \infty$ gilt. Wir lassen diesen Fall hier weg, weil er in dieser Konstellation kaum auftaucht.

$q^{(n)}$ eine Punktierung der Partition $p^{(n)}$. Für jedes $n \in \mathbb{N}$ betrachten wir die einfache Funktion

$$f_n := \sum_{k=1}^{n} f(q_k^{(n)})\, \mathbb{1}_{\left[p_{k-1}^{(n)}, p_k^{(n)}\right)} + f(b)\, \mathbb{1}_{\{b\}}.$$

Dann ist (f_n) eine monoton wachsende Folge einfacher Funktionen, die punktweise gegen f konvergiert. Somit gilt laut Satz von der monotonen Konvergenz (Theorem 2.2.3(b))

$$\int_{[a,b]} f \,\mathrm{d}\lambda_1 = \lim_{n\to\infty} \int_{[a,b]} f_n \,\mathrm{d}\lambda_1 = \lim_{n\to\infty} \underbrace{\sum_{k=1}^{n} \left(p_k^{(n)} - p_{k-1}^{(n)}\right) f(q_k^{(n)})}_{\text{Riemann-Summe } R(f,p^{(n)},q^{(n)})} = \int_a^b f(x)\,\mathrm{d}x.$$

(b) Es gilt $f = \sum_{k=1}^{n} f(k)\, \mathbb{1}_{\{k\}}$, das heißt f ist einfach. Aus Definition 2.2.2(a) erhält man somit

$$\int_{\mathbb{N}} f \,\mathrm{d}\# = \sum_{k=1}^{n} f(k) \#\big(\{k\}\big) = \sum_{k=1}^{n} f(k).$$

(c) Für jedes $n \in \mathbb{N}$ setzen wir $f_n := f\, \mathbb{1}_{\{1,\dots,n\}}$. Dann ist jedes f_n einfach, die Folge (f_n) ist monoton steigend und sie konvergiert punktweise gegen f. Wie in (b) sieht man, dass

$$\int_{\mathbb{N}} f_n \,\mathrm{d}\# = \sum_{k=1}^{n} f(k)$$

für alle $n \in \mathbb{N}$ gilt. Aus dem Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) folgt somit

$$\int_{\mathbb{N}} f \,\mathrm{d}\# = \lim_{n\to\infty} \int_{\mathbb{N}} f_n \,\mathrm{d}\# = \lim_{n\to\infty} \sum_{k=1}^{n} f(k) = \sum_{k=1}^{\infty} f(k). \qquad \square$$

Vielleicht haben diese Beispiele Sie überzeugt, dass der Satz von der monotonen Konvergenz in Theorem 2.2.3 recht nützlich sein kann. Dann beweisen wir das Theorem.

Vorlesung 12 (Fr, 21.11.) beginnt mit dem Beweis von Theorem 2.2.3

Beweis von Theorem 2.2.3. (a) Das folgt direkt aus Definition 2.2.2(b).

(b) *für endliches Integral:* Seien g und (f_n) wie in (b) und sei $\int_\Omega g \,\mathrm{d}\mu < \infty$.

Schritt 1: Wegen (a) ist die Folge $\left(\int_\Omega f_n \,\mathrm{d}\mu\right)$ monoton wachsend und durch $\int_\Omega g \,\mathrm{d}\mu$ nach oben beschränkt. Somit ist sie konvergent und erfüllt

$$\lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu \le \int_\Omega g \,\mathrm{d}\mu.$$

Schritt 2: Sei $g_0 : \Omega \to [0,\infty]$ einfach mit $g_0 \leq g$. Wegen der Definition von $\int_\Omega g \,\mathrm{d}\mu$ müssen wir nur noch die Ungleichung $\int_\Omega g_0 \,\mathrm{d}\mu \leq \lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu$ zeigen.

Dazu sei $\varepsilon \in (0,1]$ beliebig, fest. Für jedes $n \in \mathbb{N}$ betrachten wir die messbare Menge

$$E_n := \{\omega \in \Omega \mid f_n(\omega) \geq (1-\varepsilon) g_0(\omega)\} = \big(f_n - (1-\varepsilon) g_0\big)^{-1}\big([0,\infty]\big) \subseteq \Omega.$$

Da die Folge (f_n) monoton steigend ist, gilt $E_1 \subseteq E_2 \subseteq \ldots$. Für jedes ω gilt $f_n(\omega) \to g(\omega) \geq g_0(\omega)$ für $n \to \infty$, also ist $\bigcup_{n\in\mathbb{N}} E_n = \Omega$. Mit Hilfe der unteren Stetigkeit von Maßen und der Definition des Integrals von einfachen Funktionen kann man nun sofort nachrechnen, dass $\int_\Omega g_0 \,\mathbb{1}_{E_n} \,\mathrm{d}\mu \to \int_\Omega g_0 \,\mathrm{d}\mu$ gilt. Zugleich ist

$$f_n \geq (1-\varepsilon) g_0 \,\mathbb{1}_{E_n}$$

für jedes $n \in \mathbb{N}$ und somit

$$\int_\Omega f_n \,\mathrm{d}\mu \geq \int_\Omega (1-\varepsilon) g_0 \,\mathbb{1}_{E_n} \,\mathrm{d}\mu = (1-\varepsilon) \int_\Omega g_0 \,\mathbb{1}_{E_n} \,\mathrm{d}\mu \to (1-\varepsilon) \int_\Omega g_0 \,\mathrm{d}\mu;$$

für die letzte Gleichheit haben wir verwendet, dass man Skalare aus dem Integral einfacher Funktionen ziehen darf, was sofort aus der Definition des Integrals einfacher Funktionen folgt.

Insgesamt ist also $\lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu \geq (1-\varepsilon) \int_\Omega g_0 \,\mathrm{d}\mu$. Da $\varepsilon \in (0,1]$ beliebig war, folgt die Behauptung.

(c) Sei $\alpha \in [0,\infty)$ und $f : \Omega \to [0,\infty]$ messbar. Wenn f einfach ist, folgt $\int_\Omega \alpha f \,\mathrm{d}\mu = \alpha \int_\Omega f \,\mathrm{d}\mu$ direkt aus Definition 2.2.2(a).

Laut Proposition 2.1.10 gibt es eine Folge einfacher Funktionen (f_n), die punktweise gegen f konvergiert. Die Funktionen αf_n sind ebenfalls alle einfach und die Folge (αf_n) konvergiert punktweise gegen αf. Aus (b) folgt

$$\int_\Omega \alpha f \,\mathrm{d}\mu \overset{(b)}{=} \lim_{n\to\infty} \int_\Omega \alpha f_n \,\mathrm{d}\mu \overset{f_n \text{ einfach}}{=} \alpha \lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu \overset{(b)}{=} \alpha \int_\Omega f \,\mathrm{d}\mu.$$

Die Gleichheit für $f+g$ zeigt man analog. □

Der Satz von der monotonen Konvergenz zeigt, dass man Integrale und Grenzwert für monoton wachsende Folgen vertauschen darf. Für allgemeine Folgen ist dies nicht immer der Fall, wie das nächste Beispiel zeigt.

Beispiel 2.2.5 (Nicht-Vertauschbarkeit von Integral und Grenzwert)**.** Auf dem Maßraum $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ betrachten wir die Funktionenfolge $(f_n) = \big(n \,\mathbb{1}_{(0,1/n]}\big)$. Diese konvergiert punktweise gegen die konstante Nullfunktion. Es gilt aber

$$\int_\mathbb{R} f_n \,\mathrm{d}\lambda_1 = \int_\mathbb{R} n \,\mathbb{1}_{(0,1/n]} \,\mathrm{d}\lambda_1 = n\big(\frac{1}{n} - 0\big) = 1 \overset{n\to\infty}{\not\to} 0 = \int_\mathbb{R} 0 \,\mathrm{d}\lambda_1.$$

In Beispiel 2.2.5 gilt zumindest die Ungleichung

$$\int_{\mathbb{R}} \lim_{n\to\infty} f_n \,\mathrm{d}\lambda_1 \le \lim_{n\to\infty} \int_{\mathbb{R}} f_n \,\mathrm{d}\lambda_1.$$

Zum Abschluss dieses Abschnitts zeigen wir, dass diese Ungleichung immer stimme, wenn beide Grenzwerte existieren. Falls Sie nicht existieren, lässt sich die Aussage mit Hilfe des Limes Inferior formulieren. Wir wiederholen diesen Begriff nun kurz.

Für eine Folge (y_n) in $[0,\infty]$ ist der Limes Inferior $\liminf_{n\to\infty} y_n$ definiert als der kleinste Häufungspunkt der Folge in $[0,\infty]$. Man kann sich leicht überlegen, dass

$$\liminf_{n\to\infty} y_n = \lim_{k\to\infty} \lim_{\ell\to\infty} (y_k \wedge \cdots \wedge y_\ell)$$

gilt. Beachten Sie hierbei, dass für jedes feste $k \in \mathbb{N}$ die Folge $(y_k \wedge \cdots \wedge y_\ell)_{\ell\in\mathbb{N}}$ monoton fallend ist und somit gegen ein $z_k \in [0,\infty]$ konvergiert. Die Folge $(z_k)_{k\in\mathbb{N}}$ wiederum ist monoton steigend und somit ebenfalls in $[0,\infty]$ konvergent. Man kann den Limes Inferior also als einen doppelten Limes darstellen – deshalb ist der punktweise Limes inferior einer Folge messbarer Funktionen wieder messbar (vergleiche Bemerkung 2.1.7).

Lemma 2.2.6 (Fatou)**.** *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum und sei eine Folge messbarer Funktionen $f_n : \Omega \to [0,\infty]$ gegeben. Dann gilt*

$$\int_\Omega \liminf_{n\to\infty} f_n \,\mathrm{d}\mu \le \liminf_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu.$$

Beweis. Für alle $k,\ell \in \mathbb{N}$ mit $k \le \ell$ sei $g_{k,\ell} := f_k \wedge \cdots \wedge f_\ell$. Für jedes feste k ist die Folge $(g_{k,\ell})_{\ell\ge k}$ monoton fallend und somit punktweise konvergent gegen eine messbare Funktion $h_k : \Omega \to [0,\infty]$. Für $k,\ell \in \mathbb{N}$ mit $k \le \ell$ und für alle $j \in \{k,\dots,\ell\}$ gilt $h_k \le g_{k,\ell} \le f_j$, also

$$\int_\Omega h_k \,\mathrm{d}\mu \le \int_\Omega f_j \,\mathrm{d}\mu \qquad \text{und somit} \qquad \int_\Omega h_k \,\mathrm{d}\mu \le \int_\Omega f_k \,\mathrm{d}\mu \wedge \cdots \wedge \int_\Omega f_\ell \,\mathrm{d}\mu.$$

Grenzübergang für $\ell \to \infty$ liefert somit

$$\int_\Omega h_k \,\mathrm{d}\mu \le \lim_{\ell\to\infty} \Big(\int_\Omega f_k \,\mathrm{d}\mu \wedge \cdots \wedge \int_\Omega f_\ell \,\mathrm{d}\mu \Big).$$

Nun lassen wir noch $k \to \infty$ gehen und erhalten

$$\begin{aligned} \int_\Omega \liminf_{n\to\infty} f_n \,\mathrm{d}\mu &= \int_\Omega \lim_{k\to\infty} h_k \,\mathrm{d}\mu = \lim_{k\to\infty} \int_\Omega h_k \,\mathrm{d}\mu \\ &\le \lim_{k\to\infty} \lim_{\ell\to\infty} \Big(\int_\Omega f_k \,\mathrm{d}\mu \wedge \cdots \wedge \int_\Omega f_\ell \,\mathrm{d}\mu \Big) = \liminf_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu; \end{aligned}$$

für die zweite Gleichheit haben wir den Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) benutzt, der anwendbar ist, weil die Folge (h_k) monoton wachsend ist. □

2.3 Das Lebesgue-Integral für reell- und komplexwertige Funktionen

Jetzt [8] wollen wir das Lebesgue-Integral für Funktionen zu definieren, die auch negative oder komplexe Werte annehmen können. Dazu ist die folgende einfache Beobachtung nützlich.

Proposition 2.3.1 (Zerlegung in positive Teile und Integrierbarkeit)**.** *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Maßraum und seien* $f : \Omega \to \mathbb{R}$ *und* $g : \Omega \to \mathbb{C}$ *messbar.*

(a) *Die Funktion* $|f|$ *ist genau dann integrierbar, wenn* f^+ *und* f^- *integrierbar sind.*

(b) *Allgemeiner gilt: Die Funktion* $|g|$ *ist genau dann integrierbar, wenn die vier Funktionen* $(\operatorname{Re} g)^+$, $(\operatorname{Re} g)^-$, $(\operatorname{Im} g)^+$ *und* $(\operatorname{Im} g)^-$ *integrierbar sind.*

Vorlesung 13 (Mi, 26.11.) beginnt mit dem Beweis von Proposition 2.3.1

Beweis. (a) Wegen $|f| = f^+ + f^-$ folgt das aus Theorem 2.2.3(c).

(b) Die Behauptung folgt aus Theorem 2.2.3(a) und (c) wegen

$$\begin{aligned} |g| &\le \left|(\operatorname{Re} g)^+\right| + \left|(\operatorname{Re} g)^-\right| + \left|\mathrm{i}(\operatorname{Im} g)^+\right| + \left|\mathrm{i}(\operatorname{Im} g)^-\right| \\ &= (\operatorname{Re} g)^+ + (\operatorname{Re} g)^- + (\operatorname{Im} g)^+ + (\operatorname{Im} g)^- \le 4|g|. \end{aligned}$$

□

Definition 2.3.2 (Das Lebesgue-Integral skalarwertiger Funktionen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und seien $f : \Omega \to \mathbb{R}$ und $g : \Omega \to \mathbb{C}$.

(a) Die Funktion f heißt **integrierbar**, falls sie messbar ist und $\int_\Omega |f| \, \mathrm{d}\mu < \infty$ ist.[9] In diesem Fall nennt man

$$\int_\Omega f \, \mathrm{d}\mu := \int_\Omega f^+ \, \mathrm{d}\mu - \int_\Omega f^- \, \mathrm{d}\mu \in \mathbb{R}$$

das **Lebesgue-Integral von f bezüglich μ**.

(b) Allgemeiner definiert man: Die Funktion g heißt **integrierbar**, falls sie messbar ist und $\int_\Omega |g| \, \mathrm{d}\mu < \infty$ ist.[10] In diesem Fall nennt man

$$\int_\Omega g \, \mathrm{d}\mu := \int_\Omega (\operatorname{Re} g)^+ \, \mathrm{d}\mu - \int_\Omega (\operatorname{Re} g)^- \, \mathrm{d}\mu + \mathrm{i} \int_\Omega (\operatorname{Im} g)^+ \, \mathrm{d}\mu - \mathrm{i} \int_\Omega (\operatorname{Im} g)^- \, \mathrm{d}\mu \in \mathbb{C}$$

das **Lebesgue-Integral von g bezüglich μ**.

Bemerkung 2.3.3. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und sei $g : \Omega \to \mathbb{C}$. Aus Proposition 2.3.1 und Definition 2.3.2 erhält man, dass g genau dann integrierbar ist, wenn $\operatorname{Re} g$ und $\operatorname{Im} g$ integrierbar sind, und in diesem Fall gilt

$$\int_\Omega g \, \mathrm{d}\mu = \int_\Omega \operatorname{Re} g \, \mathrm{d}\mu + \mathrm{i} \int_\Omega \operatorname{Im} g \, \mathrm{d}\mu$$

[8] In diesem Abschnitt hatte ich leider vergessen, Einstiegsfragen für die Vorlesung vorzubereiten.

[9] Das heißt, falls f messbar und $|f|$ integrierbar ist.

[10] Das heißt, falls g messbar und $|g|$ integrierbar ist.

und somit

$$\operatorname{Re}\int_\Omega g\,\mathrm{d}\mu = \int_\Omega \operatorname{Re} g\,\mathrm{d}\mu \qquad \text{und} \qquad \operatorname{Im}\int_\Omega g\,\mathrm{d}\mu = \int_\Omega \operatorname{Im} g\,\mathrm{d}\mu.$$

Theorem 2.3.4 (Eigenschaften des Lebesgue-Integrals skalarwertiger Funktionen)**.** *Sei* $(\Omega,\mathcal{A},\mu)$ *ein Maßraum, seien* $f,g:\Omega\to\mathbb{C}$ *integrierbar, und seien* $\alpha,\beta\in\mathbb{C}$.

(a) *Es ist auch* $\alpha f+\beta g$ *integrierbar und es gilt*

$$\int_\Omega \alpha f+\beta g\,\mathrm{d}\mu = \alpha\int_\Omega f\,\mathrm{d}\mu+\beta\int_\Omega g\,\mathrm{d}\mu.$$

(b) *Sind* f *und* g *reellwertig mit* $f\le g$*, dann gilt* $\int_\Omega f\,\mathrm{d}\mu\le\int_\Omega g\,\mathrm{d}\mu$.

(c) *Es gilt die* ***Fundamentalabschätzung für Integrale****:* $\left|\int_\Omega f\,\mathrm{d}\mu\right|\le\int_\Omega |f|\,\mathrm{d}\mu$.

Beweis. (a) – *Integrierbarkeit:* Laut Proposition 2.1.4(b) ist wegen der Messbarkeit von f und g auch $\alpha f+\beta g$ messbar. Außerdem gilt $|\alpha f+\beta g|\le|\alpha||f|+|\beta||g|$, also folgt aus Theorem 2.2.3(a) und (c), dass auch $\alpha f+\beta g$ integrierbar ist.

(a) – *Additivität:* Wir nehmen an, dass f,g reell-wertig sind und zeigen $\int_\Omega f+g\,\mathrm{d}\mu=\int_\Omega f\,\mathrm{d}\mu+\int_\Omega g\,\mathrm{d}\mu$. Wegen Bemerkung 2.3.3 gilt dieselbe Aussage dann auch für den komplexwertigen Fall.

Hierzu machen wir zunächst folgende Beobachtung $(*)$: Seien $h_1,h_2,\ell_1,\ell_2:\Omega\to[0,\infty)$ integrierbar und seien $h_1-h_2=\ell_1-\ell_2$. Diese Differenz ist dann ebenfalls integrierbar (wie am Anfang des beweises bereits gezeigt) und es gilt

$$\int_\Omega h_1\,\mathrm{d}\mu-\int_\Omega h_2\,\mathrm{d}\mu=\int_\Omega \ell_1\,\mathrm{d}\mu-\int_\Omega \ell_2\,\mathrm{d}\mu.$$

Das sind man folgendermaßen: Wegen $h_1-h_2=\ell_1-\ell_2$ gilt $h_1+\ell_2=h_2+\ell_1$, also folgt aus Theorem 2.2.3(c)

$$\int_\Omega h_1\,\mathrm{d}\mu+\int_\Omega \ell_2\,\mathrm{d}\mu=\int_\Omega h_1+\ell_2\,\mathrm{d}\mu=\int_\Omega h_2+\ell_1\,\mathrm{d}\mu=\int_\Omega h_2\,\mathrm{d}\mu+\int_\Omega \ell_1\,\mathrm{d}\mu.$$

Weil alle auftauchenden Integrale endlich sind, ist die Beobachtung $(*)$ gezeigt.

Nun wenden wir $(*)$ auf die Gleichheit $(f+g)^+-(f+g)^-=(f^+-g^+)-(f^--g^-)$ an und erhalten somit

$$\begin{aligned}\int_\Omega f+g\,\mathrm{d}\mu &= \int_\Omega (f+g)^+\,\mathrm{d}\mu-\int_\Omega (f+g)^-\,\mathrm{d}\mu \overset{(*)}{=} \int_\Omega f^+ + g^+\,\mathrm{d}\mu-\int_\Omega f^- + g^-\,\mathrm{d}\mu\\ &= \int_\Omega f^+\,\mathrm{d}\mu+\int_\Omega g^+\,\mathrm{d}\mu-\int_\Omega f^-\,\mathrm{d}\mu-\int_\Omega g^-\,\mathrm{d}\mu=\int_\Omega f\,\mathrm{d}\mu+\int_\Omega g\,\mathrm{d}\mu.\end{aligned}$$

(a) – *Multiplikation mit Skalaren:* Wir betrachten die folgenden Fälle:

1. Fall: f ist reellwertig und $\alpha \geq 0$.

Dann ist $(\alpha f)^+ = \alpha f^+$ und $(\alpha f)^- = \alpha f^-$ und somit folgt

$$\int_\Omega \alpha f \,\mathrm{d}\mu = \int_\Omega \alpha f^+ \,\mathrm{d}\mu - \int_\Omega \alpha f^- \,\mathrm{d}\mu = \alpha\Big(\int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu\Big) = \alpha \int_\Omega f \,\mathrm{d}\mu,$$

wobei wir für die zweite Gleichheit Theorem 2.2.3(c). verwendet haben.

2. Fall: f ist reellwertig und $\alpha = -1$.

Wegen $(-f)^+ = f^-$ und $(-f)^- = f^+$ gilt

$$\int_\Omega -f \,\mathrm{d}\mu = \int_\Omega f^- \,\mathrm{d}\mu - \int_\Omega f^+ \,\mathrm{d}\mu = -\Big(\int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu\Big) = -\int_\Omega f \,\mathrm{d}\mu.$$

3. Fall: f ist reellwertig und $\alpha < 0$.

Dann folgt die Behauptung aus dem 1. und dem 2. Fall.

4. Fall: f ist komplexwertig und $\alpha \in \mathbb{C}$.

In diesem Fall folgt die Behauptung wegen der Gleichheit

$$\alpha f = \big(\operatorname{Re}\alpha \operatorname{Re} f - \operatorname{Im}\alpha \operatorname{Im} f\big) + \mathrm{i}\big(\operatorname{Re}\alpha \operatorname{Im} f + \operatorname{Im}\alpha \operatorname{Re} f\big)$$

aus dem bereits behandelten Fall, in dem f reellwertig und $\alpha \in \mathbb{R}$ ist und aus der bereits bewiesenen Additivität.

(b) Aus $f^+ - f^- = f \leq g = g^+ - g^-$ folgt $f^+ + g^- \leq g^+ + f^-$. Somit erhält man die Behauptung aus Theorem 2.2.3(a) und (c).

(c) Falls f reellwertig ist, folgt die Behauptung sofort aus

$$\left|\int_\Omega f \,\mathrm{d}\mu\right| = \left|\int_\Omega f^+ \,\mathrm{d}\mu - \int_\Omega f^- \,\mathrm{d}\mu\right| \leq \int_\Omega f^+ \,\mathrm{d}\mu + \int_\Omega f^- \,\mathrm{d} = \int_\Omega |f| \,\mathrm{d}\mu,$$

wobei wir im letzten Schritt Theorem 2.2.3(c) verwendet haben.

Sei nun f komplexwertig.[11] Es ist $\int_\Omega f \,\mathrm{d}\mu$ eine komplexe Zahl, das heißt wir können das Integral schreiben als $\int_\Omega f \,\mathrm{d}\mu = e^{\mathrm{i}\theta}\left|\int_\Omega f \,\mathrm{d}\mu\right|$ für ein $\theta \in [0, 2\pi)$. Es folgt

$$\begin{aligned}\left|\int_\Omega f \,\mathrm{d}\mu\right| &= \operatorname{Re}\left|\int_\Omega f \,\mathrm{d}\mu\right| = \operatorname{Re}\big(e^{-\mathrm{i}\theta}\int_\Omega f \,\mathrm{d}\mu\big) \\ &= \int_\Omega \operatorname{Re}\big(e^{-\mathrm{i}\theta} f\big)\,\mathrm{d}\mu \leq \int_\Omega \left|e^{-\mathrm{i}\theta} f\right| \mathrm{d}\mu = \int_\Omega |f| \,\mathrm{d}\mu,\end{aligned}$$

wobei wir für die Gleichheit zwischen den beiden Zeile Teilaussage (a) und Bemerkung 2.3.3 verwendet haben; für die Ungleichung in der zweiten Zeile haben wir außerdem (b) verwendet. □

[11] Der reellwertige Fall folgt auch als Spezialfall aus dem komplexwertigen, den wir nun beweisen. Wir haben den reellen Fall trotzdem zunächst separat aufgeschrieben, um zu zeigen, dass er noch etwas einfacher ist.

Aus den Beispielen 2.2.4 erhält man:

Beispiel 2.3.5 (Einige Integrale skalarwertiger Funktionen)**.**

(a) Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a,b] \to \mathbb{C}$ stetig. Dann ist f integrierbar und es gilt

$$\int_{[a,b]} f \, \mathrm{d}\lambda_1 = \int_a^b f(x) \, \mathrm{d}x,$$

wobei $\int_a^b f(x) \, \mathrm{d}x$ das Riemann-Integral bezeichnet.

(b) Sei $n \in \mathbb{N}$, sei $\# : 2^{\{1,\dots,n\}} \to [0,\infty]$ das Zählmaß und $f : \{1,\dots,n\} \to \mathbb{C}$. Dann ist f integrierbar und es gilt

$$\int_{\{1,\dots,n\}} f \, \mathrm{d}\# = \sum_{k=1}^{n} f(k).$$

(c) Bezeichne $\# : 2^{\mathbb{N}} \to [0,\infty]$ das Zählmaß und sei $f : \mathbb{N} \to \mathbb{C}$. Es ist f genau dann integrierbar, wenn $\sum_{k=1}^{\infty} \left| f(k) \right| < \infty$ gilt, und in diesem Fall ist

$$\int_{\mathbb{N}} f \, \mathrm{d}\# = \sum_{k=1}^{\infty} f(k).$$

Beachten Sie, dass Beispiel 2.3.5(c) unter anderem besagt: Eine Funktion $f : \mathbb{N} \to \mathbb{C}$ ist genau dann integrierbar bezüglich des Zählmaßes $\#$, wenn die Reihe $\left(\sum_{k=1}^{n} \left| f(k) \right| \right)_{n \in \mathbb{N}}$ absolut konvergiert.

Definition 2.3.6 (Zufallsvariablen und Erwartungswert)**.**

(a) Eine **Zufallsvariable** ist eine messbare Abbildung $f : \Omega \to \mathbb{C}$ für einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mu)$.[12]

(b) Sei $(\Omega, \mathcal{A}, \mu)$ ein Wahrscheinlichkeitsraum und $f : \Omega \to \mathbb{C}$ eine Zufallsvariable. Falls f integrierbar ist, nennt man $\mathbb{E}f := \int_\Omega f \, \mathrm{d}\mu$ den **Erwartungswert von f**.

Beispiele 2.3.7 (Einfaches Beispiel für einen Erwartungswert)**.** Wir werfen eine Münze 3 mal und betrachten zur mathematischen Modellierung die Menge $\Omega = \{Z, K\}^3$ und das Wahrscheinlichkeitsmaß $\mu : 2^\Omega \to [0,\infty]$ mit $\mu(A) = \frac{\#A}{2^3} = \frac{\#A}{8}$.

Sei $f : \{Z, K\} \to \mathbb{R}$ mit $f(Z) = -1$ und $f(K) = 1$. Wir betrachten die Zufallsvariable $g : \Omega \to \mathbb{R}$, $g(\omega) = \sum_{j=1}^{3} f(\omega_j)$. Anschaulich kann man das zum Beispiel so interprieren:

[12] In konkreten Situationen guckt man sich meist eher reellwertige Zufallsvariablen an. Die sind in dieser Definition natürlich als Spezialfall enthalten.

In der Stochastik würde man für ein Wahrscheinlichkeitsmaß eher $\mathbb{P}$ statt μ schreiben und Zufallsvariablen bezeichnet man eher mit Symbolen wie X, Y, Z anstelle von f, g, h.

Wenn Sie in Zahl werfen, müssen Sie einen Euro bezahlen und wenn Sie Kopf werfen, erhalten Sie einen Euro. Damit ist $g(\omega)$ Ihr Gesamtgewinn, wenn Sie in den drei Münzwürfen das Tupel $\omega \in \{K,Z\}^3$ werfen.

Der Erwartungswert $\mathbb{E}g$ von g beschreibt anschaulich, welchen Gesamtgewinn Sie "im Mittel" erwarten können.[13] Es gilt[14]

$$\begin{aligned}\mathbb{E}g = \int_\Omega g \,\mathrm{d}\mu &= \sum_{\omega\in\Omega} g(\omega)\mu\big(\{\omega\}\big)\\ &= \frac{1}{8}g\big((Z,Z,Z)\big)+\frac{1}{8}g\big((Z,Z,K)\big)+\frac{1}{8}g\big((Z,K,Z)\big)+\frac{1}{8}g\big((Z,K,K)\big)\\ &\quad+\frac{1}{8}g\big((K,Z,Z)\big)+\frac{1}{8}g\big((K,Z,K)\big)+\frac{1}{8}g\big((K,K,Z)\big)+\frac{1}{8}g\big((K,K,K)\big)\\ &=\frac{1}{8}\big(-3-1-1+1-1+1+1+3\big)=0.\end{aligned}$$

Dass $\mathbb{E}g = 0$ gilt, wird Sie vermutlich kaum überraschen.

An vorangehendem Beispiel können Sie bereits sehen, dass solche Rechnungen schnell sehr lästig und unübersichtlich werden können. Wenn Sie zum Beispiel eine Münze 10-mal werfen, dann enthalt $\Omega := \{Z,K\}^{10}$ bereits $2^{10} = 1024$ Elemente. Es gibt aber natürlich viele Methoden, um solche Summen zu vereinfachen.[15] Im Laufe der Veranstaltungen werden Sie auch noch Beispiel für Erwartungswerte sehen, bei denen Ω keine endliche Menge ist.

Theorem 2.3.8 (Satz von der dominierten Konvergenz)**.** *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum, sei eine Folge messbarer Funktionen $f_n : \Omega \to \mathbb{C}$ gegeben, die punktweise gegen eine $\mathbb{C}$-wertige Funktion konvergiert. Wenn es ein integrierbares $h : \Omega \to [0,\infty)$ gibt mit $|f_n| \le h$ für alle Indizes n, dann ist auch $\lim_{n\to\infty} f_n$ integrierbar und es gilt*

$$\int_\Omega \lim_{n\to\infty} f_n \,\mathrm{d}\mu = \lim_{n\to\infty}\int_\Omega f_n \,\mathrm{d}\mu.$$

Beweis. Sei $f : \Omega \to \mathbb{C}$ die Grenzfunktion der Folge (f_n). Für jedes $\omega \in \Omega$ gilt dann $|f(\omega)| = \lim_{n\to\infty}|f_n(\omega)| \le h(\omega)$. Also folgt aus der Integrierbarkeit von h die Intergrierbarkeit von f. Wir müssen als noch die behauptete Gleichheit zeigen. Wir tun dies in drei allgemeiner werdenden Fällen:

1. Fall: Alle f_n nehmen nur Werte in $[0,\infty)$ an.

Dann können wir das Lemma von Fatou (Lemma 2.2.6) auf (f_n) anwenden und erhalten

$$\int_\Omega f \,\mathrm{d}\mu \le \liminf_{n\to\infty}\int_\Omega f_n \,\mathrm{d}\mu.$$

[13] Diese Aussage kann man formalisieren und dann entsprechend begründen. Siehe dazu das starke Gesetz der großen Zahlen, das Sie in einer Stochastik-Vorlesung lernen können.

[14] Wobei Sie die Summenformel für das Integral in den Übungen beweisen werden.

[15] Insbesondere, wenn sogenannte **stochastische Unabhängigkeit** im Spiel ist – wir kommen darauf zurück.

Es gilt aber laut Annahme $f_n = |f_n| \le h$ für alle n: Also nehmen auch die Funktionen $h - f_n$ nur Werte in $[0,\infty)$ an und wir können das Lemma somit auch auf $(h - f_n)$ anwenden; das liefert

$$\begin{aligned}\int_\Omega h\,\mathrm{d}\mu - \int_\Omega f\,\mathrm{d}\mu &= \int_\Omega h - f\,\mathrm{d}\mu \le \liminf_{n\to\infty} \int_\Omega h - f_n\,\mathrm{d}\mu \\ &= \liminf_{n\to\infty}\Big(\int_\Omega h\,\mathrm{d}\mu - \int_\Omega f_n\,\mathrm{d}\mu\Big) = \int_\Omega h\,\mathrm{d}\mu + \liminf_{n\to\infty}\Big(-\int_\Omega f_n\,\mathrm{d}\mu\Big);\end{aligned}$$

für die ersten und die letzte Gleichheit haben wir die Integrierbarkeit von h verwendet, damit Theorem 2.3.4(a) anwendbar ist. Ebenfalls wegen der Integrierbarkeit von h können wir die Zahl $\int_\Omega h\,\mathrm{d}\mu$ auf beiden Seiten der Gleichung subtrahieren und dann mit -1 multiplizieren und erhalten somit

$$\int_\Omega f\,\mathrm{d}\mu \ge \limsup_{n\to\infty} \int_\Omega f_n\,\mathrm{d}\mu.$$

Also haben wir insgesamt

$$\limsup_{n\to\infty} \int_\Omega f_n\,\mathrm{d}\mu \le \int_\Omega f\,\mathrm{d}\mu \le \liminf_{n\to\infty} \int_\Omega f_n\,\mathrm{d},$$

was die Behauptung zeigt.

2. Fall: Alle f_n nehmen nur reelle Werte an.

Dann gilt $f_n^+, f_n^- \le |f_n| \le h$ für alle Indizes n. Außerdem gilt $f_n^+ \to f^+$ und $f_n^- \to f^-$. Also können wir den ersten Fall auf die Folgen (f_n^+) und (f_n^-) anwenden und erhalten somit

$$\int_\Omega f_n\,\mathrm{d}\mu = \int_\Omega f_n^+\,\mathrm{d}\mu - \int_\Omega f_n^-\,\mathrm{d}\mu \to \int_\Omega f^+\,\mathrm{d}\mu - \int_\Omega f^-\,\mathrm{d}\mu = \int_\Omega f\,\mathrm{d}\mu.$$

3. Fall: Die f_n dürfen komplexwertig sein. In diesem Fall teilen wir jedes f_n in seinen Real- und Imaginärteil auf und wenden Fall 2 an. □

2.4 Existenz von Produktmaßen und der Satz von Fubini

Vorlesung 14 (Fr, 28.11.) beginnt mit Abschnitt 2.4

Einstiegsfragen. (a) Kann man anschaulich das Volumen eines Kegels in $\mathbb{R}^3$ erhalten, indem man den Kegel und viele kleine Scheiben zerschneidet?

(b) Seien $(\Omega_1, \mathcal{A}_1, \mu_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2)$ zwei σ-endliche Maßräume und sei $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$. Warum gilt die Formel

$$\begin{aligned}&\int_{\Omega_1\times\Omega_2} \mathbb{1}_{A_1\times A_2}(\omega_1,\omega_2)\,\mathrm{d}(\mu_1\otimes\mu_2)(\omega_1,\omega_2) \\ &= \int_{\Omega_2}\int_{\Omega_1} \mathbb{1}_{A_1\times A_2}(\omega_1,\omega_2)\,\mathrm{d}\mu_1(\omega_1)\,\mathrm{d}\mu_2(\omega_2)?\end{aligned}$$

(c) Warum wurde in (b) vorausgesetzt, dass die Maßräume σ-endlich sind?

In diesem Abschnitt verfolgen wir zwei Ziele:

(1) Wir werden – zumindest in einem wichtigen Spezialfall – die Existenz von Produktmaßen zeigen (Theorem 2.4.3), die wir bisher in Theorem 1.4 nur behauptet hatten.

(2) Wir werden zwei wichtige Sätze beweisen, mit deren Hilfe man Integrale bezüglich Produktmaßen als iterierte Integrale bezüglich der einzelnen Maß berechnen kann (Theoreme 2.4.8 und 2.4.11).

Im Verlaufe des Abschnitts werden wir außerdem das Prinzip von Cavalieri beweisen (Korollar 2.4.5), welches recht anschaulich beschreibt, wie sich das Produktmaße einer messbaren Menge aus dem Maß von einzelnen **Mengenschnitten** zusammensetzt.

Definition 2.4.1 (Mengenschnitte)**.** Seien Ω_1, Ω_2 Mengen, sei $A \subseteq \Omega_1 \times \Omega_2$ und sei $\omega_1 \in \Omega_1$. Dann nennen wir

$$A_{\omega_1} := \{\omega_2 \in \Omega_2 \mid (\omega_1, \omega_2) \in A\} \subseteq \Omega_2$$

den **Schnitt von A an der Stelle ω_1**.

Lemma 2.4.2 (Messbarkeit von Mengenschnitten)**.** *Seien $(\Omega_1, \mu_1, \mathcal{A}_1)$ und $(\Omega_2, \mu_2, \mathcal{A}_2)$ Maßräume.*

(a) *Für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ und jedes $\omega_1 \in \Omega_1$ gilt $A_{\omega_1} \in \mathcal{A}_2$.*

(b) *Sei μ_2 nun σ-endlich. Für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ ist die Abbildung*

$$f_A : \Omega_1 \to [0, \infty], \qquad f_A(\omega_1) := \mu_2(A_{\omega_1})$$

messbar.

Beweis. Wir definieren das Mengensystem

$$\mathcal{Q} := \{A_1 \otimes A_2 \mid A_1 \in \mathcal{A}_1,\ A_2 \in \mathcal{A}_2\}$$

und erinnern daran, dass laut Definition der Produkt-σ-Algebra $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{Q})$ gilt.

(a) Wir fixieren ω_1 und müssen $A_{\omega_1} \in \mathcal{A}_2$ für alle $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ zeigen. Dazu verwenden wir das Prinzip der guten Mengen. Sei

$$\mathcal{B} := \{A \subseteq \Omega_1 \times \Omega_2 \mid A_{\omega_1} \in \mathcal{A}_2\}.$$

Für alle $A_1 \times A_2 \in \mathcal{Q}$ gilt $A_1 \times A_2 \in \mathcal{B}$, denn es ist

$$(A_1 \times A_2)_{\omega_1} = \begin{cases} A_2, & \text{falls } \omega_1 \in A_1, \\ \emptyset, & \text{falls } \omega_1 \notin A_1, \end{cases}$$

also in jedem Fall $(A_1 \otimes A_2)_{\omega_1} \in \mathcal{A}_2$. Außerdem überprüft man leicht, dass $\mathcal{B}$ eine σ-Algebra auf $\Omega_1 \times \Omega_2$ ist.[16]

Somit gilt $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{Q}) \subseteq \mathcal{B}$. Also gilt für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ auch $A \in \mathcal{B}$ und somit $A_{\omega_1} \in \mathcal{A}_2$.

Hinweis: In der Vorlesung haben wir Lemma 2.4.2(b) nur unter der Zusatzannahme $\mu_2(\Omega_2) < \infty$ bewiesen; in diesem Fall wird die Grundidee etwas klarer.

(b) Für jedes $E \in \mathcal{A}_2$ mit $\mu(E) < \infty$ und jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ betrachten wir die Abbildung

$$f_{A,E} : \Omega_1 \to [0,\infty), \qquad f_{A,E}(\omega_1) := \mu_2(A_{\omega_1} \cap E).$$

Wir fixieren nun E und zeigen, dass $f_{A,E}$ für jedes A messbar ist. Dazu verwenden wir wieder das Prinzip der guten Mengen. Sei dieses Mal

$$\mathcal{B} := \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 \mid f_{A,E} \text{ ist messbar}\}.$$

In dem man die σ-Additivität von μ_2 benutzt, kann man nachrechnen, dass $\mathcal{B}$ ein Dynkin-System auf $\Omega_1 \times \Omega_2$ ist. Um für $A \in \mathcal{B}$ die Eigenschaft $A^\complement \in \mathcal{B}$ zu seien, verwendet man die Annahme $\mu_2(E) < \infty$: Aus ihr erhält man $f_{A^\complement,E} = \mu_2(E)\,\mathbb{1}_\Omega - f_{A,E}$.

Außerdem kann man direkt überprüfen, dass $\mathcal{Q} \subseteq \mathcal{B}$ gilt. Somit ist

$$\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{Q}) = \delta(\mathcal{Q}) \subseteq \mathcal{B} \subseteq \mathcal{A}_1 \otimes \mathcal{A}_2,$$

wobei die zweite Gleichheit wegen der Stabilität von $\mathcal{Q}$ bezüglich endlicher Durchschnitte aus Theorem 1.2.12 folgt. Also gilt $\mathcal{A}_1 \otimes \mathcal{A}_2 = \mathcal{B}$.

Damit haben wir gezeigt, dass $f_{A,E}$ tatsächlich für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ und jedes $E \in \mathcal{A}_2$ mit $\mu_2(E) < \infty$ messbar ist. Nun verwenden wir die σ-Endlichkeit von μ_2. Wegen ihr gibt es eine Folge messbarer Mengen $E_1 \subseteq E_2 \subseteq \cdots \subseteq \Omega_2$ mit $\bigcup_{n=1}^\infty E_n = \Omega_2$ und $\mu_2(E_n) < \infty$ für alle $n \in \mathbb{N}$.

Sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Für jedes n ist die Funktion f_{A,E_n}, wie soeben gezeigt, messbar. Wegen der unteren Stetigkeit des Maßes μ_2 konvergiert die Folge (f_{A,E_n}) punktweise gegen f_A. Somit ist auch f_A messbar. □

Theorem 2.4.3 (Existenz von Produktmaßen im σ-endlichen Fall)**.** *Seien* $(\Omega_1, \mu_1, \mathcal{A}_1)$ *und* $(\Omega_2, \mu_2, \mathcal{A}_2)$ *Maßräume, wobei der zweite* σ*-endlich sei. Dann gibt ein Produktmaß* $\mu : \mathcal{A}_1 \otimes \mathcal{A}_2 \to [0,\infty]$ *von* μ_1 *und* μ_2.

Beweis. Für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ betrachten wir wie in Lemma 2.4.2(b) die Funktion

$$f_A : \Omega_1 \to [0,\infty], \qquad f_A(\omega_1) := \mu_2(A_{\omega_1}).$$

Laut Lemma 2.4.2(b) ist f_A messbar. Wir definieren nun $\mu : \mathcal{A}_1 \otimes \mathcal{A}_2 \to [0,\infty]$ durch $\mu(A) := \int_{\Omega_1} f_A \,\mathrm{d}\mu_1$ für alle $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Für alle $A_1 \in \mathcal{A}_1$ und $A_2 \in \mathcal{A}_2$ gilt

$$f_{A_1 \times A_2}(\omega_1) = \mu_2((A_1 \otimes A_2)_{\omega_1}) = \mu_2(A_2)\,\mathbb{1}_{A_1}(\omega_1)$$

[16] Um das im Detail nachzurechnen, überlegen Sie sich am besten zuerst die folgenden drei Beobachtungen: (a) Es gilt $\emptyset_{\omega_1} = \emptyset$. (b) Für jedes $A \subseteq \Omega_1 \times \Omega_2$ ist $(A^\complement)_{\omega_1} = (A_{\omega_1})^\complement$. (c) Für alle $A_1, A_2, \cdots \subseteq \Omega_1 \times \Omega_2$ gilt $\left(\bigcup_{n=1\infty} A_n\right)_{\omega_1} = \bigcup_{n=1}^\infty (A_n)_{\omega_1}$.

für alle $\omega_1 \in \Omega_1$. Also ist $f_{A_1 \times A_2} = \mu_2(A_2)\, \mathbb{1}_{A_1}$ und somit $\mu(A_1 \otimes A_2) = \mu_2(A_2) \int_{\Omega_1} \mathbb{1}_{A_1} \, \mathrm{d}\mu_1 = \mu_1(A_1)\mu_2(A_2)$.

Wir müssen noch nachprüfen, dass μ tatsächlich ein Maß ist. Wegen $f_\emptyset = \mu_2(\emptyset_{\omega_2}) = \mu_2(\emptyset) = 0$ gilt $\mu(\emptyset) = 0$. Seien nun $C_1, C_2, C_3 \cdots \in \mathcal{A}_1 \otimes \mathcal{A}_2$ paarweise disjunkt. Für jedes $\omega_1 \in \Omega_1$ gilt $(\bigcup_{n\in\mathbb{N}} C_n)_{\omega_1} = \bigcup_{n\in\mathbb{N}} (C_n)_{\omega_1}$ und die Mengen $(C_1)_{\omega_1}, (C_2)_{\omega_1}, (C_3)_{\omega_1}, \ldots$ sind paarweise disjunkte messbare Teilmengen von Ω_2. Somit gilt

$$f_{\bigcup_{n\in\mathbb{N}} C_n}(\omega_1) = \mu_2\Big(\bigcup_{n\in\mathbb{N}} (C_n)_{\omega_1}\Big) = \sum_{n=1}^{\infty} \mu_2\big((C_n)_{\omega_1}\big) = \sum_{n=1}^{\infty} f_{C_n}(\omega_1),$$

wobei wir für die mittlere Gleichheit die σ-Additivität von μ_2 verwendet haben. Also konvergiert die monoton wachsende Funktionenfolge $(\sum_{n=1}^{N} f_{C_n})_{N\in\mathbb{N}}$ punktweise gegen die Funktion $f_{\bigcup_{n\in\mathbb{N}} C_n}(\omega_1)$. Es folgt

$$\mu\Big(\bigcup_{n\in\mathbb{N}} C_n\Big) = \int_{\Omega_1} f_{\bigcup_{n\in\mathbb{N}} C_n} \, \mathrm{d}\mu_1 = \lim_{N\to\infty} \sum_{n=1}^{N} \int_{\Omega_1} f_{C_n} \, \mathrm{d}\mu_1 = \sum_{n=1}^{\infty} \mu(C_n),$$

wobei wir für die zweite Gleichheit den Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) verwendet haben. □

Aus der Eindeutigkeit des Produktmaßes σ-endlicher Maße (siehe Korollar 1.4.6) und aus dem Beweis von Theorem 2.4.3 erhält man sofort das nachfolgende Korollar.[17] Um es effizient aufschreiben zu können, ist die folgende Notation hilfreich.

Notation 2.4.4 (Integrationsvariablen)**.** Ein Lebesgueintegral $\int_\Omega f \, \mathrm{d}\mu$ schreibt man oft auch als $\int_\Omega f(\omega) \, \mathrm{d}\mu(\omega)$.[18] Man bezeichnet ω in dieser Schreibweise als **Integrationsvariable**.

Vorlesung 15 (Mi, 03.12.) beginnt mit Korollar 2.4.5

Korollar 2.4.5 (Cavalierisches Prinzip)**.** *Seien $(\Omega_1, \mu_1, \mathcal{A}_1)$ und $(\Omega_2, \mu_2, \mathcal{A}_2)$ σ-endlich Maßräume und sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$. Dann ist die Abbildung $\Omega_1 \ni \omega_1 \to \mu_2(A_{\omega_1}) \in [0,\infty]$ messbar und es gilt*

$$(\mu_1 \otimes \mu_2)(A) = \int_{\Omega_1} \mu_2(A_{\omega_1}) \, \mathrm{d}\mu_1(\omega_1).$$

In folgendem Beispiel veranschaulichen wir das Cavalierische Prinzip am Beispiel eines Kegelvolumes im $\mathbb{R}^3$. Bemerkenswerter Weise kann man die Grundfläche des Kegel extrem allgemein wählen und erhält immer dasselbe Ergebnis für das Volumen.

Beispiel 2.4.6 (Kegelvolumen)**.** Sei $G \subseteq \mathbb{R}^2$ Borel-messbar und sei $h \in (0,\infty)$. Wir betrachten die Menge

$$K := \big\{x \in \mathbb{R}^3 \mid \; x_1 \in [0,h], \; (x_2, x_3)^{\mathrm{T}} \in \tfrac{x_1}{h} G.\big\}.$$

[17] Eigentlich ist die Bezeichnung als "Korollar" aus terminologischer Sicht ein bisschen provokativ – kann es wirklich ein Korollar eines Beweises (statt eines Resultates) geben?

[18] Manchmal wird auch die Notation $\int_\Omega f(\omega)\mu(\mathrm{d}\omega)$ verwendet.

Also ist $K \subseteq \mathbb{R}^3$ ein auf der Seite liegender Kegel mit Grundfläche G und der Spitze im Ursprung.[19,20] Die Menge K ist Borel-messbar und es gilt $\lambda_3(K) = \frac{1}{3} h \lambda_2(G)$ – das heißt, das Volumen des Kegels ist gleich $\frac{1}{3}$ mal die Höhe mal die Grundfläche.

Beweis. Wir betrachten die stetige und somit Borel-messbare Abbildung $f : (0,h] \times \mathbb{R}^2 \to \mathbb{R}^2$, $x \mapsto \frac{h}{x_1}(x_2, x_3)^{\mathrm{T}}$. Es gilt

$$K = \{0\} \cup \left\{ x \in \mathbb{R}^3 \mid x_1 \in (0,h],\ (x_2,x_3)^{\mathrm{T}} \in \tfrac{x_1}{h} G. \right\} = \{0\} \cup f^{-1}(G),$$

also ist K als Vereinigung von zwei Borel-messbaren Mengen ebenfalls Borel-messbar.

Zur Berechnung des Volumens verwenden wir das Cavalierisches Prinzip (siehe Korollar 2.4.5), wobei wir $(\Omega_1, \mathcal{A}_1, \mu_1) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda_1)$ und $(\Omega_2, \mathcal{A}_2, \mu_2) = (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \lambda_2)$ wählen. Wegen $\mathcal{B}(\mathbb{R}^3) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^2)$ (Beispiel 1.4.9) und $\lambda_3 = \lambda_1 \otimes \lambda_2$ (Beispiel 1.4.10) können wir das Cavalierische Prinzip in der Tat anwenden und erhalten somit

$$\lambda_3(K) = (\lambda_1 \otimes \lambda_2)(K) = \int_{\mathbb{R}} \lambda_2(K_{x_1}) \, \mathrm{d}\lambda_1(x_1).$$

Für $x_1 \in \mathbb{R}$ gilt $K_{x_1} = \frac{x_1}{h} G$, falls $x_1 \in [0,h]$ ist, und $K_{x_1} = \emptyset$ sonst. Also ist

$$\lambda_2(K_{x_1}) = \lambda_2(\tfrac{x_1}{h} G) \, \mathbb{1}_{[0,h]}(x_1) = \lambda_2(G) \tfrac{x_1^2}{h^2} \, \mathbb{1}_{[0,h]}(x_1),$$

wobei wir verwendet haben, dass das zwei-dimensionale Lebesgue-Maß quadratisch mit Streckungen skaliert.[21] Insgesamt ist also

$$\begin{aligned} \lambda_3(K) &= \int_{\mathbb{R}} \lambda_2(G) \tfrac{x_1^2}{h^2} \, \mathbb{1}_{[0,h]}(x_1) \, \mathrm{d}\lambda_1(x_1) = \lambda_2(G) \frac{1}{h^2} \int_{[0,h]} x_1^2 \, \mathrm{d}\lambda_1(x_1) \\ &= \lambda_2(G) \frac{1}{h^2} \int_0^h x_1^2 \, \mathrm{d}x_1 = \lambda_2(G) \frac{1}{h^2} \cdot \frac{h^3}{3} = \frac{1}{3} \lambda_2(G) h; \end{aligned}$$

für die zweite Gleichheit haben wir Aufgabe 8.3(b) auf Hausaufgabenblatt 8 verwendet, für die dritte Gleichheit Beispiel 2.3.5(a), und für die letzte Gleichheit den Hauptsatz der Differential- und Integralrechnung aus der Analysis 1. □

Im Rest dieses Abschnitts wollen wir nun zeigen, wie man Integrale bezüglich eines Produktmaßes $\mu_1 \otimes \mu_2$ durch ein zweifaches Integral bezüglich der einzelnen Maße μ_1 und μ_2 ausdrücken kann. Unser Ziel ist es die unten stehenden Theorem 2.4.8 und 2.4.11 zu beweisen. Um die Aussagen dieser Theoreme überhaupt hinschreiben zu können, benötigen wir die folgende Beobachtung.

[19] Beachten Sie, dass die Grundfläche G möglicherweise gegenüber der Spitze verschoben ist.

[20] Die Annahmen an G hier sind sehr schwach, wodurch am auch sehr merkwürdige "Kegel" mit diesem Beispiel abdeckt. Wenn Sie Spaß an Abschweifungen haben, können Sie sich einmal überlegen, wie K aussieht, wenn G beispielsweise unbeschränkt ist. Oder wenn G ein Kreisring ist.

[21] Allgemeiner hatten wir in der Analysis 2 besprochen, dass für jede invertierbare Matrix $A \in \mathbb{R}^{d \times d}$ und jede Borel-messbare Menge $M \subseteq \mathbb{R}^d$ die folgenden Formel gilt:

$$\lambda_2\big(A(M)\big) = |\det(A)| \, \lambda_2(M)$$

Proposition 2.4.7 (Messbarkeit von Funktionenschnitten)**.** *Seien Maßräume* $(\Omega_1, \mu_1, \mathcal{A}_1)$ *und* $(\Omega_2, \mu_2, \mathcal{A}_2)$ *gegeben und sei* $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ *messbar bezüglich der Produkt-*σ*-Algebra* $\mathcal{A}_1 \otimes \mathcal{A}_2$*.*

(a) *Für jedes* $\omega_1 \in \Omega_1$ *ist die Abbildung* $f(\omega_1, \cdot) : \Omega_2 \to [0, \infty]$ *messbar.*

(b) *Sei* μ_2 *nun* σ*-endlich. Dann ist die Funktion*

$$g : \Omega_1 \to [0, \infty], \qquad g(\omega_1) := \int_{\Omega_2} f(\omega_1, \cdot)\, \mathrm{d}\mu_2 = \int_{\Omega_2} f(\omega_1, \omega_2)\, \mathrm{d}\mu_2(\omega_2)$$

messbar.

Beweis. (a) Sei $\omega_1 \in \Omega_1$ fest. Sei $\mathcal{F}$ die Menge aller messbaren Funktionen $f : \Omega_1 \times \Omega_2 \to [0, \infty]$, für die $f(\omega_1, \cdot) : \Omega_2 \to [0, \infty]$ messbar ist. Wir gehen in mehreren Schritten vor:

Schritt 1: Wir zeigen, dass $\mathbb{1}_A \in \mathcal{F}$ für jedes $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt.

Sei solch ein A gegeben. Natürlich ist $\mathbb{1}_A : \Omega_1 \times \Omega_2 \to [0, \infty]$ messbar. Für jedes $\omega_2 \in \Omega_2$ gilt

$$\mathbb{1}_A(\omega_1, \omega_2) = \begin{cases} 1, & \text{falls } \omega_2 \in A_{\omega_1}, \\ 0, & \text{falls } \omega_2 \notin A_{\omega_1}, \end{cases}$$

das heißt, es gilt $\mathbb{1}_A(\omega_1, \cdot) = \mathbb{1}_{A_{\omega_1}}$. Letztere Funktion ist messbar, denn laut Lemma 2.4.2(a) gilt $A_{\omega_1} \in \mathcal{A}_2$,

Schritt 2: Wir zeigen, dass $\mathcal{F}$ abgeschlossen ist bezüglich linear Kombinationen mit Koeffizienten in $[0, \infty)$. Seien dazu $f, g \in \mathcal{F}$ und $\alpha, \beta \in [0, \infty)$. Dann ist $\alpha f + \beta g : \Omega_1 \times \Omega_2 \to [0, \infty]$ messbar und es ist

$$(\alpha f + \beta g)(\omega_1, \cdot) = \alpha f(\omega_1, \cdot) + \beta g(\omega_1, \cdot) : \Omega_2 \to [0, \infty]$$

messbar. Also folgt $\alpha f + \beta g \in \mathcal{F}$.

Schritt 3: Wir bemerken, dass für jede Folge (f_n) in $\mathcal{F}$, die punktweise gegen eine Funktion f konvergiert, auch $f \in \mathcal{F}$ gilt.

In der Tat ist nämlich $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ messbar und es konvergiert $(f_n(\omega_1, \cdot))$ punktweise gegen die Funktion $f(\omega_1, \cdot) : \Omega_2 \to [0, \infty]$, also ist letztere Funktion messbar.

Schritt 4: Wir zeigen, dass jede messbare Funktion $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ in $\mathcal{F}$ liegt.

Wegen der Schritte 1 und 2 liegen alle einfachen Funktionen $\Omega_1 \times \Omega_2 \to [0, \infty]$ in $\mathcal{F}$. Wenn $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ eine messbare Funktion ist, dann lässt sich f laut Proposition 2.1.10 punktweise durch eine Folge einfacher Funktionen approximieren. Also gilt laut Schritt 3 auch $f \in \mathcal{F}$.

(b) Man kann den Beweis fast genauso wie in (a) führen, indem man die Menge $\mathcal{F}$ aller messbaren Funktionen $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ betrachtet, für die die Funktion

$$\Omega_1 \ni \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \cdot)\, \mathrm{d}\mu_2$$

messbar ist. In Schritt 1 vewendet man dann Lemma 2.4.2(b) statt Teil (a) desselben Lemmas. Die Anwendbarkeit dieses Resultats ist auch der einzige Grund, warum wir nun σ-Endlichkeit an μ_2 voraussetzen. Für Schritt 3 muss man den Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) verwenden. Die Schritte 2 und 4 funktionieren im Wesentlichen analog zu (a). □

Theorem 2.4.8 (Satz von Tonelli)**.** *Seien* $(\Omega_1, \mu_1, \mathcal{A}_1)$ *und* $(\Omega_2, \mu_2, \mathcal{A}_2)$ σ*-endliche Maßräume und sei* $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ *messbar bezüglich der Produkt-*σ*-Algebra* $\mathcal{A}_1 \otimes \mathcal{A}_2$*. Dann gilt*

$$\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2) \, \mathrm{d}(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) = \int_{\Omega_1} \int_{\Omega_2} f(\omega_1, \omega_2) \, \mathrm{d}\mu_2(\omega_2) \, \mathrm{d}\mu_1(\omega_1)$$

Beweis. Wenn $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ und $f = \mathbb{1}_A$ ist, dann gilt die Behauptung wegen des Cavalierischen Prinzips (Korollar 2.4.5), denn für jedes $\omega_1 \in \Omega_1$ ist $\mathbb{1}_A(\omega_1, \omega_2) = \mathbb{1}_{A_{\omega_1}}(\omega_2)$ und somit

$$\int_{\Omega_2} \mathbb{1}_A(\omega_1, \omega_2) \, \mathrm{d}\mu_2(\omega_2) = \mu_2(A_{\omega_1}).$$

Wegen der Linearität des Lebesgue-Integrals (Theorem 2.2.3(c)) gilt die Behauptung somit auch für alle einfachen Funktionen von $\Omega_1 \times \Omega_2$ nach $[0, \infty)$. Die Behauptung für alle messbaren Funktionen $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ folgt durch Approximation durch einfache Funktionen von unten (Proposition 2.1.10) nun durch dreifache Anwendung des Satzes von der monotonen Konvergenz (Theorem 2.2.3(b)). □

Lassen Sie uns den Satz von Tonelli an einem konkreten Beispiel demonstrieren. Wir werden nach dem Beispiel besprechen, dass dieses Beispiel auch eine Schwierigkeit aufzeigt, wenn man ein Analogon zum Satz von Tonelli für reell- oder komplex-wertige Funktionen beweisen will.

Beispiel 2.4.9 (Ein Integral mit Hilfe des Satzes von Tonelli berechnen)**.** Wir statten die Intervalle $[0, 1]$ und $[1, \infty)$ mit dem ein-dimensionalen Lebesgue-Maß λ_1 aus und betrachten die stetige Funktion $f : [0, 1] \times [1, \infty) \to [0, \infty)$, $f(\omega_1, \omega_2) := \omega_2^{-\sqrt{\omega_1}-1}$. Wir wollen

$$\int_{[0,1] \times [1,\infty)} f(\omega_1, \omega_2) \, \mathrm{d}\lambda_2(\omega_1, \omega_2)$$

berechnen und verwenden dazu den Satz von Tonelli (Theorem 2.4.8).

Für jedes festes $\omega_1 \in [0, 1]$ folgt aus dem Satz von der monotonen Konvergenz

$$\int_{[1,\infty)} f(\omega_1, \omega_2) \, \mathrm{d}\omega_2 = \lim_{n \to \infty} \int_{[1,\infty)} f(\omega_1, \omega_2) \, \mathbb{1}_{[1,n]}(\omega_2) \, \mathrm{d}\omega_2 = \lim_{n \to \infty} \int_1^n \omega_2^{-\sqrt{\omega_1}-1} \, \mathrm{d}\omega_2 =: (*).$$

Falls $\omega_1 \in (0, 1]$ ist, gilt

$$(*) = \lim_{n \to \infty} \frac{1}{-\sqrt{\omega_1}} \omega_2^{-\sqrt{\omega_1}} \Big|_{\omega_2 = 1}^{\omega_2 = n} = \frac{1}{\sqrt{\omega_1}}.$$

Falls hingegen $\omega_1 = 0$ ist, gilt

$$(*) = \lim_{n\to\infty} \log(\omega_2)\Big|_{\omega_2=1}^{\omega_2=n} = \infty.$$

Wenn wir der einfacheren Schreibweise $g : [0,1] \to [0,\infty]$, $g(\omega_1) = \frac{1}{\sqrt{\omega_1}}$ für $\omega_1 \in (0,1]$ und $g(0) = \infty$ setzen, gilt also insgesamt

$$\begin{aligned}
&\int_{[0,1]\times[1,\infty)} f(\omega_1,\omega_2)\,\mathrm{d}\lambda_2(\omega_1,\omega_2) = \int_{[0,1]}\int_{[1,\infty)} f(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2)\,\mathrm{d}\lambda_1(\omega_1)\\
&= \int_{[0,1]} g(\omega_1)\,\mathrm{d}\lambda_1(\omega_1) = \int_{[0,1]} g(\omega_1)\,\mathbb{1}_{\{0\}}\,\mathrm{d}\lambda_1(\omega_1) + \lim_{n\to\infty}\int_{[0,1]} g(\omega_1)\,\mathbb{1}_{[\frac{1}{n},1]}\,\mathrm{d}\lambda_1(\omega_1)\\
&= \int_{[0,1]} g(0)\,\mathbb{1}_{\{0\}}\,\mathrm{d}\lambda_1(\omega_1) + \lim_{n\to\infty}\int_{[\frac{1}{n},1]} g(\omega_1)\,\mathrm{d}\lambda_1(\omega_1)\\
&= \infty\cdot\lambda_1\big(\{0\}\big) + \lim_{n\to\infty}\int_{\frac{1}{n}}^{1}\frac{1}{\sqrt{\omega_1}}\,\mathrm{d}\omega_1 = 0\cdot\infty + \lim_{n\to\infty} 2\sqrt{\omega_1}\Big|_{\omega_1=\frac{1}{n}}^{\omega_1=1} = 2.
\end{aligned}$$

Vorlesung 16 (Fr, 05.12.) beginnt mit Beispiel 2.4.10

Beispiel 2.4.10 (Probleme mit der Integrierbarkeit)**.** Wir betrachten dieselbe Funktion $f : [0,1]\times[1,\infty) \to [0,\infty)$ wie in Beispiel 2.4.9. Wir definieren eine neue Funktion $h : [0,1]\times[1,\infty) \to \mathbb{C}$, die sich bis auf den komplexen Winkel genauso verhält wie f: wir setzen $h(\omega_1,\omega_2) := \exp(\mathrm{i}(\omega_1+\omega_2))f(\omega_1,\omega_2)$ für alle $(\omega_1,\omega_2) \in [0,1]\times[1,\infty)$.[22] Um über das Integral von h sprechen zu können, müssen wir uns laut Definition 2.3.2 zuerst überlegen, ob h integrierbar ist. Da h stetig ist, ist es messbar, und wegen $|h| = f$ folgt aus Beispiel 2.4.9, dass h tatsächlich integrierbar ist.

Wir würden nun gerne die Formel aus dem Satz von Tonelli (Theorem 2.4.8) auch für die Funktion h anwenden und das Integral

$$\int_{[0,1]\times[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_2(\omega_1,\omega_2)$$

in der Form $\int_{[0,1]}\int_{[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2)\,\mathrm{d}\lambda_1(\omega_1)$ darstellen um es dann zu berechnen. Noch bevor wir anfangen nun zu überlegen, ob die beiden Ausdrücke tatsächlich gleich sind, stellen wir aber fest, dass das Doppelintegral $\int_{[0,1]}\int_{[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2)\,\mathrm{d}\lambda_1(\omega_1)$ gar keinen Sinn ergibt, weil das innere Integral $\int_{[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2)$ nicht für jedes $\omega_1 \in [0,1]$ definiert ist. Für $\omega_1 = 0$ ist nämlich

$$\int_{[1,\infty)} |h(0,\omega_2)|\,\mathrm{d}\lambda_1(\omega_2) = \int_{[1,\infty)} f(0,\omega_2)\,\mathrm{d}\lambda_1(\omega_2) = \infty,$$

wie wir in Beispiel 2.4.9 berechnet haben. Somit ist die Funktion $h(0,\cdot) : [1,\infty) \to \mathbb{C}$ nicht integrierbar, das heißt, wir dürfen das innere Integral $\int_{[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2)$ gar nicht hinschreiben, falls $\omega_1 = 0$ ist.

[22] Das h nach $\mathbb{C}$ abbildet ist hier übrigens überhaupt nicht wichtig. Man kann dasselbe Funktion wie in diesem Beispiel auch leicht mit einer $\mathbb{R}$-wertigen Funktion erzeugen – eine $\mathbb{C}$-fertige Funktion verwenden wir hier nur, weil die Formel damit etwas einfacher ist und wir uns somit mehr auf das Wesentliche konzentrieren können.

Beispiel 2.4.10 erklärt, warum das folgende Theorem etwas technischer wirkt als Theorem 2.4.8. Zur Formulierung des Theorems verwenden wir, dass man σ-Algebren, Maß und Funktionen auf messbare Teilmengen eines Maßraums einschränken kann (Aufgabe 6.2 auf Hausaufgabenblatt 6 und Aufgabe 8.3 auf Hausaufgabenblatt 8).

Theorem 2.4.11 (Satz von Fubini)**.** *Seien $(\Omega_1, \mu_1, \mathcal{A}_1)$ und $(\Omega_2, \mu_2, \mathcal{A}_2)$ σ-endliche Maßräume, sei $\Omega_1 \times \Omega_2$ mit der Produkt-σ-Algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ und dem Produktmaß $\mu_1 \otimes \mu_2$ ausgestattet, und sei $f : \Omega_1 \times \Omega_2 \to \mathbb{C}$ integrierbar. Dann gibt es eine messbare Menge $N_1 \subseteq \Omega_1$ mit Maß $\mu_1(N_1) = 0$ derart, dass $f(\omega_1, \cdot) : \Omega_2 \to \mathbb{C}$ für jedes $\omega_1 \in \Omega_1 \setminus N_1$ integrierbar ist. Für jede solche*[23] *Menge N_1 ist die Funktion*

$$\Omega_1 \setminus N_1 \to \mathbb{C}, \qquad \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2)\, \mathrm{d}\mu_2(\omega_2)$$

integrierbar und erfüllt

$$\int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2)\, \mathrm{d}(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) = \int_{\Omega_1 \setminus N_1} \int_{\Omega_2} f(\omega_1, \omega_2)\, \mathrm{d}\mu_2(\omega_2)\, \mathrm{d}\mu_1(\omega_1).$$

Wir weisen kurz darauf hin, dass wir die rechte Seite der Gleichung in Theorem 2.4.11 etwas unpräzise aufgeschrieben haben. Genau genommen hätten wir dort den Ausdruck

$$\int_{\Omega_1 \setminus N_1} \int_{\Omega_2} f|_{\Omega_1 \setminus N_1}(\omega_1, \omega_2)\, \mathrm{d}\mu_2(\omega_2)\ \mathrm{d}\mu_1|_{\Omega_1 \setminus N_1}(\omega_1)$$

hinschreiben müssen. Das lässt die Formel aber komplizierter erscheinen, als sie eigentlich ist, und die Restriktionen von f und μ in der Notation einfach zu unterdrücken, birgt in dieser Situation kaum Verwirrungsgefahr – deshalb haben wir die Restriktionen in der Formulierung des Theorems einfach unter den Tisch fallen lassen.

Beweis von Theorem 2.4.11. Wir definieren die Funktion $g : \Omega_1 \to [0, \infty]$ durch die Formel $g(\omega_1) := \int_{\Omega_2} \big|f(\omega_1, \omega_2)\big|\, \mathrm{d}\mu_2(\omega_2)$; diese ist laut Proposition 2.4.7(b) messbar. Wir wenden den Satz von Tonelli (Theorem 2.4.8) auf die Funktion $\big|f\big|$ an und erhalten somit

$$\int_{\Omega_1} g(\omega_1)\, \mathrm{d}\mu_1(\omega_1) = \int_{\Omega_1 \times \Omega_2} \big|f(\omega_1, \omega_2)\big|\, \mathrm{d}(\mu_1 \otimes \mu_2)(\omega_1, \omega_2) < \infty,$$

da wir f als integrierbar vorausgesetzt haben. Lassen Sie uns die messbare Menge $N_1 := \{\omega_1 \in \Omega_1 \mid g(\omega_1) = \infty\}$ betrachten. Für jedes $n \in \mathbb{N}$ gilt

$$n\mu_1(N_1) = \int_{\Omega_1} n\, \mathbb{1}_{N_1}\, \mathrm{d}\mu_1 \leq \int_{\Omega_1} g\, \mathbb{1}_{N_1}\, \mathrm{d}\mu_1 \leq \int_{\Omega_1} g\, \mathrm{d}\mu_1,$$

[23] Sie fragen sich vielleicht – zurecht –, weshalb dieser Satz nicht einfach mit den Worten "für die Menge N_1" beginnt, sondern mit "für jede solche Menge N_1". Der Grund ist, dass N_1 nicht eindeutig bestimmt sein muss. Um den Satz möglichst einfach anwenden zu können, ist es aber gut, wenn man nur irgendeine Menge N_1 mit der gewünschten Eigenschaft zu finden braucht, und dann direkt mit ihr arbeiten kann.

also $\mu_1(N_1) \leq \frac{1}{n}\int_\Omega g\,\mathrm{d}\mu_1$; der rechtsstehende Ausdruck konvergiert gegen 0 für $n \to \infty$, weil $\int_\Omega g\,\mathrm{d}\mu_1 < \infty$ ist, also folgt $\mu_1(N_1) = 0$.[24]

Sei nun $N_1 \subseteq \Omega_1$ eine Menge mit der im Theorem angegebenen Eigenschaft. Dann kann man Ω_1 durch $\Omega_1 \setminus N_1$ ersetzen, die Funktion f in ihre positiven Anteile zerlegen[25] und auf diese vier Anteile den Satz von Tonelli (Theorem 2.4.8) anwenden. □

Übrigens wurden Sie mit Beispiel 2.4.10 in gewissem Sinne übers Ohr gehauen: Denn auch jetzt, wo wir eine korrekte Version von Theorem 2.4.11 gefunden haben, können wir das Integral $\int_{[0,1]\times[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_2(\omega_1,\omega_2)$ aus Beipiel 2.4.10 nicht wirklich ausrechnen, weil wir nämlich das innere Integral

$$\int_{[1,\infty)} h(\omega_1,\omega_2)\,\mathrm{d}\lambda_1(\omega_2) = \int_{[1,\infty)} \exp(\mathrm{i}(\omega_1+\omega_2))\omega_2^{-\sqrt{\omega_1}-1}\,\mathrm{d}\lambda_1(\omega_2)$$

aufgrund der recht komplizierten Form das Integranden nicht explizit ausrechnen können. Der eigentliche Grund, weshalb wir Beispiel 2.4.10 besprochen haben, war um zu sehen, dass man im Satz von Fubini die Ausnahmemenge N_1 benötigt.

2.5 Bildmaße und der Transformationssatz

Einstiegsfragen. (a) Was besagt noch einmal der Transformationssatz für Integrale aus der Analysis 2?

(b) Was stellen Sie sich unter dem Begriff "Wahrscheinlichkeitsverteilung" vor?

(c) Nochmal eine Widerholungsfrage zur Analysis 2: Was hat die Determinante von Matrizen mit dem Lebesgue-Maß zu tun?

In diesem Abschnitt besprechen wir **Bildmaße**. Sie kennen Sie bereits aus Aufgabe 3.3 auf Präsenzblatt 3 und aus Aufgabe 7.3 auf Hausaufgabenblatt 7. Damit alle wichtigen Begriffe und Resultate an einem Ort versammelt sind, wiederholen wir die entsprechenden Konzepte hier noch einmal kurz und erweitern sie noch um einige Eigenschaften und Begriffe.

Proposition 2.5.1 (Bildmaße)**.** *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum, sei $(\widehat{\Omega},\widehat{\mathcal{A}})$ ein Messraum und sei $f:\Omega\to\widehat{\Omega}$ messbar. Wir betrachten die Abbildung*

$$f_*\mu : \widehat{\mathcal{A}} \to [0,\infty], \qquad \widehat{A} \mapsto (f_*\mu)(\widehat{A}) := \mu\big(f^{-1}(\widehat{A})\big).$$

Dann ist $f_*\mu$ ein Maß.

[24] Hier könnte man stattdessen auch etwas kürzer argumentieren, dass

$$\infty\cdot\mu_1(N_1) = \int_{\Omega_1} \infty\cdot\mathbb{1}_{N_1}\,\mathrm{d}\mu_1 = \int_{\Omega_1} g\,\mathbb{1}_{N_1}\,\mathrm{d}\mu_1 \leq \int_\Omega g\,\mathrm{d}\mu_1 < \infty$$

gilt. Dann muss man sich allerdings kurz überlegen, weshalb die ersten Gleichheit gilt (denn $\infty\cdot\mathbb{1}_{N_1}$ ist laut Definition 2.1.9 keine einfache Funktion).

[25] Also in f^+ und f^-, falls f reellwertig ist, und in $(\mathrm{Re}\,f)^+$, $(\mathrm{Re}\,f)^-$, $(\mathrm{Im}\,f)^+$ und $(\mathrm{Im}\,f)^-$, falls f komplexwertig ist.

Beweis. Das haben Sie in Aufgabe 3.3(b) auf Präsenzblatt 3 bewiesen. □

Definition 2.5.2 (Bildmaß)**.** In der Situation von Proposition 2.5.1 nennt man $f_*\mu$ das **Bildmaß von μ unter f.**[26]

Beispiele 2.5.3 (Einige Bildmaße)**.** (a) Sei $A \in \mathbb{R}^{d\times d}$ eine invertierbare Matrix, die wir mit der Abbildung $\mathbb{R}^d \to \mathbb{R}^d$, $x \mapsto Ax$ identifizieren. Dann gilt $(A_*\lambda_d)(S) = \frac{1}{|\det A|}\lambda_d(S)$ für alle $S \in \mathcal{B}(\mathbb{R}^d)$.

(b) Sei $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$. Dann gilt für jedes $A \in \mathcal{B}(\mathbb{R})$

$$(f_*\lambda_1)(A) = 2\lambda_1\big(A \cap [0,\infty)\big).$$

Beweis. (a) In der Analysis 2 haben wir besprochen, dass $\lambda_d\big(A(S)\big) = |\det A|\,\lambda_d(S)$ für jede Menge $S \in \mathcal{B}(\mathbb{R}^d)$ gilt. Indem wir das auf die Matrix A^{-1} anwenden und $1 = |\det(\mathrm{id})| = |\det A|\,\big|\det(A^{-1})\big|$ verwenden, folgt die Behauptung.

(b) Sei $A \in \mathcal{B}(\mathbb{R})$. Dann ist

$$f^{-1}(A) = f^{-1}\big(A \cap [0,\infty)\big) = A \cap [0,\infty) \cup -\big(A \cap [0,\infty)\big) = A \cap [0,\infty) \cup -\big(A \cap (0,\infty)\big).$$

Die Mengen $A \cap [0,\infty)$ und $-\big(A \cap (0,\infty)\big)$ sind disjunkt und haben wegen $\lambda_1(\{0\}) = 0$ und wegen Aufgabe 3.4(b) auf Präsenzblatt 3 dasselbe eindimensionale Lebesgue-Maß. Also folgt die Behauptung. □

Theorem 2.5.4 (Transformationssatz für Bildmaße)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $(\widehat{\Omega}, \widehat{\mathcal{A}})$ ein Messraum und sei $f : \Omega \to \widehat{\Omega}$ messbar.*

(a) *Für jede messbare Funktion $g : \widehat{\Omega} \to [0,\infty]$ gilt*

$$\int_{\widehat{\Omega}} g \,\mathrm{d}(f_*\mu) = \int_\Omega g \circ f \,\mathrm{d}\mu.$$

(b) *Sei $g : \widehat{\Omega} \to \mathbb{C}$ messbar. Die Funktion g ist genau dann bezüglich $f_*\mu$ integrierbar, wenn $g \circ f : \Omega \to \mathbb{C}$ bezüglich μ integrierbar ist, und in diesem Fall gilt*

$$\int_{\widehat{\Omega}} g \,\mathrm{d}(f_*\mu) = \int_\Omega g \circ f \,\mathrm{d}\mu.$$

Beweis. (a) Das haben Sie in Aufgabe 7.3(b) auf Hausaufgabenblatt 7 bewiesen.

(b) Wegen $\big|g\big| \circ f = \big|g \circ f\big|$ folgt aus (a)

$$\int_{\widehat{\Omega}} \big|g\big| \,\mathrm{d}(f_*\mu) = \int_\Omega \big|g \circ f\big| \,\mathrm{d}\mu.$$

Somit ist tatsächlich g ist genau dann bezüglich $f_*\mu$ integrierbar, wenn $g \circ f : \Omega \to \mathbb{C}$ bezüglich μ integrierbar ist. Durch Zerlegung von g in seine vier positiven Anteile erhält man die in (b) behauptete Gleichheit der Integrale ebenfalls aus (a). □

[26] Oder, falls Sie Anglizismen mögen, das **Pushforward-Maß von μ unter f**.

Ein besonders wichtiges Beispiel für Bildmaße sind Wahrscheinlichkeitsverteilungen.

Definition 2.5.5 (Wahrscheinlichkeitsverteilungen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Wahrscheinlichkeitsraum und sei $f : \Omega \to M \subseteq \mathbb{R}$ eine Zufallsvariable. Dann bezeichnet man das Bildmaß $f_*\mu : \mathcal{B}(\mathbb{R}) \to [0, \infty)$ als **Wahrscheinlichkeitsverteilung von f.**

Bevor wir einige Beispiele für Wahrscheinlichkeitsverteilungen besprechen, geben wir geben wir zunächst die folgende extrem nützliche Konsequenz von Theorem 2.5.4 an.

Korollar 2.5.6 (Erwartungswerte via Wahrscheinlichkeitsverteilung)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Wahrscheinlichkeitsraum, $f : \Omega \to \mathbb{R}$ eine Zufallsvariable, und sei $\nu : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ die Wahrscheinlichkeitsverteilung von f, das heißt $\nu = f_*\mu$.*

Die Zufallsvariable f ist genau dann integrierbar, wenn $\int_{\mathbb{R}} |x| \, \mathrm{d}\nu(x) < \infty$ ist, und in diesem Fall ist der Erwartungswert von f gegeben durch

$$\mathbb{E}f = \int_{\mathbb{R}} x \, \mathrm{d}\nu(x).$$

Beweis. Man wende den Transformationssatz für Bildmaße (Theorem 2.5.4) auf den Messraum $(\widehat{\Omega}, \widehat{\mathcal{A}}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ und die Funktion $g = \mathrm{id}_{\mathbb{R}}$ an. □

Das so harmlos wirkende Korollar 2.5.6 hat es in Wirklichkeit in sich, denn es zeigt folgendes: Wenn man sich für die Erwartungswerte einer Zufallsvariable f interessiert, [27] muss man den zugrundeliegenden Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, \mu)$ und die Abbildungsvorschrift für f gar nicht kennen! Es genügt, lediglich die Verteilung von f zu kennen. Dies ist einer der Gründe, weshalb man in der Stochastik viel öfter über die Verteilung von Zufallsvariablen spricht als über den zugrundeliegenden Wahrscheinlichkeitsraum.

Lassen Sie uns zur Veranschaulichung ein einfaches Beispiel für eine Wahrscheinlichkeitsverteilung diskutieren.

Beispiel 2.5.7 (Binomialverteilung mit Wahrscheinlichkeit **1/2**)**.** Sei $n \in \mathbb{N}$ und sei $\Omega = \{Z, K\}^n$. Wir betrachten das Wahrscheinlichkeitsmaß $\mu : 2^\Omega \to [0, \infty]$, das durch $\mu(A) = \frac{\#A}{2^n}$ gegeben ist.[28] Sei $f : \Omega \to \mathbb{R}$ diejenigen Funktion, die jedem $\omega \in \Omega$ die Zahl zuordnet, die angibt, wie oft K in ω vorkommt. Ganz präzise geschrieben heißt das

$$f(\omega) = \#\big\{j \in \{1, \dots, n\} \mid \omega_j = K\big\}.$$

für alle $\omega \in \Omega$. Die Funktion f ist messbar, da wir auf Ω die σ-Algebra 2^Ω betrachten. Es gilt:

[27] Und eine analoge Aussage erhält man zum Beispiel auch für die sogenannte **Varianz** einer Zufallsvariablen aus dem Transformationssatz für Bildmaße.

[28] Anschaulich betrachten wir also die Situation, dass eine faire Münze n-mal geworfen wird. Die möglichen Ergebnisse des n-maligen Münzwurfs sind die n-Tupel $\omega \in \Omega$.

(a) Die Wahrscheinlichkeitsverteilung $\nu := f_*\mu : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ ist gegeben durch

$$\nu(A) = \sum_{k\in\{0,\dots,n\}\cap A} \frac{1}{2^n}\binom{n}{k},$$

für alle $A \in \mathcal{B}(\mathbb{R})$, wobei $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ wie üblich den Binomialkoeffizienten bezeichnet.

(b) Der Erwartungswert von f ist $\mathbb{E}f = \frac{n}{2}$.

Beweis. (a) Sei $A \in \mathcal{B}(\mathbb{R})$. Dann gilt

$$\begin{aligned}\nu(A) = (f_*\mu)(A) = \mu\big(f^{-1}(A)\big) = \frac{\#f^{-1}(A)}{2^n} = \frac{1}{2^n}\#f^{-1}\Big(\bigcup_{k\in\{0,\dots,n\}\cap A}\{k\}\Big)\\ = \frac{1}{2^n}\#\Big(\bigcup_{k\in\{0,\dots,n\}\cap A} f^{-1}\big(\{k\}\big)\Big) = \frac{1}{2^n}\sum_{k\in\{0,\dots,n\}\cap A}\#f^{-1}\big(\{k\}\big) = \frac{1}{2^n}\sum_{k\in\{0,\dots,n\}\cap A}\binom{n}{k}.\end{aligned}$$

Für die vierte Gleichheit haben wir verwendet, dass f nur Werte in $\{0,\dots,n\}$ annimmt. Für die letzte Gleichheit haben wir verwendet, dass es genau so viele Elemente $\omega \in \{Z,K\}^n$ gibt, die genau k-Mal das Symbol K enthalten, wie es k-elementige Teilmengen von $\{1,\dots,n\}$ gibt (und das dies genau $\binom{n}{k}$ Stück sind).

Vorlesung 17 (Mi, 10.12.) beginnt mit dem Beweis von Beispiel 2.5.7(b)

(b) Es gilt

$$\begin{aligned}\mathbb{E}f = \int_{\mathbb{R}} x\,\mathrm{d}\nu(x) = \int_{\{0,\dots,n\}} x\,\mathrm{d}(\nu|_{\{1,\dots,n\}})(x) + \int_{\mathbb{R}\setminus\{0,\dots,n\}} x\,\mathrm{d}(\nu|_{\mathbb{R}\setminus\{1,\dots,n\}})(x)\\ = \int_{\{0,\dots,n\}} x\,\mathrm{d}(\nu|_{\{1,\dots,n\}})(x) = \sum_{k\in\{0,\dots,n\}} k\frac{1}{2^n}\binom{n}{k},\end{aligned}$$

wobei die erste Gleichheit aus Beispiel 2.5.6 folgt, die zweite Gleichheit aus Aufgabe 8.3(c) auf Hausaufgabenblatt 8, die dritte Gleichheit aus $\nu|_{\mathbb{R}\setminus\{1,\dots,n\}} = 0$, und die vierte Gleichheit aus Aufgabe 7.3(a) auf Präsenzblatt 7. Wegen

$$k\binom{n}{k} = \begin{cases} 0, & \text{falls} \quad k = 0,\\ \frac{n!}{(k-1)!(n-k)!} & \text{falls} \quad k \in \{1,\dots,n\}\end{cases}$$

erhalten wir also

$$\begin{aligned}\mathbb{E}f = \frac{1}{2^n}\sum_{k=1}^{n}\frac{n!}{(k-1)!(n-k)!} = \frac{n}{2^n}\sum_{k=1}^{n}\frac{(n-1)!}{(k-1)!\big((n-1)-(k-1)\big)!}\\ = \frac{n}{2^n}\sum_{k=0}^{n-1}\frac{(n-1)!}{k!\big((n-1)-k\big)!} = \frac{n}{2^n}2^{n-1} = \frac{n}{2},\end{aligned}$$

wie behauptet. □

2.6 Absolute Stetigkeit von Maßen

Einstiegsfragen. (a) Welche Vorstellung haben Sie, wenn Sie den Begriff "Massedichte" hören? Und wenn Sie den Begriff "Ladungsdichte" hören?

(b) Würden Sie sagen, dass das Lebesgue-Maß λ_d und das Dirac-Delta-Maß δ_0 des Punktes $0 \in \mathbb{R}^d$ sich in irgendeinem Sinne fundamental unterscheiden?

Definition 2.6.1 (Absolutstetigkeit von Maßen)**.** Sei $(\Omega, \mathcal{A})$ ein Messraum und seien $\mu, \nu : \mathcal{A} \to [0,\infty]$ zwei Maße. Wir nennen ν **absolutstetig bezüglich μ**, falls für alle $A \in \mathcal{A}$ die Implikation

$$\mu(A) = 0 \qquad \Rightarrow \qquad \nu(A) = 0$$

gilt. In diesem Fall schreiben wir $\nu \ll \mu$.

Hier ist ein einfaches Beispiel zweier Maße von deinen keines bezüglich des anderen absolutstetig ist.

Beispiel 2.6.2 (Dirac-Maße und das Lebesgue-Maß)**.** Sei $x \in \mathbb{R}$ und sei $\delta_x : 2^{\mathbb{R}} \to [0,\infty]$ das Dirac-δ-Maß im Punkt x. Dann gilt:

(a) $\delta_x|_{\mathcal{B}(\mathbb{R})}$ ist nicht absolutstetig bezüglich λ_1.[29]

(b) λ_1 ist nicht absolutstetig bezüglich $\delta_x|_{\mathcal{B}(\mathbb{R})}$.

Beweis. (a) Es gilt $(\delta_x|_{\mathcal{B}(\mathbb{R})})(\{x\}) = \delta_x(\{x\}) = 1$, aber $\lambda_1(\{x\}) = 0$.

(b) Es gilt $\lambda_1([x+1, x+2]) = 1$, aber $(\delta_x|_{\mathcal{B}(\mathbb{R})})([x+1, x+2]) = \delta_x([x+1, x+2]) = 0$. □

Man kann aus einem gegebenen Maß μ leicht neue Maße konstruieren, die bezüglich μ absolut stetig sind. Wir zeigen das in Proposition 2.6.4. Der einfacheren Schreibweise halber führen wir vorher noch die folgende Notation ein.

Notation 2.6.3 (Integration über Teilmengen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und sei $f : \Omega \to [0,\infty]$ messbar oder $f : \Omega \to \mathbb{C}$ integrierbar. Für jedes $C \in \mathcal{A}$ setzt man

$$\int_C f \, \mathrm{d}\mu := \int_\Omega \mathbb{1}_C f \, \mathrm{d}\mu.$$

Aus Aufgabe 8.3(b) auf Hausaufgabenblatt 8 folgt, dass $\int_C f \, \mathrm{d}\mu = \int_C f|_C \, \mathrm{d}(\mu|_C)$ ist.[30]

Die Schreibweise mit Integrationsvariablen, die wir in Notation 2.4.4 eingeführt haben, kann man natürlich auch für Integrale über $C \in \mathcal{A}$ verwenden, das heißt, man schreibt

$$\int_C f \, \mathrm{d}\mu =: \int_C f(\omega) \, \mathrm{d}\mu(\omega).$$

[29] Wir sprechen hier von der Einschränkung $\delta_x|_{\mathcal{B}(\mathbb{R})}$ anstelle von δ_x selbst, weil man laut Definition 2.6.1 überhaupt nur dann von Absolutstetigkeit eines Maßes bezüglich eines anderen Maßes sprechen kann, wenn beide dieselbe σ-Algebra als Definitionsbereich haben.

[30] Genau genommen haben Sie das in dieser Aufgabe nur für $f : \Omega \to [0,\infty]$ gezeigt. Hieraus erhält man aber leicht: Wenn $f : \Omega \to \mathbb{C}$ integrierbar ist, dann ist $f|_C : C \to \mathbb{C}$ bezüglich $\mu|_C$ ebenfalls integriert und es gilt dieselbe Formel.

Proposition 2.6.4 (Konstruktion von Maßen aus Dichten)**.** *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Maßraum und sei* $f : \Omega \to [0, \infty)$ *messbar.*[31] *Dann ist die Abbildung*

$$\nu : \mathcal{A} \to [0, \infty], \qquad A \mapsto \nu(A) := \int_A f \, \mathrm{d}\mu$$

ein Maß und ν *ist absolutstetig bezüglich* μ*.*

Den Beweisen führen wir in den Übungen. Die Konstruktion in Proposition 2.6.4 ist so nützlich, dass wir folgenden Begriff dafür einführen.

Definition 2.6.5 (Dichte von Maßen)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $\nu : \mathcal{A} \to [0, \infty]$ ein weiteres Maß und sei $f : \Omega \to [0, \infty)$ messbar. Wir sagen, dass f eine **Dichte von ν bezüglich μ** ist, falls $\nu(A) = \int_A f \, \mathrm{d}\mu$ für alle $A \in \mathcal{A}$ gilt.

Wir werden in Proposition 2.6.7 zeigen, dass Dichten "im Wesentlichen" eindeutig sind. Um zu präzisieren, was hierbei "im Wesentlichen" bedeutet, verwenden wir den folgenden Begriff.

Definition 2.6.6 (Fast überall-Eigenschaften)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und für jedes $\omega \in \Omega$ sei eine Aussage $A(\omega)$ gegeben. Wir sagen, dass A *μ-fast überall* wahr ist, falls es eine Menge $N \in \mathcal{A}$ mit $\mu(N) = 0$ gibt derart, dass $A(\omega)$ für alle $\omega \in \Omega \setminus N$ wahr ist.

Wenn das Maß μ aus dem Kontext klar ist, sagt man oft auch einfach **fast überall** anstelle von μ-fast überall.

Proposition 2.6.7 (Eindeutigkeit von Dichten)**.** *Sei* $(\Omega, \mathcal{A}, \mu)$ *ein Maßraum, sei* $\nu : \mathcal{A} \to [0, \infty]$ *ein* σ*-endliches Maß, und seien* $f_1, f_2 : \Omega \to [0, \infty)$ *messbar und Dichten von* ν *bezüglich* μ*. Dann gilt* μ*-fast überall* $f_1 = f_2$*.*[32,33]

Hinweis: In der Vorlesung beweisen wir Proposition 2.6.7 nur im Spezialfall $\nu(\Omega) < \infty$. Der Beweis beruht in diesem Spezialfall auf derselben Idee, ist aber etwas weniger technisch.

Beweis. Wegen der σ-Endlichkeit von ν gibt es eine Folge $E_1, E_2, \dots \in \mathcal{A}$ mit $\bigcup_{j=1}^{\infty} E_j = \Omega$ und mit $\nu(E_j) < \infty$ für alle $j \in \mathbb{N}$. Wir setzen

$$L := \{\omega \in \Omega \mid \ f_1(\omega) < f_2(\omega)\} = (f_1 - f_2)^{-1}\big((-\infty, 0)\big),$$
$$M := \{\omega \in \Omega \mid \ f_1(\omega) > f_2(\omega)\} = (f_1 - f_2)^{-1}\big((0, \infty)\big)$$

und $N := L \cup M = \{\omega \in \Omega \mid \ f_1(\omega) \neq f_2(\omega)\}$. Die Mengen L, M und N sind messbar und es genügt $\mu(N) = 0$ zu zeigen. Dafür wiederum müssen wir nur $\mu(L) = 0$ und $\mu(M) = 0$ zeigen.

Wir können L schreiben als $L = \bigcup_{n=1}^{\infty} L_n$ mit

$$L_n := \{\omega \in \Omega \mid \ f_1(\omega) + \tfrac{1}{n} < f_2(\omega)\} = (f_1 - f_2)^{-1}\big((-\infty, -\tfrac{1}{n})\big) \in \mathcal{A}$$

[31] Man könnte hier auch messbare Funktionen f zulassen, die Werte in $[0, \infty]$ annehmen, aber dann werden manche Stellen in diesem Abschnitt etwas technischer. Darauf verzichten wir lieber.

[32] Das heißt also laut Definition 2.6.6, dass es ein $N \in \mathcal{A}$ mit $\mu(N) = 0$ gibt derart, dass $f_1(\omega) = f_2(\omega)$ für alle $\omega \in \Omega \setminus N$ gilt.

[33] Und somit auch ν-fast überall, da ν laut Proposition 2.6.4 absolut stetig bezüglich μ ist.

für alle $n \in \mathbb{N}$. Es gilt für all $n, j \in \mathbb{N}$

$$\nu(L_n \cap E_j) = \int_{L_n \cap E_j} f_2 \, \mathrm{d}\mu \geq \int_{L_n \cap E_j} f_1 + \frac{1}{n} \, \mathrm{d}\mu = \nu(L_n \cap E_j) + \frac{1}{n}\mu(L_n \cap E_j)$$

Wegen $\nu(L_n \cap E_j) \leq \nu(E_j) < \infty$ folgt daraus $\mu(L_n \cap E_j) = 0$ für alle $n, j \in \mathbb{N}$ und somit $\mu(L_n) \leq \sum_{j=1}^{\infty} \mu(L_n \cap E_j) = 0$ für alle $n \in \mathbb{N}$. Insgesamt ist also $\mu(L) \leq \sum_{n=1}^{\infty} \mu(L_n) = 0$, das heißt $\mu(L) = 0$.

Indem man die Rollen von f_1 und f_2 vertauscht folgt auch $\mu(M) = 0$ und somit, wie behauptet, $\mu(N) = 0$. □

Vorlesung 18 (Fr, 12.12.) beginnt mit Proposition 2.6.8

Proposition 2.6.8 (Integral bezüglicher von Maßen mit Dichte)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $\nu : \mathcal{A} \to [0, \infty]$ ein weiteres Maß und sei $f : \Omega \to [0, \infty)$ messbar und eine Dichte von ν bezüglich μ.*

(a) *Sei $g : \Omega \to [0, \infty]$ messbar. Dann gilt*

$$\int_\Omega g \, \mathrm{d}\nu = \int_\Omega g f \, \mathrm{d}\mu.$$

(b) *Sei $g : \Omega \to \mathbb{C}$ messbar. Es ist g genau dann bezüglich ν integrierbar, wenn gf bezüglich μ integrierbar ist, und in diesem Fall gilt*

$$\int_\Omega g \, \mathrm{d}\nu = \int_\Omega g f \, \mathrm{d}\mu.$$

Beweis. (a) Sei zunächst $g : \Omega \to [0, \infty)$ einfach, das heißt, $g = \sum_{k=1}^{m} \alpha_k \mathbb{1}_{A_k}$ für geeignete Skalare $\alpha_1, \ldots, \alpha_m \in [0, \infty)$ und Mengen $A_1, \ldots, A_m \in \mathcal{A}$. Dann gilt

$$\int_\Omega g f \, \mathrm{d}\mu = \sum_{k=1}^{m} \alpha_k \int_\Omega \mathbb{1}_{A_k} f \, \mathrm{d}\mu = \sum_{k=1}^{m} \alpha_k \int_{A_k} f \, \mathrm{d}\mu = \sum_{k=1}^{m} \alpha_k \nu(A_k) = \int_\Omega g \, \mathrm{d}\nu.$$

Für allgemeine messbare Funktionen $g : \Omega \to [0, \infty]$ erhält man nun die Behauptung, indem man g von unten durch messbare Funktionen approximiert (Proposition 2.1.10) und den Satz von der monotonen Konvergenz (Theorem 2.2.3(b)) verwendet.

(b) Aus (a) folgt

$$\int_\Omega |g| \, \mathrm{d}\nu = \int_\Omega |g| f \, \mathrm{d}\mu = \int_\Omega |gf| \, \mathrm{d}\mu.$$

Also ist g genau dann bezüglich ν integrierbar, wenn gf bezüglich μ integrierbar ist. In diesem Fall erhält man aus (a) die Gleichheit $\int_\Omega g \, \mathrm{d}\nu = \int_\Omega g f \, \mathrm{d}\mu$, indem man g in seine positiven Anteile zerlegt. □

Die vorangehende Proposition motiviert die folgende Notation für ein Maß ν, welche eines Dichte f bezüglich eines Maßes μ besitzt.

Notation 2.6.9 (Maße mit Dichte)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $\nu : \mathcal{A} \to [0, \infty]$ ein weiteres Maß und sei $f : \Omega \to [0, \infty)$ messbar. Wenn man sagen möchte, dass f eine Dichte von ν bezüglich μ ist, schreibt man hierfür oft $\mathrm{d}\nu = f\,\mathrm{d}\mu$.

Wenn auf einem Messraum $(\Omega, \mathcal{A})$ zwei Maße $\mu, \nu : \mathcal{A} \to [0, \infty]$ gegeben sind und ν eine Dichte bezüglich μ besitzt, dann ist ν laut Proposition 2.6.4 absolutstetig bezüglich μ. Bemerkenswerter Weise gilt (im σ-endlichen Fall) auch die umgekehrte Implikation, das heißt, man hat insgesamt das folgende Resultat.

Theorem 2.6.10 (Absolutstetigkeit und Dichten – der Satz von Radon-Nikodým)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum und $\nu : \mathcal{A} \to [0, \infty]$ ein weiteres σ-endliches Maß. Dann sind äquivalent:*

(i) *Das Maß ν ist absolutstetig bezüglich μ, das heißt $\nu \ll \mu$.*

(ii) *Es gibt eine messbare Funktion $f : \Omega \to [0, \infty)$ mit $\mathrm{d}\nu = f\,\mathrm{d}\mu$.*

Das Theorem werden wir aus Zeitgründen nicht beweisen. Bei Interesse können Sie einen Beweis zum Beispiel in [1, Satz VII.2.3 auf Seite 303] finden. Ein sehr effizienter Beweis ist zu dem mit den Methoden der Funktionalanalysis möglich. Falls sich in den *Grundlagen der Funktionalanalysis* die Gelegenheit ergibt, werden wir den Satz dort beweisen.[34]

[34] Das war zwar kein Versprechen, aber auf jeden Fall ein Versuch Werbung zu machen!

Kapitel 3

Funktionenräume

3.1 Banachräume und lineare Operatoren

Einstiegsfragen. (a) Als welche Reihe war nochmal die Exponentialfunktion auf $\mathbb{R}$ definiert? Warum konvergiert diese Reihe und was hat das mit der Vollständigkeit von $\mathbb{R}$ zu tun?

(b) Wiederholung aus der Linearen Algebra 1: Was haben Matrizen mit linearen Abbildungen zu tun?

(c) Wiederholung aus der Analysis 2: Was war noch einmal die induzierte Norm einer Matrix?

Wir wiederholen kurz den Begriff der **Norm**, den Sie aus der Analysis 2 kennen, und definieren außerdem den etwas allgemeineren Begriff **Halbnorm**.

Definition 3.1.1 (Normen und Halbnormen)**.** Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $\|\cdot\| : V \to [0,\infty)$ heißt **Halbnorm**, falls sie die folgenden Axioma erfüllt:

(I) **Absolute Homogenität:** Für alle $v \in V$ und alle $\alpha \in \mathbb{K}$ gilt $\|\alpha v\| = |\alpha| \, \|v\|$.

(II) **Dreiecksungleichung:** Für alle $v, w \in V$ gilt $\|v + w\| \leq \|v\| + \|w\|$.

Man bezeichnet eine Halbnorm $\|\cdot\| : V \to [0,\infty)$ als **Norm**, wenn Sie zusätzlich die folgenden Eigenschaft erfüllt:

(III) **Positive Definitheit:** Ist $v \in V$ und $\|v\| = 0$, dann folgt $v = 0$.

Wenn $\|\cdot\|$ eine Norm auf V ist, dann bezeichnet man das Paar $(V, \|\cdot\|)$ als **normierten Raum**. Häufig unterdrückt man hierbei das Symbol $\|\cdot\|$ in der Schreibweise, das heißt, man schreibt einfach nur, dass "V ein normierter Raum" ist, anstatt zu schreiben, dass "$(V, \|\cdot\|)$ ein normierter Raum" ist.

Halbnormen werden wir in Abschnitt 3.2 verwenden. Im aktuellen Abschnitt sehen wir uns einige grundlegenden Eigenschaften von Normen und normierten Räumen an.

Zunächst erinnern wir daran, dass jeder normierte Raum V in natürlicher Weise auch ein metrischer Raum ist, denn wenn man für alle $v, w \in V$

$$\mathrm{d}(v, w) := \|v - w\|$$

definiert, dann ist d eine Metrik. Damit sind für normierte Räume auch automatische Begriffe wie Stetigkeit, Konvergenz von Folgen, und Cauchy-Folgen definiert, denn diese Begriffe wurden in der Analysis 2 allgemein für metrische Räume eingeführt. Folgende Eigenschaft ist elementar, aber besonders nützlich, weshalb wir sie hier kurz wiederholen. Für eine Folge (v_n) in V und ein $v \in V$ gilt $v_n \to v$ genau dann, wenn $\mathrm{d}(v_n, v) \to 0$ gilt, und dies ist wiederum äquivalent zu

$$\|v_n - v\| \to 0.$$

Diese Eigenschaft ist äußerst nützlich um Konvergenz von Folgen in normierten Räumen zu beweisen, und wir werden sie sehr oft verwenden.

Proposition 3.1.2 (Stetigkeit der Vektorraum-Operationen)**.** *Sei E ein normierter Raum, seien (x_n) und (y_n) konvergente Folgen in E mit Grenzwerten x, y, und seien (α_n) und (β_n) konvergente Folgen im Skalarkörper mit Grenzwerten α, β. Dann gilt*

$$\alpha_n x_n + \beta_n y_n \to \alpha x + \beta y \qquad \text{und} \qquad \|x_n\| \to \|x\|$$

für $n \to \infty$.

Beweis. Die Behauptung $\|x_n\| \to \|x\|$ zeigen Sie in einer Übungsaufgabe. Für die anderen beiden Aussagen genügt es $\alpha_n x_n \to \alpha x$ und $x_n + y_n \to x + y$ zu zeigen.

Skalare Multiplikation: Es ist

$$\begin{aligned}\|\alpha_n x_n - \alpha x\| &\le \|\alpha_n x_n - \alpha_n x\| + \|\alpha_n x - \alpha x\| \\ &= \underbrace{|\alpha_n|}_{\text{beschr.}} \underbrace{\|x_n - x\|}_{\to 0} + \underbrace{|\alpha_n - \alpha|}_{\to 0} \|x\| \to 0\end{aligned}$$

für $n \to \infty$.

Addition: Es gilt

$$\big\|(x_n + y_n) - (x + y)\big\| \le \|x_n - x\| + \big\|y_n - y\big\| \to 0$$

für $n \to \infty$, womit die Behauptung bewiesen ist. □

Lassen Sie uns kurz ein paar Beispiele für normierte Räume besprechen; manche davon haben wir auch in der Analysis 2 schon besprochen. Weitere wichtige Beispiele für normierte Räume werden wir im nachfolgenden Abschnitt 3.2 mit Hilfe der Maß- und Integrationstheorie konstruieren.

Vorlesung 19 (Mi, 17.12.) beginnt mit Beispiel 3.1.9

Beispiele 3.1.3 (Ein paar normierte Räume)**.** Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $d \in \mathbb{N}$.

(a) Sei $p \in [1, \infty]$. Es ist $\mathbb{K}^d$ zusammen mit der p-Norm $\|\cdot\|_p$, die durch

$$\|x\|_p := \begin{cases} \left(\sum_{k=1}^d |x_k|^p\right)^{1/p} & \text{falls } p \in [1,\infty), \\ \max\left\{|x_k| \mid k \in \{1,\dots,d\}\right\} \end{cases}$$

für alle $x \in \mathbb{K}^n$ gegeben ist, ein normierter Raum.

Wie Sie aus der Analysis 2 wissen, führt jede Norm in $\mathbb{K}^d$ zurselben Charakterisierung von Konvergenz: Eine Folge $(x^{(n)})_{n\in\mathbb{N}}$ in $\mathbb{K}^d$ konvergiert genau dann gegen ein $x \in \mathbb{K}^d$, wenn für jedes $k \in \{1,\dots,d\}$ die Konvergenz $x_k^{(n)} \to x_k$ in $\mathbb{K}$ gilt.

(b) Sei (M, d) ein kompakter metrischer Raum. Dann ist der Vektorraum $C(M;\mathbb{K})$ aller stetigen Funktionen von M nach $\mathbb{K}$ zusammen der **Supremumsnorm** (auch ∞-**Norm** genannt), die durch

$$\|f\|_{\sup} := \|f\|_\infty := \sup\left\{|f(\omega)| \mid \omega \in M\right\}$$

für alle $f \in C(M;\mathbb{K})$ gegeben ist, ein normierter Raum. Die Konvergenz einer Funktionenfolge (f_n) gegen eine Funktion f in diesem Raum bedeutet gerade, dass (f_n) gleichmäßig gegen f konvergiert.

(c) Sei c_{00} die Menge all derjenigen Folgen $x = (x_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$, die nur endlich viele von 0 verschiedene Glieder haben. Das ist Vektorraum bezüglich der komponentenweisen Addition und Skalarmultiplikation, und man kann leicht nachrechnen, dass durch

$$\|x\|_\infty := \max_{n\in\mathbb{N}} |x_n|$$

eine Norm auf c_{00} gegeben ist.

In der Analysis 2 haben wir häufig lineare Abbildungen zwischen $\mathbb{K}^c$ und $\mathbb{K}^d$ – oder, äquivalent dazu, Matrizen in $\mathbb{K}^{d\times c}$ – verwendet.[1] Praktischer Weise sind lineare Abbildungen zwischen $\mathbb{K}^d$ und $\mathbb{K}^c$ immer stetig. In unendlicher Dimension ist das allerdings anders, wie das folgende Beispiel zeigt:

Beispiel 3.1.4 (Eine unstetige lineare Abbildung)**.** Betrachten wir den Raum c_{00} mit der Norm $\|\cdot\|_\infty$ aus Beispiel 3.1.3(c). Es sei $T : c_{00} \to c_{00}$ gegeben durch

$$Tx := (n^2 x_n)_{n\in\mathbb{N}}$$

[1] Zum Beispiel haben wir die Ableitung $\mathrm{D}f(x)$ einer differenzierbaren Funktion $f : \mathbb{K}^c \to \mathbb{K}^d$ in einem Punkt $x \in \mathbb{K}^c$ als eine Matrix in $\mathbb{K}^{d\times c}$.

für alle $x = (x_n)_{n\in\mathbb{N}}$. Man kann sich T auch als Matrixmultiplikation mit einer unendlich großen Matrix vorstellen: Für alle $x \in c_{00}$ ist

$$Tx = \begin{pmatrix} 1^2 x_1 \\ 2^2 x_2 \\ 3^2 x_3 \\ 4^2 x_4 \\ \vdots \end{pmatrix} = \begin{pmatrix} 1^2 & & & & \\ & 2^2 & & & \\ & & 3^2 & & \\ & & & 4^2 & \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ \vdots \end{pmatrix}$$

Die Abbildung T ist linear und nicht stetig.

Beweis. Linearität: Für alle $x, y \in c_{00}$ und alle $\alpha, \beta \in \mathbb{C}$ gilt

$$\begin{aligned} \big(T(\alpha x + \beta x)\big)_n &= n^2(\alpha x + \beta y)_n = n^2(\alpha x_n + \beta y_n) = n^2\alpha x_n + n^2\beta y_n \\ &= \alpha n^2 x_n + \beta n^2 y_n = \alpha(Tx)_n + \beta(Ty)_n = \big(\alpha Tx + \beta Tx\big)_n \end{aligned}$$

für alle $n \in \mathbb{N}$ und somit $T(\alpha x + \beta y) = \alpha Tx + \beta Ty$.

Unstetigkeit: Für jedes $n \in \mathbb{N}$ bezeichne $e^{(n)} \in c_{00}$ den n-ten kanonischen Einheitsvektor, das heißt, diejenige Folge in c_{00}, die an der Position n den Eintrag 1 und sonst den Eintrag 0 besitzt. Dann ist $\left\| \frac{1}{n} e^{(n)} \right\|_\infty = \frac{1}{n} \to 0$ und somit $\frac{1}{n} e^{(n)} \to 0$ für $n \to \infty$. Zugleich gilt aber $Te^{(n)} = n^2 e^{(n)}$ für alle $n \in \mathbb{N}$ und somit

$$\left\| T \frac{1}{n} e^{(n)} \right\|_\infty = \left\| n e^{(n)} \right\|_\infty = n \to \infty$$

für und somit $\frac{1}{n} e^{(n)} \not\to 0 = T0$ für $n \to \infty$. □

Da es in unendlicher Dimension unstetige lineare Abbildungen gibt, liegt es nahe sich Gedanken darüber zu machen, wann eine lineare Abbildung zwischen zwei normieren Räumen stetig ist. Das tun wir in der folgenden Proposition.

Proposition 3.1.5 (Stetigkeit linearer Abbildungen)**.** *Seien V, W normierte Räume über demselben Körper und sei $T : V \to W$ linear. Dann sind äquivalent:*

(i) *T ist stetig.*

(ii) *T ist stetig im Nullpunkt $0 \in V$.*

(iii) *Es gilt $\sup\{ \|Tx\|_W \mid x \in V,\ \|x\|_V \le 1 \} < \infty$.*

(iv) *Es gibt eine Zahl $C \in [0, \infty)$ mit $\|Tx\|_W \le C \|x\|_V$ für alle $x \in V$.*

Beweis. "(i) ⇒ (ii)" Diese Implikation ist offensichtlich.

"(ii) ⇒ (i)" Wir verwenden mehrmals Proposition 3.1.2. Sei $x \in V$ und sei (x_n) eine Folge in V, die gegen x konvergiert. Dann gilt $x_n - x \to 0$ und somit wegen (ii) $Tx_n - Tx = T(x_n - x) \to 0$. Folglich gilt $Tx_n \to Tx$, also ist T stetig in x.

"(ii) ⇒ (iii)" Wir zeigen die Kontraposition. Sei das Supremum in (iii) gleich ∞. Dann gibt es für jedes $n \in \mathbb{N}$ einen Vektor $x_n \in V$ mit $\|Tx_n\|_W > n^2$ und $\|x_n\|_W \leq 1$. Damit gilt $\|x_n/n\| \leq \frac{1}{n} \to 0$ und somit $x_n/n \to 0$ für jedes $n \in \mathbb{N}$. Zugleich ist aber $\|Tx_n/n\|_W = \|Tx_n\|_W/n > n \to \infty$ für $n \to \infty$. Somit gilt $Tx_n/n \not\to 0 = T0$ für $n \to \infty$, also ist T im Punkt 0 unstetig.

"(iii) ⇒ (iv)" Bezeichne $C \in [0,\infty)$ das Supremum in (iii). Sei $x \in V$ mit $\|x\|_V \leq 1$. Falls $x = 0$ ist, gilt $\|Tx\|_W = \|0\|_W = 0 = \|0\|_V$. Sei nun also $x \neq 0$. Dann können wir $v := x/\|x\|_V$ setzen und erhalten $\|v\|_V = 1$ und somit

$$\frac{\|Tx\|_W}{\|x\|_V} = \|Tv\|_W \leq C.$$

Also folgt $\|Tx\|_W \leq C\|x\|_V$.

"(iv) ⇒ (ii)" Sei (x_n) eine Folge in V, die gegen 0 konvergiert. Dann gilt

$$0 \leq \|Tx_n\|_W \leq C\|x_n\|_V \to 0$$

und somit $Tx_n \to 0$. Also ist T stetig in 0. □

Proposition 3.1.6 (Summe und skalare Vielfache stetiger linearer Operatoren)**.** *Seien V, W Vektorräume über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, seien $S, T : V \to W$ linear und stetig und seien $\alpha, \beta \in \mathbb{K}$. Dann ist auch $\alpha S + \beta T : V \to W$ stetig.*

Beweis. Das folgt leicht aus Proposition 3.1.2. □

Übrigens stimmt Proposition 3.1.6 auch dann, wenn S und T nicht stetig sind (mit dem selben Beweis). An dieser Stelle benötigen wir die Aussage insbesondere für lineare Abbildungen, damit die nachfolgende Definition 3.1.7 Sinn ergibt.

Lineare Abbildungen bezeichnet man, insbesondere wenn sie auf unendlich-dimensionalen Vektorräumen definiert sind, oft als **lineare Operatoren**. Das erklärt zum Beispiel die Verwendung des Begriffs **Operatornorm** in der folgenden Definition.

Definition 3.1.7 (Operatorraum und Operatornorm)**.** Seien V, W normierte Räume über demselben Körper.

(a) Mit $\mathcal{L}(V;W)$ bezeichne wir den Vektorraum aller stetigen linearen Abbilungen von V nach W (ausgestattet mit der punktweisen Addition und der skalaren Multiplikation).[2]

(b) Für jedes $T \in \mathcal{L}(V;W)$ nennt man

$$\|T\|_{\mathcal{L}(V;W)} := \|T\|_{W \leftarrow V} := \|T\| := \sup\left\{\|Tx\|_W \mid x \in V,\ \|x\|_V \leq 1\right\} \in [0,\infty)$$

die **Operatornorm von T**.

[2] Manchmal nennt man $\mathcal{L}(V;W)$ auch kurz den **Operatorraum** zwischen V und W, aber das ist etwas vage.

Bevor wir einige Eigenschaften der Operatornorm anschreiben, führen wir noch den folgenden Begriff ein. Er nimmt eine ganz zentrale Stelle in der unendlich-dimensionalen Analysis ein.

Definition 3.1.8 (Banachräume)**.** Ein normierter Raum heißt **Banachraum**, wenn er bezüglich der von der Norm induzierten Metrik vollständig ist (das heißt, wenn jede Cauchy-Folge konvergiert).

Proposition 3.1.9 (Die Operatornorm ist eine Norm)**.** *Seien V, W normierte Räume über demselben Körper.*

(a) *Die Operatornorm $\|\cdot\|_{W\leftarrow V} : \mathcal{L}(V,W) \to [0,\infty)$ ist tatsächlich eine Norm.*

(b) *Wenn der Bildraum W ein Banachraum ist, dann ist auch $\mathcal{L}(V;W)$ mit der Operatornorm ein Banachraum.*

Den Beweis von (a) werden wir ebenfalls in den Übungen führen. Auf einen Beweis von (b) verzichten wir an dieser Stelle.[3]

Sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ und sei $A \in \mathbb{K}^{c\times d}$ eine Matrix, die wir als einen linearen Operator $A : \mathbb{K}^d \to \mathbb{K}^c$ auffassen. Seien $p, q \in [1,\infty]$. Wenn der Definitionsbereich $\mathbb{K}^d$ der Abbildung A mit der ℓ^p-Norm ausgestattet ist und der Wertebereich $\mathbb{K}^c$ mit der ℓ^q-Norm, dann schreiben wir für die Operatornorm von A oft kurz $\|A\|_{q\leftarrow p}$. In einigen wichtigen Spezialfällen kann man diese Operatornorm durch eine explizite Formel ausdrücken. Für die folgenden Fälle hatten wir das in der Analysis 2 besprochen.

Hinweis: Das Beispiel 3.1.10 kennen Sie bereits aus der Analysis 2. Deshalb – und um zügig zu Abschnitt 3.2 zu gelangen – haben wir es in der Vorlesung nur mündlich besprochen.

Beispiel 3.1.10 (Induzierte Matrixnormen bzgl. der **1**-Norm und der ∞-Norm)**.** Sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$ und sei $A \in \mathbb{K}^{c\times d}$. Dann gilt

$$\|A\|_{1\leftarrow 1} = \max\left\{\left\|A_{\bullet,1}\right\|_1, \dots, \left\|A_{\bullet,d}\right\|_1\right\}$$
$$\text{und} \quad \|A\|_{\infty\leftarrow\infty} = \max\left\{\left\|A_{1,\bullet}\right\|_1, \dots, \left\|A_{c,\bullet}\right\|_1\right\},$$

wobei $A_{\bullet,1}, \dots, A_{\bullet,d} \in \mathbb{R}^c$ die Spalten von A und $A_{1,\bullet}, \dots, A_{c,\bullet} \in \mathbb{R}^{1\times d}$ die Zeilen von A bezeichnen.

3.2 L^p-Räume

Vorlesung 20 (Fr, 19.12.) beginnt mit Abschnitt 3.1.9

Einstiegsfragen. (a) Haben Sie eine Vorstellung von dem Begriff "quadratischer Fehler"?

(b) Falls Sie sich für Physik interessieren: Was ist ein reiner Zustand eines Quantensystems?

[3] Er ist nicht besonders kompliziert, aber nimmt etwas Zeit in Anspruch. Die Aussage (b) ist für das aktuelle Kapitel 3 auch nicht so zentral; wir haben es hier vor allem kurz angesprochen, damit Sie es irgendwann schon einmal gehört haben, falls diese Eigenschaft später (in dieser oder in anderen Vorlesungen) auftauchen sollte.

In den *Grundlagen der Funktionalanalysis* im Sommersemester 2026 werden wir das Resultat aber auf jeden Fall beweisen, denn für die Funktionalanalysis ist diese Aussage tatsächlich enorm wichtig.

Definition 3.2.1. Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $f : \Omega \to \mathbb{C}$ messbar, und sei $p \in [1, \infty)$. Man setzt

$$\|f\|_p := \Big(\int_\Omega |f|^p \, \mathrm{d}\mu \Big)^{1/p} \in [0, \infty],$$

wobei wir die Konvention $\infty^{1/p} := \infty$ verwenden.

Man beachte, dass $\|\cdot\|_p$ keine Norm ist, auch wenn die Notation das vielleicht nahelegen würde. Genauer:

- Es kann $\|\cdot\|_p$ den Wert ∞ annehmen, während eine Norm laut Definition immer nach $[0, \infty)$ abbilden muss.
- Es kann messbare Funktionen f geben, die nicht 0 sind, aber trotzdem $\|f\|_p = 0$ erfüllen. Das passiert dann (und, wie wir gleich in Theorem 3.2.2(b) sehen werden, nur dann), wenn $f = 0$ fast überall gilt.

Wenn man sich gerne einen Vektorraum basteln möchte, der aus messbaren Funktionen besteht und auf dem $\|\cdot\|_p$ eine Norm ist, muss man also einerseits alle Funktionen f wegwerfen, für die $\|f\|_p = \infty$ gilt. Und andererseits muss man einen Weg finden, diejenigen Funktionen mit $\|f\|_p = 0$ "zum Nullvektor zu machen". Beides werden im Laufe dieses Abschnitts tun.

Wir brauchen die folgende simple Beobachtung: Wenn $p \in (1, \infty)$ ist, dann gibt es genaue eine Zahl $p' \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{p'} = 1$. Das kann man schlicht nachrechnen (und erhält $p' = \frac{p}{p-1}$).

Theorem 3.2.2 (Eigenschaften von $\|\cdot\|_p$)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und seien $f, g : \Omega \to \mathbb{C}$ messbar. Sei $p \in [1, \infty)$.*

(a) *Es gilt $\big\| |f| \big\|_p = \|f\|_p$.*

(b) *Es gilt $\|f\|_p = 0$ genau dann, wenn $f = 0$ fast überall gilt.*

(c) *Für alle $\alpha \in \mathbb{C}$ gilt $\|\alpha f\|_p = |\alpha| \, \|f\|_p$.*

(d) *Es gilt die Dreiecks-Ungleichung (die man für $\|\cdot\|_p$ manchmal auch als **Minkowksi-Ungleichung** bezeichnet)*

$$\|f + g\|_p \le \|f\|_p + \|g\|_p .$$

(e) *Sei nun sogar $p \in (1, \infty)$ und sei $p' \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{p'} = 1$. Dann gilt die **Hölder-Ungleichung**[4]*

$$\int_\Omega |fg| \, \mathrm{d}\mu \le \|f\|_p \, \|g\|_{p'} .$$

[4] Dieselbe Aussage gilt übrigens auch für $p = 1$. Allerdings muss man dann $p' = \infty$ wählen um $\frac{1}{p} + \frac{1}{p'} = 1$ zu erreichen, und $\|\cdot\|_\infty$ für messbare Funktionen so zu definieren, dass es sich im Kontext dieses Abschnitts "passend" verhält, ist etwas subtil. An dieser Stelle verzichten wir deshalb darauf, verweisen aber gerne und immer wieder mit Freude auf die Vorlesung *Grundlagen der Funktionalanalysis* im nächsten Semester.

Lemma 3.2.3 (Youngsche Ungleichung)**.** *Seien $p, p' \in (1,\infty)$ mit $\frac{1}{p} + \frac{1}{p'} = 1$ und seien $a, b \in [0,\infty)$. Dann gilt*

$$ab \leq \frac{a^p}{p} + \frac{b^{p'}}{p'}.$$

Beweis. Falls $a = 0$ oder $b = 0$ gilt, ist die Behauptung klar, also seien $a, b > 0$. Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}$ ist konvex.[5] Wegen $\frac{1}{p} > 0$, $\frac{1}{p'} > 0$ und $\frac{1}{p} + \frac{1}{p'} = 1$ gilt also für alle $x, y \in \mathbb{R}$

$$\exp(\frac{1}{p}x + \frac{1}{p'}y) \leq \frac{1}{p}\exp(x) + \frac{1}{p'}\exp(y)$$

(das ist gerade die Definition von Konvexität!). Wenn man in dieser Ungleichung $x = p\log a = \log(a^p)$ und $y = p'\log b = \log(b^{p'})$ wählt, erhält man die behauptete Ungleichung. □

Nun beweisen wir Theorem 3.2.2. Beachten Sie, dass wir Aussage (e) vor (d) zeigen, weil wir (e) im Beweis von (d) verwenden.

Beweis von Theorem 3.2.2.

(a) Das folgt sofort aus $\big||f|\big| = |f|$.

(b) $\Rightarrow$ Sei $\|f\|_p = 0$. Dann gilt $0 \leq \int_A |f|^p \,\mathrm{d}\mu \leq \int_\Omega |f|^p \,\mathrm{d}\mu = 0$ für alle $A \in \mathcal{A}$, also ist $|f|^p$ eine Dichte des Nullmaßes bezüglich μ: Zugleich ist aber auch auch die konstante Nullfunktion $0 : \Omega \to [0,\infty)$ eine Dichte des Nullmaßes bezüglich μ. Laut Proposition 2.6.7 gilt somit $|f|^p = 0$ fast überall.

$\Leftarrow$ Sei $f = 0$ fast überall. Dann gibt es eine Menge $N \in \mathcal{A}$ mit $\mu(N) = 0$ und mit $f(\omega) = 0$ für alle $\omega \in N^\complement$. Also ist

$$\int_\Omega |f|^p \,\mathrm{d}\mu = \int_N |f|^p \,\mathrm{d}\mu = 0,$$

wobei die zweite Gleichheit in Aufgabe 10.1 auf Hausaufgabenblatt 10 gezeigt wurde. Somit ist $\|f\|_p = 0$.

(c) Das folgt direkt aus der Definition von $\|\cdot\|_p$.

(e) Wenn $\|f\|_p = 0$ oder $\|g\|_{p'} = 0$ ist, dann gilt laut (b) $f = 0$ fast überall oder $g = 0$ fast überall, und somit sind beide Seiten der behaupteten Ungleichung gleich 0.

Wenn $\|f\|_p > 0$ und $\|g\|_{p'} = \infty$ gilt (oder umgekehrt), dann ist die rechte Seite der behaupteten Ungleichung gleich ∞, also ist die Ungleichung wahr.

[5] Wiederholung zur Analysis 1: Können Sie den Graphen von exp skizzieren und erklären, was es geometrisch bedeutet, dass exp konvex ist? Und woher weiß man, dass exp konvex ist?

Sei nun also $0 < \|f\|_p < \infty$ und $0 < \|g\|_{p'} < \infty$. Wir setzen $\tilde{f} := f/\|f\|_p$ und $\tilde{g} := g/\|g\|_{p'}$. Wegen (c) gilt dann $\|\tilde{f}\|_p = 1$ and $\|\tilde{g}\|_{p'} = 1$ und somit $\int_\Omega |\tilde{f}|^p \, d\mu = 1$ und $\int_\Omega |\tilde{g}|^{p'} \, d\mu = 1$. Also folgt

$$\frac{1}{\|f\|_p \|g\|_{p'}} \int_\Omega |fg| \, d\mu = \int_\Omega |\tilde{f}| \, |\tilde{g}| \, d\mu \overset{\text{Lemma 3.2.3}}{\leq} \int_\Omega \frac{|\tilde{f}|^p}{p} + \frac{|\tilde{g}|^{p'}}{p'} \, d\mu = \frac{1}{p} + \frac{1}{p'} = 1.$$

(d) Für $p = 1$ folgt die Behauptung sofort aus $|f+g| \leq |f| + |g|$, also sei nun $p \in (1,\infty)$. Sei $p' \in (1,\infty)$ diejenige Zahl mit $\frac{1}{p} + \frac{1}{p'} = 1$, also $p' = \frac{p}{p-1}$. Wenn $\|f+g\|_p = 0$ gilt, ist die behauptete Ungleichung klar, also sei für den Rest des Beweises $\|f+g\|_p > 0$. Wegen

$$|f+g|^p = |f+g| \, |f+g|^{p-1} \leq |f| \, |f+g|^{p-1} + |g| \, |f+g|^{p-1}$$

gilt

$$\begin{aligned} \|f+g\|_p^p = \int_\Omega |f+g|^p \, d\mu &\leq \int_\Omega |f| \, |f+g|^{p-1} \, d\mu + \int_\Omega |g| \, |f+g|^{p-1} \, d\mu \\ &\overset{\text{(e), (a)}}{\leq} \|f\|_p \left\| |f+g|^{p-1} \right\|_{p'} + \|g\|_p \left\| |f+g|^{p-1} \right\|_{p'} \end{aligned}$$

Es ist

$$\left\| |f+g|^{p-1} \right\|_{p'} = \left(\int_\Omega |f+g|^p \, d\mu \right)^{1/p'} = \|f+g\|_p^{\frac{p}{p'}} = \|f+g\|_p^{p-1}$$

und somit folgt insgesamt

$$\|f+g\|_p^p \leq \left(\|f\|_p + \|g\|_p \right) \|f+g\|_p^{p-1}.$$

Division durch $\|f+g\|_p^{p-1}$ liefert die behauptete Ungleichung. □

Das vorangehende Theorem motiviert die folgende Definition.

Definition 3.2.4 ($\mathcal{L}^p$-Räume)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1,\infty)$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Dann setzt man[6]

$$\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K}) := \{ f : \Omega \to \mathbb{K} \mid f \text{ ist messbar und } \|f\|_p < \infty \}.$$

Aus Theorem 3.2.2 erhalten wir nun ganz leicht:

[6] Vielleicht fällt Ihnen auf, dass das Symbol $\mathcal{L}$ in $\mathcal{L}^p$ dasselbe ist, dass man auch zur Notation das Raums $\mathcal{L}(V;W)$ der beschränkten linearne Operatoren zwischen zwei normierten Räumen verwendet. Das ist aber Zufall und steht nicht in direktem Zusammenhang miteinander. Das $\mathcal{L}$ in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ steht für **Lebesgue-Raum**; das $\mathcal{L}$ in $\mathcal{L}(V;W)$ ist hingegen durch das Wort "linear" motiviert (zumindest sieht ihr Dozent das so; ob das historisch auch so war, ist eine andere Frage…).

Korollar 3.2.5 (Die $\mathcal{L}^p$-Räume sind Vektorräume und $\|\cdot\|_p$ ist eine Halbnorm auf ihnen)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1, \infty)$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Dann ist $\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ ein Untervektorraum des $\mathbb{K}$-Vektorraum aller messbaren Abbildungen von Ω nach $\mathbb{K}$. Die Einschränkung von $\|\cdot\|_p$ auf diesen Untervektorraum[7] ist eine Halbnorm.*

Beweis. Für alle $f, g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ und alle $\alpha, \beta \in \mathbb{K}$ folgt aus Theorem 3.2.2

$$\|\alpha f + \beta g\|_p \leq |\alpha| \|f\|_p + |\beta| \|g\|_p < \infty,$$

also ist $\alpha f + \beta g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$. Dass $\|\cdot\|_p$ auf diesem Raum eine Halbnorm ist, gilt wegen Theorem 3.2.2(c) und (d). □

Vorlesung 21 (Mi, 07.01.) beginnt mit Wiederholung 3.2.6

Dass es sich bei $\|\cdot\|_p$ nur um eine Halbnorm statt um eine Norm handelt, bedeutet anschaulich, dass man manche Funktionen in $\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ mit Hilfe von $\|\cdot\|_p$ nicht voneinander unterscheiden kann. Wir möchten nun gerne derartige Funktionen "miteinander identifizieren". Um diese Idee zu präzisieren verwenden wir sogenannte **Quotientenvektorräume** von Vektorräumen. Lassen Sie uns vor der Definition dieses Begriffs[8] kurz einen anderen Begriff wiederholen, den Sie bereits kennen:

Wiederholung 3.2.6 (Äquivalenzrelationen und Äquivalenzklassen)**.** Sei X eine Menge.

(1) Eine Relation $\sim$ auf X nennt man eine **Äquivalenzrelation**, wenn Sie die folgenden drei Axiome erfüllt:

   (I) **Reflexivität:** Für jedes $x \in X$ gilt $x \sim x$.

   (II) **Symmetrie:** Für alle $x, y \in X$ gilt die folgende Implikation:

$$x \sim y \quad \Rightarrow \quad y \sim x$$

   (III) **Transitivität:** Für alle $x, y, z \in X$ gilt die folgende Implikation:

$$x \sim y \wedge y \sim z \qquad \Rightarrow \qquad x \sim z$$

(2) Sei nun $\sim$ eine Äquivalenzrelation auf X. Für jedes $x \in X$ nennt man die Menge

$$\{y \in X \mid y \sim x\}$$

die **Äquivalenzklasse** von x: Jedes Elemente einer Äquivalenzklasse bezeichnet man als einen **Repräsentanten** dieser Äquivalenzklasse. Es gelten die folgenden Eigenschaft:

   (a) Zwei Elemente $x, y \in X$ liegen genau dann in derselben Äquivalenzklasse, wenn $x \sim y$ gilt.

   (b) Zwei Äquivalenzklassen sind stets disjunkt oder gleich.

[7] Die wir der einfacheren Notation halber auch mit $\|\cdot\|_p$ bezeichnen.
[8] Den Sie vermutlich bereits aus der Linearen Algebra kennen.

(c) Die Vereinigung aller Äquivalenzklassen ist gleich ganz X. Somit bilden die Äquivalenzklassen eine disjunkte Zerlegung von X.

Mithilfe von Äquivalenzrelationen kann man aus gegebenen mathematischen Stukturen neue Strukturen vom selben Typ bauen. Hier ist ein einfaches und sehr wichtiges Beispiel hierfür.[9] Für zwei Zahlen $a, b \in \mathbb{Z}$ schreiben wir $a|b$ um zu sagen, dass a die Zahl b teilt – das heißt, dass es ein $d \in \mathbb{Z}$ gibt mit $ad = b$.

Beispiel 3.2.7 ($\mathbb{Z}$ modulo $\boldsymbol{n}\mathbb{Z}$)**.** Sei $n \in \mathbb{Z} \setminus \{0\}$.

Äquivalenz modulo $\boldsymbol{n}$: Wir definieren eine Relation $\sim$ auf $\mathbb{Z}$ folgendermaßen: Für alle $a, b \in \mathbb{Z}$ sei $a \sim b$ genau dann, wenn $n|(b-a)$ gilt. Man kann nachrechnen, dass $\sim$ eine Äquivalenzrelation ist und dass für jedes $a \in \mathbb{Z}$ die Äquivalenzklasse von a gleich der Menge

$$\{a + h \mid h \in n\mathbb{Z}\} =: a + n\mathbb{Z}$$

ist. Damit gibt es insgesamt n verschiedene Äquivalenzklassen. Mit

$$\mathbb{Z}/n\mathbb{Z} := \{a + \mathbb{Z} \mid a \in \mathbb{Z}\} = \{a + \mathbb{Z} \mid a \in \{0, \dots, n-1\}\}$$

bezeichnet man die Menge aller Äquivalenzklassen.

Addition auf $\boldsymbol{\mathbb{Z}/n\mathbb{Z}}$: Wie Sie wissen ist $\mathbb{Z}$ zusammen mit der Addition $+$ eine kommutative Gruppe. Man definiert nun auf $\mathbb{Z}/n\mathbb{Z}$ eine binäre Verknüpfung[10] $+ : \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$ folgendermaßen: Für alle $a + n\mathbb{Z}, b + n\mathbb{Z} \in \mathbb{Z}/n\mathbb{Z}$ sei

$$(a + n\mathbb{Z}) + (b + n\mathbb{Z}) := (a + b) + n\mathbb{Z}.$$

Hierbei muss man sicher im Detail überlegen, dass diese Abbbildung tatsächlich Wohldefiniert ist, denn: Die beiden Äquivalenzklassen $a+n\mathbb{Z}$ und $b+n\mathbb{Z}$ kann man ja auch in der Form $a+n\mathbb{Z} = \tilde{a}+n\mathbb{Z}$ und $b+n\mathbb{Z} = \tilde{b}+n\mathbb{Z}$ mit anderen Zahlen $\tilde{a}, \tilde{b} \in \mathbb{Z}$ schreiben, und dann folgt aus der vorangehenden Definition plötzlich, dass die Summe dieser beiden Äquivalenzklassen gleich $(\tilde{a} + \tilde{b}) + n\mathbb{Z}$ ist. Wir hatten aber gesagt, dass Sie gleich $(a + b) + n\mathbb{Z}$ ist.

Das ergibt nur Sinn, wenn $(a + b) + n\mathbb{Z} = (\tilde{a} + \tilde{b}) + n\mathbb{Z}$ ist. Um zu überprüfen, dass diese Eigenschaft erfüllt ist, muss man nachrechnen, dass aus $a \sim \tilde{a}$ und $b \sim \tilde{b}$ die Äquivalenz $a + b \sim \tilde{a} + \tilde{b}$ folgt. Wir tun das an dieser Stelle nicht, weil Sie diesen Beweis aller Wahrscheinlichkeit nach bereits in den Grundlagen der Mathematik und/oder in der Linearen Algebra geführt haben[11] und weil wir einen ähnlichen Beweis nachher gleich für Quotientenvektorräume führen.

[9] Das tendenziell eher in den Bereich der Zahlentheorie und Algebra gehört als in die Analysis – wir besprechen es hier vor allem, damit Sie den Bezug der anschließend diskutierten Quotientenvektorräume zu verwandten Konzepten möglichst gut erkennen können.

[10] Die man bewusst ebenfalls mit dem Symbol + bezeichnet.

[11] In der Vorlesung sprechen wir allerdings zumindest darüber, wie man sich die Addition auf $\mathbb{Z}/n\mathbb{Z}$ plastisch vorstellen kann.

Es ist nicht schwer nachzuprüfen, dass $(\mathbb{Z}/n\mathbb{Z}, +)$ eine kommutative Gruppe ist und dass die Abbildung $\varphi : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$, $a \mapsto a + n\mathbb{Z}$ ein Gruppenhomomorphismus ist.

Man nennt $\mathbb{Z}/n\mathbb{Z}$ die **Quotientengruppe von $\mathbb{Z}$ modulo $n\mathbb{Z}$** oder die **Faktorgruppe von $\mathbb{Z}$ modulo $n\mathbb{Z}$**.

Anstelle der Gruppe $(\mathbb{Z}, +)$ und ihrer Untergruppe $n\mathbb{Z}$ betrachten wir nun denselben Spaß für Vektorräume und Untervektorräume:

Definition 3.2.8 (Äquivalenz modulo eines Untervektorraums)**.** Sei $\mathbb{K}$ ein Körper,[12] sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Wir definieren eine Relation $\sim_U$ auf V folgendermaßen: Für alle $v, w \in V$ sei $v \sim_U w$ genau dann, wenn $v - w \in U$ ist.

Proposition 3.2.9 (Äquivalenz modulo eines Untervektorraums ist eine Äquivalenzrelation)**.** *Sei $\mathbb{K}$ ein Körper, sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Dann ist $\sim_U$ eine Äquivalenzrelation auf V und für jedes $v \in V$ ist die Äquivalenzklasse gleich*

$$\{v + u \mid u \in U\} =: v + U.$$

Beweis. Dass $\sim_U$ tatsächlich eine Äquivalenzrelation ist, kann man leicht nachrechnen, indem man verwendet, dass U ein Untervektorraum ist.

Sei nun $v \in V$ fest und bezeichne $A_v \subseteq V$ die Äquivalenzklasse von v. Wir müssen $A_v = v + U$ zeigen.

"$\subseteq$" Sei $w \in A_v$. Dann ist $w \sim_U v$, das heißt, es gilt $u := w - v \in U$. Somit ist $w = v + u \in v + U$.

"$\supseteq$" Sei $w \in v + U$. Dann gibt es ein $u \in U$ mit $w = v + u$. Somit folgt $w - v = u \in U$, also ist $w \sim_U v$ und somit $w \in A_v$. □

Definition 3.2.10 (Quotientenvektorräume)**.** Sei $\mathbb{K}$ ein Körper, sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Dann setzt man

$$V/U := \{v + U \mid v \in V\}.$$

Man definiert Verknüpfungen $+ : (V/U) \times (V/U) \to V/U$ und $\cdot : \mathbb{K} \times (V/U) \to V/U$ durch

$$(v + U) + (w + U) := (v + w) + U \qquad \text{und} \qquad \alpha(v + U) := \alpha v + U$$

für alle $v + U, w + U \in V/U$ und alle $\alpha \in \mathbb{K}$. Man bezeichnet V/U zusammen mit den beiden Funktionen $+$ und $\cdot$ als den **Quotientenvektorraum von V modulo U.**

Die folgende Proposition zeigt, dass Definition 3.2.10 gerechtfertigt ist.

[12] Nachher interessieren wir uns, wie meist in der Analysis, vor allem für die beiden Körper $\mathbb{R}$ und $\mathbb{C}$.

Proposition 3.2.11 (Quotientenvektorräume)**.** *Sei $\mathbb{K}$ ein Körper, sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ ein Untervektorraum. Die beiden Funktionen $+ : (V/U) \times (V/U) \to V/U$ und $\cdot : \mathbb{K} \times (V/U) \to V/U$ aus Definition 3.2.10 sind wohldefiniert und machen V/U zu einem Vektorraum über $\mathbb{K}$ mit Nullvektor $0 + U$.*

Außerdem ist die sogenannte ***Quotientenabbildung*** *$V \to V/U$, $v \mapsto v + U$ linear.*

Beweis. Wohldefiniertheit der skalaren Multiplikation: Sei $\alpha \in \mathbb{K}$ und seien $v, \tilde{v} \in V$ zwei Repräsentanten derselben Äquivalenzklasse modulo U, das heißt, sei $v + U = \tilde{v} + U$. Wir müssen $(\alpha v) + U = (\alpha \tilde{v}) + U$ zeigen.

Weil v und $\tilde{v}$ dieselbe Äquivalenzklasse haben, gilt $v \sim_U \tilde{v}$, das heißt, es gilt $v - \tilde{v} \in U$. Weil U ein Untervektorraum von V ist, folgt

$$\alpha v - \alpha \tilde{v} = \alpha(v - \tilde{v}) \in U.$$

Also ist $\alpha v \sim_U \alpha \tilde{V}$ und somit wie behauptet $\alpha v + U = \alpha \tilde{v} + U$.

Wohldefiniertheit der Addition: Das zeigt man analog zur Wohldefiniertheit der skalaren Multiplikation.

***V/U** ist ein $\mathbb{K}$-Vektorraum mit Nullvektor $\mathbf{0} + \boldsymbol{U}$:* Die Vektorraumaxiome für V/U kann man direkt nachrechnen; wir verzichten an dieser Stelle darauf.

Linearität der Quotientenabbildung: Das folgt direkt aus der Definition der Addition und der skalaren Multiplikation auf V/U. □

Beispiel 3.2.12 ($\mathbb{R}^2$ modulo einer Ursprungsgeraden)**.** Sei $U \subseteq \mathbb{R}^2$ eine Ursprungsgerade. Die Äquivalenzklassen von $\sim_U$ sind dann die Geraden $v + U$ für $v \in \mathbb{R}^2$. Das sind genau diejenigen Geraden in $\mathbb{R}^2$, die zu U parallel sind.

Vorlesung 22 (Fr, 09.01.) beginnt mit Proposition 3.2.13

Proposition 3.2.13 (Verschwindungsraum einer Halbnorm und Quotientenraum)**.** *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei V ein Vektorraum über $\mathbb{K}$ und sei $\|\cdot\| : V \to [0, \infty)$ eine Halbnorm. Dann ist der* ***Kern***[13]

$$\ker \|\cdot\| := \{u \in V \mid \|u\| = 0\}$$

ein Untervektorraum von V. Für alle $v \in V$ und alle $u \in \ker \|\cdot\|$ gilt $\|v + u\| = \|v\|$ und die Abbildung

$$\|\cdot\|_/ : V/\ker\|\cdot\| \to [0, \infty), \qquad v + \ker\|\cdot\| \mapsto \left\| v + \ker\|\cdot\| \right\|_/ := \|v\|$$

ist wohldefiniert und eine Norm auf $V/\ker\|\cdot\|$.

Hinweis: In der Vorlesung hatten wir die Norm $\|\cdot\|_/$ auf $V/\ker\|\cdot\|$ mit demselben Symbol $\|\cdot\|$ bezeichnet wie die Halbnorm auf V. Das hat im Beweis aber etwas Verwirrung gestiftet, deshalb verwenden wir hier die Notation $\|\cdot\|_/$.

[13] Wie Sie sehen definiert man den Kern einer Halbnorm analog zum Kern einer linearen Abbildung. Dadurch sollten Sie sich aber nicht zu der falschen Annahmen verleiten lassen, dass eine Halbnorm linear wäre.

Beweis. **ker $\|\cdot\|$** *ist ein Untervektorraum von* $\boldsymbol{V}$*:*

Wegen $\|0\| = \|0 \cdot 0\| = 0 \cdot \|0\| = 0$ gilt $0 \in \ker \|\cdot\|$. Seien nun $u, u_1, u_2 \in \ker \|\cdot\|$ und $\alpha \in \mathbb{K}$. Wegen

$$0 \leq \|u_1 + u_2\| \leq \|u_1\| + \|u_2\| = 0$$

gilt $\|u_1 + u_2\| = 0$ und somit $u_1 + u_2 \in \ker \|\cdot\|$. Außerdem ist $\|\alpha u\| = |\alpha| \, \|u\| = 0$ und folglich $\alpha u \in \ker \|\cdot\|$.

Gleichheit $\|\boldsymbol{u} + \boldsymbol{v}\| = \|\boldsymbol{v}\|$ *für alle* $\boldsymbol{v} \in \boldsymbol{V}$ *und* $\boldsymbol{u} \in$ **ker** $\|\cdot\|$*:*

Sei $v \in V$ und $u \in U$. Dann ist einerseits $\|v + u\| \leq \|v\| + \|u\| = \|v\|$. Anderseits gilt ebenso

$$\|v\| = \|(v + u) + (-u)\| \leq \|v + u\| + \|-u\| = \|v + u\|$$

wegen $-u \in \ker \|\cdot\|$.

Wohldefiniertheit von $\|\cdot\|_{/}$*:*

Seien $v_1, v_2 \in V$ mit $v_1 + \ker \|\cdot\| = v_2 + \ker \|\cdot\|$. Dann gibt es ein $u \in \ker \|\cdot\|$ mit $v_1 = v_2 + u$ und somit folgt aus der im vorangehenden Punkt gezeigten Gleichheit $\|v_2\| = \|v_2 + u\| = \|v_1\|$.

$\|\cdot\|_{/}$ *ist eine Norm auf* $\boldsymbol{V}$**/ker** $\|\cdot\|$*:*

Die Abbildung $\|\cdot\|_{/}$ bildet laut ihrer Definition vom Vektorraum $V/\ker \|\cdot\|$ nach $[0, \infty)$ ab. Wir zeigen, dass Sie die drei Axiome einer Norm aus Definition 3.1.1 erfüllt:

(I) *Absolute Homogenität:* Sei $v + V/\ker \|\cdot\| \in V/\ker \|\cdot\|$ und $\alpha \in \mathbb{K}$. Dann ist

$$\|\alpha(v + \ker \|\cdot\|)\|_{/} = \|(\alpha v) + \ker \|\cdot\|\|_{/} = \|\alpha v\| = |\alpha| \, \|v\| = |\alpha| \, \|v + \ker \|\cdot\|\|_{/}.$$

(II) *Dreiecksungleichung:* Seien $v + V/\ker \|\cdot\|$, $w + V/\ker \|\cdot\| \in V/\ker \|\cdot\|$. Dann gilt

$$\begin{aligned} \|(v + \ker \|\cdot\|) + (w + \ker \|\cdot\|)\|_{/} &= \|(v + w) + \ker \|\cdot\|\|_{/} \\ &= \|v + w\| \\ &\leq \|v\| + \|w\| \\ &= \|v + \ker \|\cdot\|\|_{/} + \|w + \ker \|\cdot\|\|_{/}, \end{aligned}$$

(III) *Positive Definitheit:* Sei $v + V/\ker \|\cdot\| \in V/\ker \|\cdot\|$ und $\|v + \ker \|\cdot\|\|_{/} = 0$. Damit ist $\|v\| = 0$, also $v \in \ker \|\cdot\|$ und somit

$$v + \ker \|\cdot\| = 0 + \ker \|\cdot\|,$$

das heißt, $v + \ker \|\cdot\|$ ist der Nullvektor in $V/\ker \|\cdot\|$. □

Definition 3.2.14 (Die $\boldsymbol{L^p}$-Räume)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1, \infty)$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

(a) Man setzt

$$L^p(\Omega,\mathcal{A},\mu;\mathbb{K}) := \mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})/\ker\|\cdot\|_p.$$

Für jedes $f\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ verwenden wir der Einfachheit halber die Schreibtweise

$$\tilde{f} := f+\ker\|\cdot\|_p.$$

(b) Man stattet den Quotientenraum $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ mit der in Proposition 3.2.13 beschriebenen Norm aus und notiert diese der Einfachheit halber wieder als $\|\cdot\|_p$.

Damit ist $\|\tilde{f}\|_p = \|f\|_p$ für alle $f\in L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$, wobei $\|\tilde{f}\|_p$ die Norm der Äquivalenzklasse $\tilde{f}=f+\ker\|\cdot\|_p\in L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ bezeichnet und $\|f\|_p$ die Halbnorm der Funktion $f\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$.

Beachten Sie den etwas subtilen Unterschied in der Notation: Der Raum $\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ mit dem geschwungenen $\mathcal{L}$ besteht aus messbaren Funktionen $f:\Omega\to\mathbb{K}$ (nämlich aus denjenigen mit $\|f\|_p<\infty$), während der Raum $L^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ mit dem gedruckten L ein Quotientenvektorraum hiervon ist und somit aus Äquivalenzklassen $\tilde{f}$ mit $f\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ besteht.[14] Es ist gut sich klar zu machen, dass man die Äquivalenzklasse $\tilde{f}$ einer Funktion f genau beschreiben kann. Das folgt aus Theorem 3.2.2(b):

Korollar 3.2.15. *Sei $(\Omega,\mathcal{A},\mu)$ ein Maßraum, $p\in[1,\infty)$ und $\mathbb{K}\in\{\mathbb{R},\mathbb{C}\}$. Für jede Funktion $f\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ gilt*

$$\tilde{f} = f+\ker\|\cdot\|_p = \{g\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K}) \mid g=f \text{ fast überall}\}$$

Beweis. Der Untervektorraum $\ker\|\cdot\|_p$ von $\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ besteht aus genau denjenigen $h\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$, die $\|f\|_p=0$ erfüllen; laut Theorem 3.2.2(b) sind dies genau diejenigen $h\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$, die fastüberall gleich 0 sind.

Somit gilt: In $f+\ker\|\cdot\|_p$ liegen genau diejenigen Elemente $g\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$, die von der Form $g=f+h$ sind, wobei $h\in\mathcal{L}^p(\Omega,\mathcal{A},\mu;\mathbb{K})$ ist und h fast überall gleich 0 ist. □

Wir haben in diesem Abschnitt ja durchaus ein bisschen Aufwand betrieben um die Räume L^p zu definieren. Das zahlt sich aber langfristig aus, denn: Erstens sind p-Normen sehr praktisch (insbesondere die 2-Norm taucht sehr häufig in der Analysis auf, und auch die 1-Norm kommt recht häufig vor; für allgemeine p braucht man die p-Norm insbesondere in der Theorie der partiellen Differentialgleichungen). Zweitens ist es angenehmer mit Normen als mit Halbnormen zuarbeiten, also war es eine gute Idee, den Kern der p-Norm herauszuteilen. Und drittens haben die L^p-Räume die folgende extrem nützliche Eigenschaft:

[14] Sobald man mit diesen Räumen etwas vertrauter ist, ist man häufig bequem und macht diese Unterscheidung zwischen Funktionen und ihren Äquivalenzklassen nicht mehr ganz so rigoros. Aber um zunächst mal zu verstehen, wovon man überhaupt spricht, ist die Unterscheidung sehr wichtig.

Theorem 3.2.16 (Die L^p-Räume sind Banachräume)**.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $p \in [1, \infty)$ und $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Dann ist $L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ zusammen mit der Norm $\|\cdot\|_p$ ein Banachraum.*

Für den Beweis ist die folgende Charakterisierung der Vollständigkeit normierter Räume mithilfe von Reihenkonvergenz sehr nützlich.

Proposition 3.2.17 (Banachräume via Reihenkonvergenz)**.** *Sei V ein normierter Raum. Dann sind die folgenden Aussagen äquivalent:*

(i) *Die Norm auf V ist vollständig, das heißt, V ist ein **Banachraum**.*

(ii) *Für jede Folge (v_k) in V gilt: Falls $\sum_{k=1}^{\infty} \|v_k\| < \infty$ ist, dann ist $(\sum_{k=1}^{n} v_k)_{n\in\mathbb{N}}$ in V konvergent.*

Den Beweis der Proposition führen Sie auf Hausaufgabenblatt 12.

Beweis von Theorem 3.2.16. Sei $(\tilde{f}_k)$ eine Folge in $L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ mit $\sum_{k=1}^{\infty} \|\tilde{f}_k\|_p < \infty$. Unser Ziel ist es zu zeigen, dass ein $\tilde{f} \in L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ existiert mit $\sum_{k=1}^{n} \tilde{f}_k \to \tilde{f}$ bezüglich $\|\cdot\|_p$ für $n \to \infty$. Um solch ein $\tilde{f}$ zu finden, konstruieren wir ein $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$, dessen Äquivalenzklasse $\tilde{f}$ in $L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ die gewünschte Eigenschaft hat.

Für jedes $k \in \mathbb{N}$ bezeichnen wir mit $f_k \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ einen Repräsentanten von $\tilde{f}_k$. Dann gilt $\|\tilde{f}_k\|_p = \|f_k\|_p = \||f_k|\|_p$ für jedes k. Der vermutlich naheliegendste Gedanke ist es, f als punktweise Reihe zu definieren, das heißt $f : \Omega \to \mathbb{K}$ durch

$$f(\omega) := \lim_{n\to\infty} \sum_{k=1}^{n} f_k(\omega) \tag{3.2.1}$$

für alle $\omega \in \Omega$ zu definieren, und anschließend zu überprüfen, dass die entsprechende Konvergenz auch bezüglich der p-Norm stattfindet. So ähnlich werden wir auch vorgehen, aber es gibt dabei ein Problem: Wir wissen gar nicht, ob der Reihengrenzwert in (3.2.1) für jedes $\omega \in \Omega$ existiert.[15] Deshalb muss man ein paar Umwege nehmen um zum Ziel zukommen:

Wir summieren die Beträge der f_k punktweise auf: Lassen Sie uns eine Funktion $g : \Omega \to [0, \infty]$ definieren durch

$$g(\omega) := \sum_{k=1}^{\infty} |f_k(\omega)|$$

für alle $\omega \in \Omega$. Wenn für ein $\omega \in \Omega$ gilt, dass $g(\omega) < \infty$ ist, dann ist für dieses ω die Reihe in (3.2.1) absolut konvergent in $\mathbb{K}$ und somit konvergent in $\mathbb{K}$. Diese Beobachtung motiviert den nächsten Schritt.

Vorlesung 23 (Mi, 14.01.) beginnt im Beweis von Theorem 3.2.16 mit dem Beweisschritt, dass fast überall $g < \infty$ gilt.

[15] Genau genommen wissen wir im Moment noch nicht einmal, ob er überhaupt für irgendein ω existiert.

Wir zeigen, dass fast überall $g < \infty$ gilt Machen Sie sich zunächst klar, dass man im Allgmeinen nicht erwarten kann, dass für wirklich alle $\omega \in \Omega$ die Ungleichung $g(\omega) < \infty$ gilt:

Falls zum Beispiel Ω einen Punkt $\omega_0 \in \Omega$ mit $\{\omega_0\} \in \mathcal{A}$ und $\mu(\{\mu_0\}) = 0$ enthält, dann gibt es keine Möglichkeit zu wissen, welchen Wert die Repräsentanten f_k von $\tilde{f}_k$ an der Stelle ω_0 haben – denn zwei verschiedene Repräsentanten der Äquivalenzklasse $\tilde{f}_k$ können sich auf einer Menge vom Maß 0 um einen beliebigen Wert unterscheiden. Zum Beispiel könnte also $f_k(\omega_0) = 1$ für alle $k \in \mathbb{N}$ sein und in diesem Fall ist $g(\omega_0) = \infty$.

Wir zeigen nun aber, dass $g < \infty$ fast überall gilt – oder anders gesagt, dass $g(\omega) < \infty$ für fast alle $\omega \in \Omega$ gilt. Dazu genügt es zu zeigen, dass $\int_\Omega g^p \, \mathrm{d}\mu < \infty$ ist. Und tatsächlich gilt

$$\begin{aligned}\int_\Omega g^p \, \mathrm{d}\mu &= \int_\Omega \Big(\lim_{n\to\infty} \sum_{k=1}^n |f_k| \Big)^p \, \mathrm{d}\mu \\ &= \lim_{n\to\infty} \Big\| \sum_{k=1}^n |f_k| \Big\|_p^p \leq \lim_{n\to\infty} \Big(\sum_{k=1}^n \|f_k\|_p \Big)^p = \sum_{k=1}^\infty \|\tilde{f}_k\|_p < \infty,\end{aligned}$$

wobei wir für die Gleichheit zwischen der ersten und zweiten Zeile den Satz von der monotonen Konvergenz verwendet haben. Also besitzt die Menge $N := \{\omega \in \Omega \mid g(\omega) = \infty\} \in \mathcal{A}$ das Maß $\mu(N) = 0$.

Wir konstruieren einen Kandidaten für die Grenzfunktion f: Die Formel (3.2.1) können wir zur Definition von f nur für $\omega \in \Omega \setminus N$ verwenden, denn für $\omega \in N$ ist die Reihe in (3.2.1) zumindest nicht absolut konvergent und wir wissen nicht, ob sie für diese ω überhaupt konvergiert. Also definieren wir $f : \Omega \to \mathbb{K}$ stattdessen durch

$$f(\omega) := \lim_{n\to\infty} \sum_{k=1}^n (f_k \mathbb{1}_{\Omega\setminus N})(\omega) = \begin{cases} \lim_{n\to\infty} \sum_{k=1}^n f_k(\omega) & \text{falls} \quad \omega \in \Omega \setminus N, \\ 0 \quad \text{falls} & \omega \in N. \end{cases}$$

Wir zeigen, dass f p-integrierbar ist: Das ist einfach, denn es gilt $|f| \leq g \mathbb{1}_{\Omega\setminus N}$, und somit $\|f\|_p \leq \|g \mathbb{1}_{\Omega\setminus N}\|_p < \infty$.

Somit ist $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ und folglich besitzt f eine Äquivalenzklasse $\tilde{f} \in L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$.

Wir zeigen die gewünschte Reihenkonvergenz in der p-Norm: Wir wollen beweisen, dass $\sum_{k=1}^n \tilde{f}_k \to \tilde{f}$ im normierten Raum $L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ gilt. Dazu müssen wir

$$\Big\| \tilde{f} - \sum_{k=1}^n \tilde{f}_k \Big\|_p \to 0$$

für $n \to \infty$ zeigen. Für jedes $k \in \mathbb{N}$ ist die Funktion $f_k \mathbb{1}_{\Omega\setminus N}$ ebenfalls ein Repräsentant von f_k. Somit gilt

$$\Big\| \tilde{f} - \sum_{k=1}^n \tilde{f}_k \Big\|_p^p = \Big\| f - \sum_{k=1}^n f_k \mathbb{1}_{\Omega\setminus N} \Big\|_p^p = \int_\Omega \Big| f - \sum_{k=1}^n f_k \mathbb{1}_{\Omega\setminus N} \Big|^p \, \mathrm{d}\mu$$

$$
\begin{aligned}
&= \int_{\Omega\setminus N} \left| \lim_{\ell\to\infty} \sum_{k=n+1}^{\ell} f_k(\omega) \right|^p \mathrm{d}\mu(\omega) \le \int_\Omega \lim_{\ell\to\infty} \Big(\sum_{k=n+1}^{\ell} \big|f_k(\omega)\big| \Big)^p \mathrm{d}\mu(\omega) \\
&= \lim_{\ell\to\infty} \int_\Omega \Big(\sum_{k=n+1}^{\ell} \big|f_k(\omega)\big| \Big)^p \mathrm{d}\mu(\omega) = \lim_{\ell\to\infty} \Big\| \sum_{k=n+1}^{\ell} |f_k| \Big\|_p^p \\
&\le \lim_{\ell\to\infty} \Big(\sum_{k=n+1}^{\ell} \|f_k\|_p \Big)^p = \Big(\sum_{k=n+1}^{\infty} \|f_k\|_p \Big)^p \to 0
\end{aligned}
$$

für $n \to \infty$ wegen $\sum_{k=1}^{\infty} \|f_k\|_p < \infty$. Für die erste Gleichung haben wir verwendet, dass die Quotientenabbildung $\mathcal{L}^p(\Omega, \mathcal{A}, \mu; \mathbb{K}) \to L^p(\Omega, \mathcal{A}, \mu; \mathbb{K})$ linear ist. □

Kapitel 4

Einführung in gewöhnliche Differentialgleichungen

4.1 Wohlgestelltheit

Einstiegsfragen. (a) Können Sie sich an eine Differentialgleichung aus der Analysis 2 erinnern?

(b) Was soll Ihnen das Wort "gewöhnlich" im Titel von Kapitel 4 sagen?

(c) Warum wurde während der COVID-19-Pandemie ständig von "exponentiellem Wachstum" gesprochen? Was ist gewachsen? Weshalb ist es exponentiell gewachsen und nicht zum Beispiel quadratisch? Und was heißt das überhaupt?[1]

Beispiel 4.1.1 (Papierschiffchen und Vektorfelder)**.** Lassen Sie uns die Strömung an der Oberfläche eines Flusses mathematisch modellieren.[2] Wir stellen uns die Landschaft als flache Ebene, also als $\mathbb{R}^2$ vor, und die Flussoberfläche als eine, sagen wir offene, Teilmenge $\emptyset \neq \Omega \subseteq \mathbb{R}^2$. Als Mengen aller Zeiten, zu denen wir uns den Fluss ansehen, nehmen wir das Intervall $[0,\infty)$. Für jeden Ort $x \in \Omega$ und jede Zeit $t \in [0,\infty)$ gilt kann man sich die Geschwindigkeit $v(t,x) \in \mathbb{R}^2$ angucken, mit der sich das Wasser an der Stelle x zum Zeitpunkt t bewegt. Somit ist v eine Abbildung

$$v : [0,\infty) \times \Omega \to \mathbb{R}^2.$$

Man bezeichnet v oft als ein **Geschwindigkeitsfeld**.[3] Nehmen wir nun an, Sie setzen

[1] Es ist etwas in Mode gekommen, den Begriff "exponentielles Wachstum" als Metapher für "schnelles Wachstum" oder ähnliches zu verwenden. In der Mathematik hat der Begriff aber natürlich eine präzise Bedeutung, und diese hier tatsächlich gemeint.

[2] Wir gucken hier nur ein sehr einfaches Modell zur Motivation an. In Wirklichkeit ist natürlich alles immer viel komplizierter.

[3] Der Begriff **Feld** wird vor allem in der Physik benutzt um Größen zu beschreiben, die nicht nur eine Zahl oder ein Vektor mit endlich vielen Einträge sind, sondern die vom Ort abhängen. Weitere typische Beispiele für Felder, die Sie vielleicht schon einmal gehört haben: Elektrisches Feld, magnetisches Feld, Gravitationsfeld.

zum Zeitpunkt 0 ein kleines Papierschiffchen auf den Fluss, und zwar der Stelle $x_0 \in \Omega$.[4] Das Schiffchen wird dann von der Strömung mitgenommen. Wenn wir die Position des Schiffchens zum Zeitpunkt $t \in [0,\infty)$ mit $x_S(t)$ bezeichnen, dann ist $x_S : [0,\infty) \to \Omega$, das heißt, x_S beschreibt eine Kurve in Ω. Die Geschwindigkeit des Schiffchens zum Zeitpunkt t ist somit die Ableitung $\dot{x}_S(t) = \frac{\mathrm{d}}{\mathrm{d}t} x_S(t)$.

Weil das Schiffchen aus Papier ist und somit eine sehr geringe Masse hat, ignorieren wir hier Trägheitseffekte und tun so, als würde das Schiffchen exakt dem Geschwindigkeitsfeld f folgen – das bedeutet also $\dot{x}_S(t) = v(t, x_S(t))$ für jedes $t \in [0,\infty)$. Weil wir auch $x_S(0) = x_0$ wissen, haben nun insgesamt zwei Gleichungen, die uns etwas über die Funktion x_S sagen:

$$\begin{cases} \dot{x}_S(t) = v(t, x_S(t)) & \text{für alle } t \in [0,\infty), \\ x_S(0) = x_0. \end{cases}$$

Die erste Zeile nennt man eine **Differentialgleichung** – denn sie beschreibt, wie die – noch unbekannte – Funktion x_S und ihre Ableitung zusammenhängen. Genauer spricht man von einer **gewöhnlichen Differentialgleichung**; damit ist gemeint, dass die unbekannte Funktion x_S in der Differentialgleichung nur nach einer Variablen abgeleitet wird. Weil nur die erste Ableitung von x_S vorkommt, spricht man von einer **Differentialgleichung erster Ordnung**.

Die zweite Zeile, also die Gleichung $x_S(0) = x_0$ bezeichnet man als **Anfangsbedingung** und den Vektor $x_0 \in \Omega$ als **Anfangswert**. Beide Zeilen – also die Differentialgleichung und die Anfangsbedingung zusammengenommen – bezeichnet man als **Anfangswertproblem**.

Die Situation aus dem vorangehenden Beispiel kann man natürlich auch ganz allgemein betrachten. Wir lassen in der folgenden Definition und in den folgenden Wohlgestelltheitssätzen nicht nur Teilmengen $\Omega \subseteq \mathbb{R}^d$, sondern auch Teilmengen $\Omega \subseteq \mathbb{C}^d$ zu. Das mag aus Anwendungssicht zunächst etwas skurril erscheinen. Aber zum einen ist es aus mathematischer Sicht sehr hilfreich für die Behandlung von linearen Differentialgleichungen weiter hinten in diesem Kapitel. Zum anderen gibt es tatsächlich auch Anwendungsprobleme, die sich besser mit komplexen Zahlen als mit reellen Zahlem modellieren lassen (zum Beispiel in der Elektrotechnik oder in der Quantenmechanik). Hingegen nehmen wir I hier immer als ein Teilintervall von $\mathbb{R}$ an, da wir die Elemente von I meist als Zeiten interpretieren und da würde man intuitiv auf jeden Fall erwarten, dass sie reell sind.[5]

Vorlesung 24 (Fr, 16.01.) beginnt mit Defintion 4.1.2.

Definition 4.1.2 (Differentialgleichungen erster Ordnung)**.** Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $I \subseteq \mathbb{R}$ ein Intervall mit mehr als einem Punkt und sei $\emptyset \neq \Omega \subseteq \mathbb{K}^d$ offen. Außerdem sei $f : I \times \Omega \to \mathbb{K}^d$.[6] Seien $t_0 \in I$ und $x_0 \in \Omega$.

[4] Hier sehen Sie noch eine vereinfachende Annahme: Wir tun so, als wäre das Schiffchen nur ein Punkt.

[5] Aber Vorsicht: Es gibt kann manchmal rein aus mathematischen Gründen durchaus nützlich sein, bei der Entwicklung einer Theorie auch komplexe Zeiten zuzulassen, selbst wenn sie keine Interpretation in der realen Welt haben. Aber wir wollen es an dieser Stelle mal nicht übertreiben.

[6] Viele Bücher lassen auch etwas allgemeiner Funktionen f zu, die auf einer geeigneten Teilmenge von

(a) Die Gleichung

$$\dot{x}(t) = f(t, x(t)) \qquad \text{für alle } t \in I,$$

wobei die Funktion $x : I \to \Omega$ die Unbekannte ist, bezeichnet man als ein **gewöhnliche Differentialgleichung erster Ordnung**.

(b) Man bezeichnet

$$\begin{cases} \dot{x}(t) = f(t, x(t)) & \text{für alle } t \in I, \\ x(t_0) = x_0 \end{cases} \tag{4.1.1}$$

als ein **Anfangswertproblem**.

(c) Eine **globale Lösung** des Anfangswertproblems (4.1.1) ist eine C^1-Funktion $x : I \to \Omega$, die die beiden Eigenschaften in (4.1.1) erfüllt, das heißt, $x(t_0) = x_0$ und $\dot{x}(t) = f(t, x(t))$ für alle $t \in I$.

(d) Eine **lokale Lösung** des Anfangswertproblems (4.1.1) ist eine Funktion $x : J \to \Omega$, wobei folgendes gilt:

(I) Es ist $J \subseteq I$ ein Intervall mit mehr als einem Punkt und mit $t_0 \in J$, und t_0 ist im metrischen Raum $(I, \mathrm{d}_{\mathrm{Eukl}})$ ein innerer Punkt von J.[7]

(II) Es gilt $x(t_0) = x_0$.

(III) Die Funktion x ist C^1 und erfüllt $\dot{x}(t) = f(t, x(t))$ für alle $t \in J$.

(e) Wir nennen das Anfangswertproblem (4.1.1) **höchstens eindeutig lösbar**,[8] falls für alle lokalen Lösungen $x : J \to \Omega$ und $\tilde{x} : \tilde{J} \to \Omega$ von (4.1.1) die Gleichheit

$$x|_{J \cap \tilde{J}} = \tilde{x}|_{J \cap \tilde{J}}$$

gilt – das heißt, falls zwei lokale Lösungen immer dort, wo sie beide definiert sind, übereinstimmen.

(f) Das Anfangswertproblem (4.1.1) heißt **global eindeutig lösbar**, falls es höchstens eindeutig lösbar ist und eine globale Lösung besitzt.

Es heißt **lokal eindeutig lösbar**, falls es höchstens eindeutig lösbar ist und eine lokale Lösung besitzt.

$\mathbb{R} \times \mathbb{K}^d$ definiert sind, welche nicht unbedingt von der Form $I \times \Omega$ ist. Man erhält damit geringfügig allgemeinere Sätze, die aber zugleich auch etwas unintuitiver formuliert sind. Wir beschränken uns deshalb auf den hier genannten Fall.

[7] Anders gesagt, es gibt eine offene Menge $U \subseteq \mathbb{R}$ mit $t_0 \in U \cap J \subseteq I$.

Das bedeutet nichts weiter, als dass man von t_0 aus noch ein Stück nach links und nach rechts gehen kann ohne J zu verlassen – es sei denn, t_0 ist selbst schon ein Randpunkt von I, dann wird nur verlangt, dass man von t_0 aus in diejenige Richtung, in der I liegt, ein Stück laufen kann ohne J zu verlassen.

[8] Diese Terminologie soll betonen, dass es auch gar keine Lösung geben könnte.

Also erfüllt eine lokale Lösung des Anfangswertproblems (4.1.1) die Differentialgleichung möglicherweise nur auf einem Teilintervall von I. Natürlich ist jede globale Lösung des Anfangswertproblems (4.1.1) auch eine lokale Lösung. Beachten Sie außerdem, dass jede lokale Lösung von (4.1.1) natürlich eine globale Lösung desjenigen Anfangswertproblems ist, dass man erhält, wenn man des Intervall I in (4.1.1) entsprechend verkleinert.

Es drängt sich natürlich die Frage auf, wie man sicherstellen kann, dass ein Anfangswertproblem lokal (oder sogar global) eindeutig lösbar ist. Dazu beweisen wir im folgenden eine globale und eine lokale Variante des **Existenz- und Eindeutigkeitssatzes von Picard–Lindelöf**. Wir besprechen zuerst die globale Variante; Sie kennen sie im Wesentlichen schon aus der Analysis 2 [4, Theorem 2.2.3].[9]

Theorem 4.1.3 (Existenz- und Eindeutigkeit globaler Lösungen)**.** *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und sei $I \subseteq \mathbb{R}$ ein Intervall mit mehr als einem Punkt. Sei $f : I \times \mathbb{K}^d \to \mathbb{K}^d$ stetig und für jedes $t \in I$ sei $f(t, \cdot) : \mathbb{K}^d \to \mathbb{K}^d$ Lipschitz-stetig mit einer Lipschitzkonstanten, die unabhängig von t gewählt werden kann – das heißt, es gebe eine $L \in [0, \infty)$ mit*

$$\left\| f(t,y) - f(t,x) \right\|_2 \leq L \left\| y - x \right\|_2 \qquad \text{für alle } t \in I,\ x, y \in \mathbb{K}^d. \tag{4.1.2}$$

Dann ist das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = v(t, x(t)) & \text{für alle } t \in I, \\ x(t_0) = x_0 \end{cases}$$

global eindeutig lösbar.

Der Satz ist aus vier Gründen "global", drei davon in den Annahmen des Satzes und einer in der Konklusion:

- Es wird angenommen, dass die rechte Seite f auf ganz $I \times \mathbb{K}^d$ definiert und nicht nur auf $I \times \Omega$ für ein $\Omega \subseteq \mathbb{K}^d$.
- Es wird angenommen, dass für jedes $t \in I$ die Funktion $f(t, \cdot) : \mathbb{K}^d \to \mathbb{K}^d$ Lipschitzstetig auf ganz $\mathbb{K}^d$ ist.
- Es wird angenommen, dass man eine Lipschitz-Konstante finden kann, die zugleich für alle $t \in I$ funktioniert.
- Als Konklusion erhält man die Existent einer globalen Lösung.[10]

Zum Beweis des Satz verwenden wir zwei Resultate: Zum einen, dass man ein Anfangswertproblem in eine Integralgleichung umschreiben kann:

[9] Genau genommen haben wir das Resultat dort etwas weniger allgemein formuliert, aber der Beweis, den wir gleich führen werden, ist tatsächlich genau derselbe.

[10] Zusätzlich zur höchstens eindeutigen Lösbarkeit; aber die ist ja für globale und lokale Lösbarkeit gleich definiert.

Proposition 4.1.4 (Differentialgleichungen vs. Integralgleichungen)**.** *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $I \subseteq \mathbb{R}$ ein Intervall mit mehr als einem Punkt und sei $\emptyset \neq \Omega \subseteq \mathbb{K}^d$ offen. Außerdem sei $f : I \times \Omega \to \mathbb{K}^d$ stetig, $t_0 \in I$ und $x_0 \in \Omega$. Für jede Funktion $x : I \to \Omega$ sind folgende Aussagen äquivalent:*

(i) *Die Funktion x eine globale Lösung des Anfangswertproblems* (4.1.1) *– das heißt, x ist stetig differenzierbar und erfüllt $x(t_0) = x_0$ und $\dot{x}(t) = f(t, x(t))$ für alle $t \in I$.*

(ii) *Die Funktion x stetig und erfüllt*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s)) \, \mathrm{d}s \qquad \text{für alle } t \in I.$$

Beweis. Das folgt aus dem Hauptsatz der Differential- und Integralrechnung. □

Zum anderen verwenden wir den Banachschen Fixpunktsatz, den wir in der Analysis 2 besprochen haben [4, Theorem 1.4.3] und hier wiederholen. Zur Erinnerung: Ist Abbildung $f : M \to M$ eine Abbildung auf einer Menge M, dann heißt ein Punkt $x^* \in M$ ein **Fixpunkt von f**, falls $f(x^*) = x^*$ gilt.

Theorem 4.1.5 (Wiederholung: Der Banachsche Fixpunktsatz)**.** *Sei (M, d) ein nichtleerer und vollständiger metrischer Raum, sei $f : M \to M$ eine Lipschitz-stetige Abbildung. Für jedes $k \in \mathbb{N}$ sei $L_k \in [0, \infty)$ eine Lipschitzkonstante der k-fachen Iterierten $f^k = f \circ \cdots \circ f$. Falls*

$$\sum_{k=1}^{\infty} L_k < \infty$$

gilt, dann besitzt f genau einen Fixpunkt $x^* \in M$. Außerdem gilt für jedes $x \in M$, dass $f^k(x) \to x^*$ für $k \to \infty$.

Beweis von Theorem 4.1.3. Wir zeigen für alle reellen Zahlen a, b mit $a < b$ und $t_0 \in [a, b] \subseteq I$, dass das Anfangswertproblem

$$(*) \qquad \begin{cases} \dot{x}(t) = f(t, x(t)) & \text{für alle } t \in [a, b], \\ x(t_0) = x_0 \end{cases}$$

genau eine globale Lösung besitzt. Weshalb dies die Aussage des Theorems impliziert, besprechen wir in Aufgabe 13.2 auf Hausaufgabenblatt 13.

Mit $\mathrm{C}([a, b]; \mathbb{K}^d)$ bezeichnen wir den Vektorraum aller stetigen Funktionen von $[a, b]$ nach $\mathbb{K}^d$, ausgestattet mit der Norm $\|\cdot\|_\infty$, die durch

$$\|x\|_\infty := \sup \big\{ \|x(t)\|_2 \mid t \in [a, b] \big\}$$

für alle $x \in \mathrm{C}([a, b]; \mathbb{K}^d)$ gegeben ist. Völlig analog zu Aufgabe 12.2 auf Präsenzblatt 12 kann man zeigen, dass $\mathrm{C}([a, b]; \mathbb{K}^d)$ ein Banachraum ist.

Wir definieren eine Abbildung $T : \mathrm{C}([a,b];\mathbb{K}^d) \to \mathrm{C}([a,b];\mathbb{K}^d)$ durch

$$(Tx)(t) := x_0 + \int_{t_0}^{t} f(s, x(s))\,\mathrm{d}s$$

für alle $x \in \mathrm{C}([a,b];\mathbb{K}^d)$ und alle $t \in [a,b]$. Eine Funktion $x : [a,b] \to \mathbb{K}^d$ ist laut Proposition 4.1.4 genau dann eine globale Lösung von des Anfangswertproblems $(*)$, wenn $x \in \mathrm{C}([a,b];\mathbb{K}^d)$ gilt und $x = Tx$ gilt. Somit müssen wir nur beweisen, dass T genau einen Fixpunkt in $\mathrm{C}([a,b];\mathbb{K}^d)$ besitzt.

Dies tun wir nun mit Hilfe des Banachschen Fixpunktsatzes (Theorem 4.1.5). Um die Lipschitzstetigkeits-Voraussetzung für T^k im Theorem nachzuprüfen, betrachten wir beliebige, feste Elemente $x, \tilde{x} \in \mathrm{C}([a,b];\mathbb{K}^d)$. Wir zeigen zunächst per Induktion über k, dass für jedes $k \in \mathbb{N}_0$ die Abschätzung

$$\left\|(T^k x - T^k \tilde{x})(t)\right\|_2 \leq \frac{L^k (t-t_0)^k}{k!} \|x - \tilde{x}\|_\infty \qquad \text{für alle } t \in [t_0, b] \tag{4.1.3}$$

gilt. Für $k = 0$ kann man das sofort nachrechnen. Wenn (4.1.3) bereits für ein festes $k \in \mathbb{N}_0$ bewiesen ist, dann folgt für jedes $t \in [t_0, b]$

$$\begin{aligned}
\left\|(T^{k+1}x - T^{k+1}\tilde{x})(t)\right\|_2 &= \left\|\int_{t_0}^{t} f(s,(T^k x)(s))\,\mathrm{d}s - \int_{t_0}^{t} f(s,(T^k\tilde{x})(s))\,\mathrm{d}s\right\|_2 \\
&\overset{t\geq t_0}{\leq} \int_{t_0}^{t} \left\|f(s,(T^k x)(s)) - f(s,(T^k\tilde{x})(s))\right\|_2 \mathrm{d}s \\
&\overset{(4.1.2)}{\leq} L\int_{t_0}^{t} \left\|(T^k x - T^k\tilde{x})(s)\right\|_2 \mathrm{d}s \\
&\overset{\text{Ind.hyp.}}{\leq} L\int_{t_0}^{t} \frac{L^k (s-t_0)^k}{k!} \|x-\tilde{x}\|_\infty \,\mathrm{d}s \\
&= \frac{L^{k+1}}{k!}\int_{t_0}^{t} (s-t_0)^k \,\mathrm{d}s\, \|x-\tilde{x}\|_\infty \\
&= \frac{L^{K+1}(t-t_0)^{k+1}}{(k+1)!}\|x-\tilde{x}\|_\infty .
\end{aligned}$$

Damit ist (4.1.3) für alle $k \in \mathbb{N}_0$ gezeigt. Völlig analog zu (4.1.3) kann man dieselbe Abschätzung für $t \in [a, t_0]$ zeigen, wenn man auf der rechten Seite anstelle von $(t-t_0)^k$ den Ausdruck $(t_0 - t)^k$ verwendet. Insgesamt gilt somit

$$\left\|(T^k x - T^k \tilde{x})(t)\right\|_2 \leq \frac{L^k |t_0 - t|^k}{k!} \|x - \tilde{x}\|_\infty$$

für alle $t \in [a,b]$ und alle $k \in \mathbb{N}_0$. Indem wir in dieser Ungleichung das Supremum über alle $t \in [a,b]$ nehmen, erhalten wir[11]

$$\left\|T^k x - T^k \tilde{x}\right\|_\infty \leq \frac{L^k (b-a)^k}{k!} \|x - \tilde{x}\|_\infty ,$$

[11] Genau genommen ist die folgende Ungleichung nicht ganz optimal. Anstelle von $(b-a)^k$ erhält man auf der rechten Seite sogar den kleineren Ausdruck $(b-t_0)^k \vee (t_0-a)^k$. Das nützt für den Rest des Arguments aber nichts und $(b-a)^k$ lässt sich übersichtlicher aufschreiben.

also ist T^k Lipschitz-stetig mit der Lipschitzkonstanten $L_k := \frac{L^k(b-a)^k}{k!}$. Wegen $\sum_{k=0}^{\infty} L_k = e^{L(b-a)} < \infty$ ist der Banachsche Fixpunktsatz (Theorem 4.1.5) anwendbar und liefert wie gewünscht, dass T genau einen Fixpunkt in $\mathrm{C}([a,b];\mathbb{K}^d)$ besitzt. □

Das folgende Theorem ist eine lokale Variante des Satzes von **Picard–Lindelöf**. Der Beweis folgt genau demselben Prinzip wie der Beweis von Theorem 4.1.3, aber er ist deutlich technischer aufzuschreiben. Deshalb verzichten wir darauf ihn an dieser Stelle zu besprechen.

Vorlesung 25 (Mi, 21.01.) beginnt mit Theorem 4.1.6.

Theorem 4.1.6 (Existenz und Eindeutigkeit lokaler Lösungen)**.** *Sei $\mathbb{K} \in \{\mathbb{R},\mathbb{C}\}$, sei $I \subseteq \mathbb{R}$ ein Intervall mit mehr als einem Punkt und $\Omega \subseteq \mathbb{K}^d$ offen. Sei $f : I \times \Omega \to \mathbb{K}^d$ stetig und es erfülle die folgende lokale Lipschitzbedingung: für jedes $s \in I$ und jedes $z \in \Omega$ gebe es ein $L \geq 0$ derart,*

$$\left\| f(t,y) - f(t,x) \right\|_2 \leq L \left\| y - x \right\|_2 \tag{4.1.4}$$

für alle $t \in I$ und alle $x, y \in \Omega$ gilt, die genügend nahe bei s beziehungsweise z liegen. Dann besitzt das Anfangswertproblem

$$\begin{cases} \dot{x}(t) = v(t, x(t)) & \text{für alle } t \in I, \\ x(t_0) = x_0 \end{cases}$$

eine lokale Lösung und sie ist eindeutig.

Beispiel 4.1.7 (Ein Anfangswertproblem mit lokaler aber ohne globale Lösung)**.** Wir betrachten das $\mathbb{R}$-wertige Anfangswertproblem

$$\begin{cases} \dot{x}(t) = x(t)^2 & \text{für } t \in \mathbb{R}, \\ x(0) = 1 \end{cases}$$

Das Anfangswertproblem ist lokal eindeutig lösbar und es besitzt die lokale Lösung $x : (-\infty, 1) \to \mathbb{R}$, $x(t) = \frac{1}{1-t}$. Aber es besitzt keine globale Lösung.

Beweis. Behauptete Lösung des Anfangsswertproblems: Man kann sofort nachrechnen, dass die gegebene Funktion x eine lokale Lösung des Anfangswertproblems ist.

Lokal eindeutige Lösbarkeit: Die rechte Seite der Differentialgleichung ist gegeben durch $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$, $f(t,y) = y^2$ und ist somit stetig. Wir rechnen nach, dass f die lokale Lipschitz-Bedingung (4.1.4) aus Theorem 4.1.6 erfüllt. Sei dazu $s \in I$ und $z \in \mathbb{R}$. Für alle $t \in [s-1, s+1]$ und $x, y \in [z-1, z+1]$ gilt

$$\left| f(t,y) - f(t,x) \right| = \left| y^2 - x^2 \right| = \left| y + x \right| \left| y - x \right| \leq \left(2|z| + 2 \right) \left| y - x \right|.$$

Also gilt die Abschätzung (4.1.4) mit $L = 2|z| + 2$.[12] Somit ist Theorem 4.1.6 anwendbar und sagt aus, dass das Anfangswertproblem lokal eindeutig lösbar ist.

[12] In diesem Beispiel hängt L also von z ab, aber nicht von s.

Nicht-Existenz einer globalen Lösung: Angenommen, es gäbe eine globale Lösung $\tilde{x} : \mathbb{R} \to \mathbb{R}$. Da das Anfangswertproblem lokal eindeutig lösbar ist, ist es inesbesondere höchstes eindeutig lösbar und folglich gilt $\tilde{x}|_{(-\infty,1)} = x$.

Für $t \in [0,1)$ folgt darauf $\tilde{x}(t) = \frac{1}{1-t} \to \infty$ für $t \to 1$; das widerspricht aber der Stetigkeit von $\tilde{x}$ im Punkt 1. □

Im vorangehenden Beispiel sehen Sie, dass es selbst für eine ziemlich einfache rechte Seite lästig sein kann, eine der Lipschitz-Bedingungen (4.1.2) und (4.1.4) per Hand nachzuprüfen. In vielen Fällen geht es mit folgendem Resultat einfacher.

Proposition 4.1.8 (Lipschitz-Bedingungen via stetiger Differenzierbarkeit)**.** *Sei $\emptyset \neq I \subseteq \mathbb{R}$ ein offenes Intervall.*[13]

(a) *Sei $f : I \times \mathbb{R}^d \to \mathbb{R}^d$ eine stetig differenzierbare Funktion. Wenn die Ableitung $\mathrm{D}f : I \times \mathbb{R}^d \to \mathbb{R}^{d\times(1+d)}$ beschränkt ist, dann erfüllt f die globale Lipschitz-Bedingung* (4.1.2) *mit $L = \sup\{\|\mathrm{D}f(t,x)\|_{2\leftarrow 2} \mid t \in I,\ x \in \mathbb{R}^d\}$.*

(b) *Sei $\emptyset \neq \Omega \subseteq \mathbb{R}^d$ offen und sei $f : I \times \Omega \to \mathbb{R}^d$ stetig differenzierbar. Dann erfüllt f die lokale Lipschitz-Bedingung* (4.1.4)*.*

Beweis. (a) Wir setzen

$$L := \sup\{\|\mathrm{D}f(t,x)\|_{2\leftarrow 2} \mid t \in I,\ x \in \mathbb{R}^d\}.$$

Für jedes $t \in I$ und jedes $x \in \mathbb{R}^d$ sei $C(t,x) \in \mathbb{R}^{d\times d}$ dieselbe Matrix wie $\mathrm{D}f(t,x)$, aber ohne die erste Spalte. Dies ist die Ableitung der Abbildung $f(t,\cdot) : \mathbb{R}^d \to \mathbb{R}^d$ an der Stelle x. Wir bezeichnen mit $P \in \mathbb{R}^{(1+d)\times d}$ die Matrix, die in der ersten Zeile nur Nullen enthält, und unterhalb der ersten Zeile die $d \times d$-Einheitsmatrix. Dann ist $C(t,x) = \mathrm{D}f(t,x)P$ für alle $t \in I$ und alle $x \in \mathbb{R}^d$; wegen $\|P\|_{2\leftarrow 2} = 1$ gilt außerdem

$$\|C(t,x)\|_{2\leftarrow 2} \leq \|\mathrm{D}f(t,x)\|_{2\leftarrow 2}\,\|P\|_{2\leftarrow 2} \leq L.$$

Seien nun $t \in I$ und $x, y \in \mathbb{R}^d$ fest. Wir verbinden x und y durch die Kurve $\gamma : [0,1] \to \mathbb{R}^d$, $\gamma(s) = x + (y-x)s = sy + (1-s)x$, die geraden die Strecke von x nach y abläuft. Wegen des Hauptsatzes der Differential- und Integralrechnung gilt

$$f(t,y) - f(t,x) = f(t,\gamma(1)) - f(t,\gamma(0)) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s} f(t,\gamma(s))\,\mathrm{d}s = \int_0^1 C(t,\gamma(s))(y-x)\,\mathrm{d}s,$$

wobei wir für die letzte Gleichheit die Kettenregeln verwendet haben. Somit ist

$$\|f(t,y) - f(t,x)\|_2 \leq \int_0^1 \|C(t,\gamma(s))(y-x)\|_2\,\mathrm{d}s$$

[13] Die Bedingung offen brauchen wir hier nur, damit $I \times \mathbb{R}^d$ in $\mathbb{R}^{1+d}$ offen und somit für Funktionen $I \otimes \mathbb{R}^d \to \mathbb{R}^d$ klar ist, was unter ihrer Ableitung (sofern sie existiert) zu verstehen ist.

Für konkrete Differentialgleichungen hat man häufig die Situation $I = [0,\infty)$ und dieses Intervall ist nicht offen. Allerdings kann man dann oft die Funktion f, die die rechte Seite der Differentialgleichung beschreibt, auf ein etwas größeres offenes Intervall fortsetzen und somit doch Proposition 4.1.8 anwenden.

$$\leq \int_0^1 \left\| C(t,\gamma(s)) \right\|_{2\leftarrow 2} \left\| y-x \right\|_2 \mathrm{d}s$$
$$\leq \int_0^1 L \left\| y-x \right\|_2 \mathrm{d}s = L \left\| y-x \right\|_2,$$

was die Behauptung zeigt.

(b) Das beweist man analog zu (a). Man muss nur verwenden, dass $\mathrm{D}f$ auf abgeschlossenen Kugeln, die in $I \times \Omega$ enthalten sind, beschränkt ist, weil stetige Funktionen auf kompakten Mengen beschränkt sind. □

Mit Hilfe von Proposition 4.1.8 kann man sich das Leben in Beispiel 4.1.7 viel einfacher machen: Weil die Abbildung $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f(t,y) = y^2$ stetig differenzierbar ist, erfüllt sie laut Proposition 4.1.8(b) die lokale Lipschitz-Bedingung 4.1.4 in Theorem 4.1.6.

Wie Sie in Beispiel 4.1.7 gesehen haben, gibt es Anfangswertprobleme, die lokal eindeutig lösbar sind, die aber keine globale Lösung besitzen. Da liegt es nahe zu fragen, ob man besser verstehen kann, für welche Zeiten die Differentialgleichung eine Lösung besitzt. Wenn man, unter den Voraussetzungen des lokalen Existenz- und Eindeutigkeitstheorems 4.1.6, zwei lokale Lösungen $x : J \to \Omega$ und $\tilde{x} : \tilde{J} \to \Omega$ desselben Anfangswertproblems hat, dann stimmen sie auf $J \cap \tilde{J}$ überein, da das Theorem insbesondere impliziert, dass das Anfangswertproblem höchstens eindeutig lösbar ist. Man kann jetzt die beiden Lösungen "verkleben", das heißt man bastelt sich eine Lösung $\hat{x} : J \cup \tilde{J} \to \Omega$, die man auf J als x und auf $\tilde{J}$ als $\tilde{x}$ definiert. Wegen $x|_{J\cap\tilde{J}} = \tilde{x}|_{J\cap\tilde{J}}$ ist $\hat{x}$ wohldefiniert und man hat somit eine Lösung, die auf einem größeren Intervall definiert ist. Dasselbe kann man nicht nur mit zwei Lösungen tun, sondern sogar mit unendlich vielen – und somit sogar mit allen (!) Lösungen.[14] Damit erhält man eine Lösung des Anfangswertproblems, die auf einem größt möglichen Teilintervall von I definiert ist. Wenn I offen ist, kann man dieses Intervall sehr genau beschrieben. Wir tun das im folgenden Theorem, beweisen es in dieser Vorlesung aber nicht.[15]

Vorlesung 26 (Fr, 23.01.) beginnt mit Theorem 4.1.9.

Theorem 4.1.9. *Seien die Voraussetzungen von Theorem 4.1.6 erfüllt und sei I offen. Dann gibt es eindeutig bestimmte Elemente $t_- < t_+$ in $[-\infty,\infty]$ mit folgenden Eigenschaften:*

(a) *Es ist $(t_-, t_+) \subseteq I$.*

(b) *Das Anfangswertproblem* (4.1.4) *aus Theorem 4.1.6 besitzt eine lokale Lösung $x : (t_-, t_+) \to \Omega$.*

(c) *Für jede lokale Lösung $\tilde{x} : \tilde{J} \to \Omega$ des Anfangswertproblems* (4.1.4) *gilt $\tilde{J} \subseteq (t_-, t_+)$.*

*Man nennt (t_-, t_+) das **maximale Existenzintervall** des Anfangswertproblems* (4.1.4).[16] *Für seinen rechten Endpunkt t_+ gilt mindestes eine der folgenden drei Eigenschaften:*[17]

[14] Ganz ähnlich, wie es in Aufgabe 13.2(b) auf Hausaufgabenblatt 13 passiert.

[15] Der Beweis ist nicht so wahnsinnig kompliziert, aber er dauert ein bisschen und wir wollen die Zeit lieber nutzen um andere Themen zu gewöhnlichen Differentialgleichungen zu besprechen.

[16] Wobei diese Bezeichnung eigentlich terminologisch etwas ungenau ist. Präziser wäre es, (t_-, t_+) das **größte Existenzintervall** des Anfangswertproblems (4.1.4) zu nennen.

[17] Beachten Sie, dass auch mehrere dieser Eigenschaften zugleich auftreten können.

(1) $t_+ = \sup I$.

(2) $\|x(t)\|_2 \to \infty$ *für* $t \to t_+$.

(3) $\liminf_{t \to t_+} \operatorname{dist}(x(t), \partial\Omega) = 0$.

Analoges gilt für t_-.

In Beispiel 4.1.7 kann man gut erkennen, dass dort die Option (2) in Theorem 4.1.9 eintritt.

Zum Abschluss dieses Abschnitts führen wir noch einen Begriff ein, den wir in diesem Kapitel häufig verwenden werden. Wenn das Vektorfeld, entlang dessen sich die Lösungen bewegen, sich nicht mit der Zeit ändert, dann nennt man eine Differentialgleichung **autonom**. Genauer:

Definition 4.1.10 (Autonome Differentialgleichung)**.** Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $I \subseteq \mathbb{R}$ ein Intervall mit mehr als einem Punkt, und sei $\emptyset \neq \Omega \subseteq \mathbb{K}^d$ offen. Sei $f : I \times \Omega \to \mathbb{K}^d$. Die Differentialgleichung

$$\dot{x}(t) = f(t, x(t)) \quad \text{für alle } t \in I$$

heißt **autonom**, falls es eine Funktion $g : \Omega \to \mathbb{K}^d$ gibt derart, dass $f(t, y) = g(y)$ für alle $y \in \Omega$ gilt – das heißt, falls sich die Differentialgleichung in der Form

$$\dot{x}(t) = g(x(t)) \quad \text{für alle } t \in I$$

schreiben lässt.

Eine Bemerkung zum Schluss dieses Abschnitts: Beachten Sie, dass wir in diesem Abschnitt überhaupt nicht über gewöhnliche Differentialgleichung von Ordnung 2 oder mehr gesprochen haben. Praktischerweise lassen sich solche Differentialgleichungen jedoch stets auf eine gewöhnliche Differentialgleichung erster Ordnung zurückführen. Das besprechen wir anhand eines Beispiels auf dem Hausaufgabenblatt 14.

4.2 Gewöhnliche Differentialgleichungen in einer Dimension

Einstiegsfragen. (a) Was versteht man unter logistischem Wachstum?

(b) Wieso genau hängt eigentlich die Durchschnittstemperatur auf der Erde vom CO2-Gehalt der Atmosphäre ab?

In diesem Abschnitt besprechen wir Differentialgleichungen, in denen die gesuchte Funktion reellwertig ist. Besonders gut lassen sich Differentialgleichungen untersuchen, deren rechte Seite $f(t, x(t))$ sich schreiben lässt als Produkt einer Funktion, die nur von t abhängt und einer Funktion, die nur von $x(t)$ abhängt. Das illustriert man am besten mit einem Beispiel:

Beispiel 4.2.1 (Eine Differentialgleichung mit getrennten Variablen)**.** Wir betrachten das Anfangswertproblem

$$(*) \qquad \begin{cases} \dot{x}(t) = tx(t) & \text{für alle } t \in \mathbb{R}, \\ x(0) = x_0 \end{cases}$$

für ein $x_0 \in \mathbb{R}$, wobei die gesuchte Funktion x nach $\mathbb{R}$ abbildet.

Weil die Abbildung $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $(t, y) \mapsto ty$ C^1 ist, folgt aus Proposition 4.1.8(b), dass die lokale Lipschitzbedingung (4.1.4) aus Theorem 4.1.6 erfüllt ist. Somit ist das Anfangswertproblem $(*)$ lokal eindeutig lösbar. In Aufgabe 13.2 auf Präsenzblatt 13 haben Sie mit Hilfe der Picard–Lindelöf–Iteration die Lösung explizit ausgerechnet,[18] aber das war etwas mühselig. Im Folgenden zeigen wir eine Technik, mit der man die Lösung des Anfangswertproblems in diesem Beispiel (und auch in einigen anderen einfachen Beispielen) sehr leicht berechnen kann.[19]

Sehen wir uns zunächst den Fall $x_0 > 0$ an. Sei $x : J \to \mathbb{R}$ eine lokale Lösung von $(*)$, wobei wir J als eine solch kleine Umgebung von 0 wählen, dass $x(t) > 0$ für alle $t \in J$ gilt. Solch ein J gibt es wegen der Stetigkeit von x.

Für $t \in J$ können wir die Differentialgleichung $\dot{x}(t) = tx(t)$ durch $x(t)$ teilen – hier verwenden wir natürlich, dass $x(t)$ eine Zahl ist und nicht etwas ein Vektor in $\mathbb{R}^d$ mit $d \geq 2$ – und erhalten somit

$$\frac{\dot{x}(t)}{x(t)} = t.$$

Wir bennen die Zeitvariable in dieser Gleichung nun anders – sagen wir zum Beispiel s – und integrieren dann beiden Seiten von 0 bis $t \in J$ bezüglich der Integrationsvariablen s. Das liefert

$$\int_0^t \frac{\dot{x}(s)}{x(s)} \, \mathrm{d}s = \int_0^t s \, \mathrm{d}s.$$

Die rechte Seite dieser Gleichung können wir einfach ausrechnen, und auf der linken Seite führen wir die Substitution $u := x(s)$ durch und erhalten somit wegen $x(0) = x_0$

$$\int_{x_0}^{x(t)} \frac{1}{u} \, \mathrm{d}u = \frac{1}{2} t^2,$$

Wegen $x_0 > 0$ und $x(t) > 0$ ist $u \mapsto \log u$ eine Stammfunktion das Integranden $u \mapsto \frac{1}{u}$ und somit folgt

$$\log \frac{x(t)}{x_0} = \frac{1}{2} t^2.$$

Das Gleichung können wir nun nach $x(t)$ auflösen und erhalten somit

$$x(t) = \exp(\frac{1}{2} t^2) x_0$$

[18] Nur auf dem Intervall $[0, 1]$, aber dieselbe Rechnung funktioniert auch auf jedem anderen kompakten Intervall in $\mathbb{R}$, und somit erhält man insgesamt sogar eine globale Lösung.

[19] Natürlich gelangen wir am Ende zum selben Ergebnis, wie in Aufgabe 13.2 auf Präsenzblatt 13.

für alle $t \in J$.

Wenn man nun dreist ist und die Funktion $\tilde{x} : \mathbb{R} \to \mathbb{R}$ durch dieselbe Formel definiert – also durch $\tilde{x}(t) = \exp(\frac{1}{2}t^2)x_0$ für alle $t \in \mathbb{R}$, dann kann man einfach nachrechnen, dass $\tilde{x}$ sogar eine globale Lösung von $(*)$ ist. Ebenso kann man leicht nachrechnen, dass dies auch dann stimmt, wenn $x_0 \leq 0$ ist – obwohl Teile unserer Rechnung verwendet hatten, dass $x(t) > 0$ ist.[20]

Die Methode, die wir im vorangehenden Beispiel verwendet haben, nennt man **Trennung der Variablen**, aus dem Grund, der direkt vor dem Beispiel erläutert wurde.

Einfache Differentialgleichungen kann man mit dieser Methode manchmal explizit lösen. Darauf sollten wir uns aber nicht allzuviel einbilden:

- Viele – man könnte auch sagen, die meisten – Differentialgleichungen, die bei der Modellierung realer Phänomene auftreten, sind so kompliziert, dass man sie nicht explizit lösen kann. In großem Detail über explizite Lösungsverfahren von Differentialgleichungen zu sprechen, ist deshalb von zweifelhaftem Nutzen.

- Sie haben gesehen, dass die Argumentation in Beispiel 4.2.1 ein wenig verschlungen war. Wir haben zuerst angenommen, dass $x_0 > 0$ ist und nur Zeiten t in der Nähe von 0 betrachtet, damit $x(t) > 0$ gilt – und am Ende haben wir festgestellt, dass die so gefundene Lösung auch ohne diese Annahmen richtig ist.

  Man kann darauf auf zwei verschiedene Arten reagieren: Entweder man schreibt ein präzises Theorem auf, das genau beschreibt, was bei der Trennung der Variablen passieren kann, und das soeben beschrieben Problem löst. Solch ein Theorem finden Sie zum Beispiel in [3, Satz 6.2.5]. Der Vorteil ist, dass man dann ein präzises Resultat hat und die Trennung der Variablen sich nicht mehr so wackelig anfühlt. Der Nachteil ist, dass das Theorem relativ technisch ist.

  Oder man geht, sagen wir, pragmatisch vor, rechnet bei einer Trennung der Variablen einfach darauf los, ignoriert dabei potentielle Probleme (zum Beispiel, dass man eigentlich aufpassen muss, nicht durch Null zu teilen) und überprüft dann am Ende einfach, ob die gefundene Lösung wirklich eine Lösung ist. Der Vorteil ist, dass das oft sehr schnell geht. Der Nachteil ist, dass man eventuell verwirrt ist, wenn etwas mal nicht so funktioniert, wie erwartet.

  Im Rahmen dieser Vorlesung begnügen wir uns mit der zweiten Möglichkeit.

Trennung der Variablen kann man in einer Dimension inbesondere bei autonomen Differentialgleichungen anwenden. Wir illustrieren das mit einem ganz einfachen Beispiel, dessen Lösung Sie auch ohne weiteres direkt erraten können (und das ein Spezialfall von Aufgabe 13.1(a) auf Hausaufgabenblatt 13 ist.

[20] Das illustriert sehr schön das folgende Phänomen: Nur weil die Argumente, die wir verwende, bestimmte Voraussetzungen benötigen, ist noch nicht ausgeschlossen, dass die Konklusion vielleicht auch ohne diese Voraussetzungen gelten könnte.

Beispiel 4.2.2 (Lineare Differentialgleichung in einer Dimension)**.** Betrachten wir das skalare Anfangswertproblem

$$\begin{cases} \dot{x}(t) = x(t) & \text{für alle } t \in \mathbb{R}, \\ x(0) = 2. \end{cases}$$

Dividieren durch $x(t)$, Ersetzen der Zeitvariable durch s, und Integration von 0 bis t liefert

$$\int_0^t \frac{\dot{x}(s)}{x(s)} \, \mathrm{d}s = \int_0^t 1 \, \mathrm{d}s.$$

Mit der Substitution $u := x(s)$ und durch Verwendung der Anfangsbedingung $x(0) = 2$ folgt somit

$$\int_2^{x(t)} \frac{1}{u} \, \mathrm{d}u = t.$$

Wenn wir zunächst einmal davon ausgehen, dass $x(t) > 0$ für alle t ist, erhalten wir links als Stammfunktion $u \mapsto \log u$ und somit

$$\log \frac{x(t)}{2} = t,$$

also $x(t) = 2e^t$.

Nachdem wir diese Formel etwas hemdsärmelig – und ohne auf Details zu achten – hergeleitet haben, definieren wir motiviert durch diese Formel die Funktion $x : \mathbb{R} \to \mathbb{R}$, $x(t) = 2e^t$. Man prüft leicht nach, dass x tatsächlich das Anfangswertproblem löst.[21]

Ein etwas interessanteres Beispiel für eine autonome skalare Differentialgleichung, die man durch Trennung der Variablen explizit lösen kann, ist die logistische Wachstumsgleichung. Wir besprechen Sie in den Übungen.

Obwohl sich Trennung der Variablen mit autonomen, $\mathbb{R}$-wertigen Differentialgleichungen im Prinzip immer anwenden lässt,[22] bedeutet das nicht, dass man die Lösung damit immer wirklich explizit ausrechnen kann. Das Problem ist zum einen, dass das Integral, dass bei der Trennung der Variablen auftritt, zu kompliziert sein könnte, um es explizit zum berechnen. Und zum anderen muss man nach dem Berechnung des Integrals die so erhaltende Gleichung noch nach $x(t)$ auflösen, was ebenfalls nicht immer möglich ist. Wir illustrieren diese Probleme anhand der folgenden Differentialgleichung, die ein einfaches zeitabhängiges Klimamodell für die Erde beschreibt. Das Beispiel folgt im Wesentlichen der Darstellung in Abschnitt 3.2.1 der Bachelorthesis [6].

[21] Wie oben gesagt, hätte man diese Formel für x aber auch einfach erraten können.

[22] Zumindest, so lange die rechte Seite genügend regulär ist.

Beispiel 4.2.3 (Zeitabhängige Strahlungsbilanzgleichung)**.** Wir betrachten ein stark vereinfachtes Klimamodell zur Beschreibung der Temperatur auf der Erde. Zu jedem Zeitpunkt t bezeichne $T(t)$ die Durchschnittstemperatur der Atmosphäre auf der Erdoberfläche,[23] angegeben in Kelvin. Die von der Erde absorbierte Leistung an Sonnenstrahlung ist durch die Formel

$$R_{\text{in}} = \pi r_{\text{E}}^2 S(1-\alpha)$$

gegeben. Hierbei ist $r_{\text{E}} > 0$ der Erdradius, $S > 0$ die sogenannte **Solarkonstante**, and $\alpha \in (0,1)$ die sogenannte **Albedo** der Erde, die angibt, wieviel von der einkommenden Strahlung wieder zurück ins All reflektiert wird.[24] Die von der Erde abgestrahle Leistung hängt von der momentanen mittleren Temperatur $T(t)$ ab. Sie ist durch die Formel

$$R_{\text{out}} = 4\pi r_{\text{E}}^2 \varepsilon\sigma T(t)^4,$$

gegeben.[25] Hierbei ist $\sigma > 0$ die sogenannte *Stefan–Boltzmann-Konstante* aus der Thermodynamik, und $\varepsilon \in (0,1)$ die **Emissivität** der Erdatmosphäre bezeichnet, die angibt, welcher Anteil der von der Oberfläche abgestrahlen Strahlung tatsächlich die Erde verlässt. Es ist $\varepsilon < 1$, weil ein Teil der von der Oberfläche abstrahlten Energie von der Atmosphäre reflektiert wird. Ein höherer Anteil an Treibhausgasen führt zu einem niedrigeren Wert von ε.

Die Energieänderung pro Zeiteinheit beträgt somit

$$R_{\text{in}} - R_{\text{out}} = \pi r_{\text{E}}^2 \big(S(1-\alpha) - 4\varepsilon\sigma T(t)^4\big).$$

Zugleich ist die Energieänderung propotional zur Temperaturänderung desjenigen Stoffes, der die Energie speichert.

Wenn man davon ausgeht, dass die Energie von der untersten Atmosphärenschicht, der sogenannte **Troposphäre**, aufgenommen wird, dann ist die Proportionalitätskonstante gleich der spezifischen Wärmekapazität von Luft mal dem Volumen der Troposphäre, also

$$R_{\text{in}} - R_{\text{out}} = 4\pi r_{\text{E}}^2 h\rho C\dot{T}(t),$$

wobei $h > 0$ die Höhe der Troposphäre bezeichnet, $\rho > 0$ die Dichte der Luft, und $C > 0$ die spezifische Wärme von Luft. Insgesamt erhält man somit die Differentialgleichung[26]

$$\dot{T}(t) = \frac{S(1-\alpha)}{4h\rho C} - \frac{\varepsilon\sigma}{h\rho C} T(t)^4.$$

[23] Wobei "Durchschnittstemperatur" bedeutet, dass wir einerseits über die komplette Erdoberfläche mitteln.

[24] Die Albedo hängt eng mit der Oberflächenbeschaffenheit der Erde zusammen, weil dunkle Oberflächen mehr Strahlung reflektieren als helle Flächen.

[25] Für diesen Zusammenhang braucht man, dass die Temperatur tatsächlich relativ zum absoluten Nullpunkt gemessen wird – deshalb haben wir vorausgesetzt, dass $T(t)$ in Kelvin angegeben wird.

[26] Aus der der Erdradius sich bemerkenswerterweise herauskürzt.

Mit den Abkürzungen

$$a := \frac{S(1-\alpha)}{4h\rho C} > 0 \qquad \text{und} \qquad b := \frac{\varepsilon\sigma}{h\rho C} > 0$$

lautet die Differentialgleichung also

$$\dot{T}(t) = a - bT(t)^4. \tag{4.2.1}$$

Die Differentialgleichung (4.2.1) ist autonom, also liegt es nahe sie durch Trennung der Variablen zu lösen. Lassen Sie uns als Startzeitpunkt 0 wählen und die Starttemperatur T_0 nennen. Dann liefert Trennung der Variablen die Gleichung

$$\int_0^t \frac{\dot{T}(s)}{a - bT(s)^4}\,\mathrm{d}s = \int_0^t 1\,\mathrm{d}s.$$

Mit der Substitution $u := T(s)$ erhält man dann

$$\int_{T_0}^{T(t)} \frac{1}{a - bu^4}\,\mathrm{d}u = t.$$

Eine Stammfunktion von $u \mapsto \frac{1}{a-bu^4}$ kann man mit Hilfe von Partialbruchzerlegung tatsächlich explizit bestimmen (und in [6] wird diese Rechnung tatsächlich durchgezogen), aber das führt zu fürchterlich komplizierten Formeln – und am Ende kann man die Gleichung, die man so erhält, trotzdem nicht explizit nach $T(t)$ auflösen, da die Gleichung zu kompliziert ist.

Die Diskussion am Ende des vorangehenden Beispiels legt nahe, dass wir, um Gleichungen wie (4.2.1) verstehen zu können, Methoden zu brauchen, mit denen man Differentialgleichungen analysieren kann ohne ihre Lösung explizit auszurechnen. Für Differentialgleichungen, bei denen die gesuchte Funktion reellwertig ist – und die wir uns im aktuellen Abschnitt ansehen –, klappt das besonders gut. Damit werden wir den Rest dieses Abschnitts verbringen.

Wir brauchen zuerst den Begriff eines **Gleichgewichts**. Obwohl wir uns im aktuellen Abschnitt 4.2 auf Differentialgleichungen in einer Dimension konzentrieren, führen wir Begriff der Vollständigkeit halber auch für $\mathbb{K}^d$-wertige Differentialgleichungen ein. Wir definieren den Begriff nur für autonome Differentialgleichungen.

Definition 4.2.4 (Gleichgewichte)**.** Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $\emptyset \neq \Omega \subseteq \mathbb{K}^d$ offen, und sei $g : \Omega \to \mathbb{K}^d$. Ein Punkt $x^* \in \Omega$ heißt **Gleichgewicht** oder **Equilibrium** oder **stationärer Punkt** der Differentialgleichung[27]

$$\dot{x}(t) = g(x(t)) \quad \text{für alle } t \in \mathbb{R},$$

falls $g(x^*) = 0$ ist (oder, äquivalent dazu, falls die konstante Funktion $\mathbb{R} \to \Omega$, $t \mapsto x^*$ die Differentialgleichung erfüllt).

[27] Als Zeitintervall verwenden wir hier ganz $\mathbb{R}$. Da die Differentialgleichung autonom ist, gibt es keinen naheliegenden Grund, die Differentialgleichung nur auf einem Teilintervall von $\mathbb{R}$ zu studieren.

Gleichgewichte sind also Positionen in Ω, von denen aus der Zustand sich nicht von allein ändert. Das Verhalten von $\mathbb{R}$-wertigen autonomen Differentialgleichungen kann man sehr gut mit Hilfe des folgenden Theorems beschreiben.

Theorem 4.2.5 (Autonome skalare Differentialgleichungen)**.** *Sei $\emptyset \neq \Omega \subseteq \mathbb{R}$ ein offenes Intervall*[28] *und sei $g : \Omega \to \mathbb{R}$ stetig differenzierbar.*[29] *Sei $t_0 \in \mathbb{R}$ und $x_0 \in \Omega$ und sei $x : (t_-, t_+) \to \Omega$ diejenigen Lösung des Anfangswertproblems*

$$\begin{cases} \dot{x}(t) = g(x(t)) & \text{für alle } t \in \mathbb{R}, \\ x(t_0) = x_0, \end{cases}$$

die auf dem maximalen Existenzintervall[30] *definiert ist. Es gilt:*

(a) *Sei $g(x_0) = 0$. Dann ist $(t_-, t_+) = \mathbb{R}$ und $x(t) = x_0$ für alle $t \in \mathbb{R}$.*

(b) *Sei $g(x_0) > 0$. Dann gilt $g(x(t)) > 0$ für alle $t \in (t_-, t_+)$ und x ist streng monoton steigend. Außerdem tritt genau einer der beiden Fälle auf:*

   (1) *Die Funktion g hat keine Nullstelle rechts von x_0. In diesem Fall gilt $x(t) \to \sup \Omega$ für $t \uparrow t_+$.*

   (2) *Die Funktion g hat eine Nullstelle rechts von x_0. Dann ist $t_+ = \infty$ und $x(t)$ konvergiert für $t \to \infty$ gegen die kleinste Nullstelle von g, die rechts von x_0 liegt.*

Vorlesung 27 (Mi, 28.01.) beginnt mit Theorem 4.2.5(c).

(c) *Sei $g(x_0) < 0$. Dann gilt $g(x(t)) < 0$ für alle $t \in (t_-, t_+)$ und x ist streng monoton fallend. Außerdem tritt genau einer der beiden Fälle auf:*

   (1) *Die Funktion g hat keine Nullstelle links von x_0. In diesem Fall gilt $x(t) \to \inf \Omega$ für $t \uparrow t_+$.*

   (2) *Die Funktion g hat eine Nullstelle links von x_0. Dann ist $t_+ = \infty$ und $x(t)$ konvergiert für $t \to \infty$ gegen die größte Nullstelle von g, die links von x_0 liegt.*

Für den Beweis des Theorems ist die folgende Beobachtung über die Konvergenz von Lösungen von autonomen Differentialgleichungen nützlich. Sie gilt, wie Sie sehen, auch für Differentialgleichungen im $\mathbb{K}^d$:

Lemma 4.2.6. *Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $\emptyset \neq \Omega \subseteq \mathbb{K}^d$ offen, sei $g : \Omega \to \mathbb{K}^d$ stetig und $t_0 \in \mathbb{R}$. Sei $x : [t_0, \infty) \to \Omega$ eine C^1-Funktion, die die Differentialgleichung*

$$\dot{x}(t) = g(x(t)) \quad \text{für alle } t \in [t_0, \infty)$$

[28] Denken Sie daran, dass Ω anschaulich den Ort beschreibt, nicht etwa die Zeit. Der folgende Satz lässt sich etwas leichter formulieren, wenn Ω ein Intervall ist. Außerdem kann die Lösung einer Differentialgleichung wegen ihrer Stetigkeit nicht zwischen zwei nichtverbundenen Intervallen hin- und herspringen, deshalb genügend es sich die Situation anzusehen, in der Ω ein Intervall ist.

[29] Oder allgemeiner lokal Lipschitzstetig, das heißt für jedes $z \in \Omega$ sei die Einschränkung von g auf eine genügend kleine Umgebung von z Lipschitzstetig (wobei die Lipschitzkonstante von z abhängen kann).

[30] Siehe Theorem 4.1.9.

erfüllt. Wenn $x(t)$ für $t \to \infty$ gegen ein $x^* \in \Omega$ konvergiert, dann ist x^* ein Equilibrium, das heißt, $g(x^*) = 0$.

Beweis. Wegen der Stetigkeit von g folgt aus der Differentialgleichung, dass $\dot{x}(t) \to g(x^*)$ für $t \to \infty$ gilt. Sei $\varepsilon > 0$. Es genügt $\|g(x^*)\|_2 \leq 3\varepsilon$ zu zeigen.

Sei $\tau \geq t_0$ so groß, dass $\|x(t) - x^*\|_2 \leq \varepsilon$ und $\|\dot{x}(t) - g(x^*)\|_2 \leq \varepsilon$ für alle $t \geq \tau$ gilt. Dann folgt

$$g(x^*) - \big(x(\tau+1) - x(\tau)\big) = g(x^*) - \int_\tau^{\tau+1} \dot{x}(t)\,\mathrm{d}t = \int_0^1 g(x^*)\,\mathrm{d}s - \int_0^1 \dot{x}(s+\tau)\,\mathrm{d}s$$

und somit

$$\begin{aligned}\|g(x^*)\|_2 &\leq \|g(x^*) - \big(x(\tau+1) - x(\tau)\big)\|_2 + \|x(\tau+1) - x(\tau)\|_2 \\ &\leq \int_0^1 \|g(x^*) - \dot{x}(s+\tau)\|_2\,\mathrm{d}s + \|x(\tau+1) - x^*\|_2 + \|x^* - x(\tau)\|_2 \leq 3\varepsilon. \qquad \square\end{aligned}$$

Beweis von Theorem 4.2.5. Da g als C^1 vorausgesetzt wurde, ist auch die Abbildung $f : \mathbb{R}\times\Omega \to \mathbb{R}$, $f(y,t) = g(y)$ eine C^1-Funktion. Somit folgt aus Proposition 4.1.8(b) und Theorem 4.1.6, dass das Anfangswertproblem lokal eindeutig lösbar ist. Wir zeigen jetzt die behaupteten Eigenschaften der Lösung $x : (t_-, t_+) \to \Omega$.

(a) Wenn $g(x_0) = 0$ gilt, dann ist x_0 ein Equilibrium der Differentialgleichung und somit die konstante Funktion $\mathbb{R} \to \Omega$, $t \mapsto x_0$ ein Lösung des Anfangswertproblems. Wegen der Maximalität von (t_-, t_+) folgt $(t_-, t_+) = \mathbb{R}$ und wegen der höchstens eindeutigen Lösbarkeit folgt, dass x die konstante Funktion mit Wert x_0 ist.

(b) Wir zeigen zuerst, dass $g(x(t)) > 0$ für alle $t \in (t_-, t_+)$ gilt. Wäre dem nicht so, dann gäbe es wegen des Zwischenwertsatzes ein $t_1 \in (t_-, t_+)$ mit $g(x(t_1)) = 0$. Damit sind aber x und die die konstante Funktion $\mathbb{R} \to \Omega$, $t \mapsto x(t_1)$ Lösungen desselben Anfangswertproblems

$$\begin{cases}\dot{\tilde{x}}(t) = g(\tilde{x}(t)) & \text{für } t \in \mathbb{R},\\ \tilde{x}(t_1) = x(t_1).\end{cases}$$

Da dieses Anfangsproblem höchstens eindeutig lösbar ist, folgt $x(t) = x(t_1)$ für alle $t \in (t_-, t_+)$ und somit insbesondere $x(t_0) = x(t_1)$, also $g(x(t_0)) = g(x(t_1)) = 0$. Widerspruch.

Also ist $g(x(t)) > 0$ für alle $t \in (t_-, t_+)$. Da x die Differentialgleichung löst, folgt $\dot{x}(t) > 0$ für alle $t \in (t_-, t_+)$; somit ist x streng monoton steigend.

Insbesondere konvergiert $x(t)$ somit für $t \uparrow t_+$ gegen ein $x^* \in \mathbb{R} \cup \{\infty\}$. Wir zeigen nun in den beiden Fällen (1) und (2) die behaupteten Eigenschaften.

(1) Besitze g keine Nullstelle rechts von x_0. Wir müssen $x^* = \sup\Omega$ zeigen. Weil x komplett in Ω verläuft, gilt natürlich $x^* \leq \sup\Omega$. Angenommen, es wäre $x^* < \sup\Omega$.

Da Ω ein Intervall ist und $x^* > x_0$ gilt, folgt dann $x^* \in \Omega$. Die Charakterisierung von t_+ in Theorem 4.1.9 (angewendet auf $I = \mathbb{R}$) zeigt nun, dass $t_+ = \infty$ gilt. Somit zeigt Lemma 4.2.6, dass $g(x^*) = 0$ ist. Das widerspricht der Annahme, dass g rechts von x_0 keine Nullstelle besitzt.

(2) Besitze g eine Nullstelle rechts von x_0. Wegen der Stetigkeit von g besitzt g dann auch eine kleinste Nullstelle rechts von x_0; nennen wir sie x_1. Wegen $g(x(t)) > 0$ für alle $t \in (t_-, t_+)$ folgt $x(t) \neq x_1$ für alle $t \in (t_-, t_+)$. Da x monoton steigt und $x(t_0) = x_0 < x_1$ ist, folgt somit aus dem Zwischenwertsatz $x_0 \leq x(t) < x_1$ für alle $t \in [t_0, t_+)$. Daraus folgt zum einen $x_0 < x^* \leq x_1$; und zum anderen folgt daraus wegen der Charakterisierung von t_+ in Theorem 4.1.9 (angewendet auf $I = \mathbb{R}$), dass $t_+ = \infty$ gilt.

Lemma 4.2.6 zeigt somit, dass x^* eine Nullstelle von g ist. Da x_1 die kleinste Nullstelle von g ist, die rechts von x_0 liegt, folgt $x_1 = x^*$.

(c) Das beweist man völlig analog zu (b). □

Beispiel 4.2.7 (Zeitabhängige Strahlungsbilanzgleichung, noch einmal)**.** Wir betrachten noch einmal die Differentialgleichung (4.2.1) aus Beispiel 4.2.1, also die Gleichung

$$\dot{T}(t) = a - bT(t)^4,$$

wobei $T(t)$ die Durchschnittstemperatur der Atmosphäre an der Erdoberfläche zum Zeitpunkt t beschreibt und die Zahlen $a, b > 0$ sich aus physikalischen Konstanten berechnen lassen. Sei[31] $\Omega = (0, \infty)$ und $g : \Omega \to \mathbb{R}$, $g(y) = a - by^4$. Sei $T_0 > 0$ die Temperatur zum Zeitpunkt 0. Das Anfangswertproblem

$$\begin{cases} \dot{T}(t) = a - bT(t)^4 & \text{für alle } t \in \mathbb{R}, \\ T(0) = T_0 \end{cases}$$

ist lokal eindeutig lösbar und besitzt eine lokale Lösung $x : [0, \infty) \to (0, \infty)$. Außerdem gilt $x(t) \to T^*$ für $t \to \infty$, wobei

$$T^* := \left(\frac{a}{b}\right)^{1/4} = \left(\frac{S(1-\alpha)}{4\varepsilon\sigma}\right)^{1/4} \in (0, \infty)$$

die einzige Nullstelle von g ist.

Beweis. Das folgt aus Theorem 4.2.5, indem man die drei Fälle $T_0 < T^*$, $T_0 = T^*$ und $T_0 > T^*$ unterscheidet. Das diskutiert man am besten anhand einer Skizze; wir tun das in der Vorlesung. □

Im vorangehenden Beispiel sind unter anderem folgende Beobachtungen interessant:

[31] Da die Temperatur in Kelvin angegeben wird, muss sie über dem absoluten Nullpunkt liegen. Deshalb ist $\Omega = (0, \infty)$ eine vernünftige Wahl.

- Die Gleichgewichtstemperatur T^*, gegen die $T(t)$ für $t \to \infty$ konvergiert, steigt, wenn die Emissivität ε kleiner wird. Man kann bereits in diesem sehr einfachen Modell also mathematisch einen Effekt sehen, denn man natürlich auch anschaulich erwarten würde: Wenn die Erdatmosphäre weniger Strahlung ins Weltall emitieren lässt[32] und somit die Emissivität sinkt, dann steigt die Durchschnittstemperatur, bei der sich das Gleichgewicht zwischen absorbierter und emittierter Strahlung einstellt.

- Die Gleichgewichtstemperatur T^* ist natürlich genau diejenige Temperatur, bei der absorbierte und emitierte Strahlung die Differenz 0 haben und sich somit ausgleichen, das heißt T^* erfüllt die stationäre Strahlungsbilanzgleichung

$$0 = R_{\text{in}} - R_{\text{out}} = \pi r_{\text{E}}^2 \big(S(1-\alpha) - 4\varepsilon\sigma(T^*)^4\big).$$

  Indem man direkt mit dieser stationären Gleichung beginnt, kann man die Formel für T^* auch direkt ohne eine Differentialgleichung herleiten. Das Schöne an der Diskussion der Differentialgleichung in Beispiel 4.2.7 ist aber, dass man sehen kann, dass die Durchschnittstemperatur der Erde in diesem Modell tatsächlich gegen die Gleichgewichtstemperatur konvergiert.

4.3 Autonome lineare Differentialgleichungen und die Matrix-e-Funktion

Einstiegsfragen. (a) Wann ist eine Matrix $A \in \mathbb{C}^{d\times d}$ diagonalisierbar?

(b) Was versteht man unter der Jordan-Normal-Form einer Matrix $A \in \mathbb{C}^{d\times d}$?

Vorlesung 28 (Fr, 30.01.) beginnt mit Abschnitt 4.3.

In diesem Abschnitt betrachten wir den besonders wichtigen Spezialfall autonomer Differentialgleichungen, bei denen die rechte Seite eine lineare Funktion des aktuellen Zustands ist. Das heißt, wir betrachten für eine gegebene Matrix $A \in \mathbb{C}^{d\times d}$ das $\mathbb{C}^d$-wertige Anfangswertproblem

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{für alle } t \in \mathbb{R}, \\ x(0) = x_0, \end{cases} \tag{4.3.1}$$

mit einem Startwert $x_0 \in \mathbb{C}^d$.[33] Im Fall $d = 1$ ist es sehr leicht, die Lösung dieses Anfangswertproblems zu sehen: es ist $x(t) = e^{tA}x_0$ für alle $t \in \mathbb{R}$.

Lassen Sie uns nun dreist sein und versuchen, ob dasselbe auch für größere Dimensionen d klappt. Dazu muss man natürlich zuerst festlegen, was es heißt, eine Matrix in die Exponentialfunktion einzusetzen.

[32] Zum Beispiel aufgrund eines höheren Konzentration an Treibhausgasen als in der Vergangenheit.

[33] Man kann natürlich auch den Skalarkörper $\mathbb{R}$ betrachten. Für das Anfangswertproblem 4.3.1 macht das aber keinen Unterschied, weil man jede reelle Matrix ja immer auch als komplexe Matrix auffassen kann.

Definition 4.3.1 (Matrix-Exponentialfunktion)**.** Sei $A \in \mathbb{C}^{d\times d}$.[34] Man setzt

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

Die Abbildung $\mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$, $B \mapsto \exp(B)$ nennt man die **Matrix-Exponentialfunktion**.

In Aufgabe 14.2(a) und (b) auf Hausaufgabenblatt 14 zeigen Sie, dass diese Reihe tatsächlich konvergiert.[35] Die Matrix-Exponentialfunktion ist, wie wir vor ihrer Definition bereits angedeutet haben, sehr nützlich um das Anfangswertproblem (4.3.1) zu lösen:

Theorem 4.3.2 (Lösung linearer autonomer Anfangswertprobleme)**.** *Sei $A \in \mathbb{C}^{d\times d}$ und $x_0 \in \mathbb{C}^d$.*

(a) *Das Anfangswertproblem* (4.3.1) *ist global eindeutig lösbar und seine globale Lösung $x : \mathbb{R} \to \mathbb{C}^d$ ist gegeben durch*

$$x(t) = e^{tA}x_0 \quad \text{für alle } t \in \mathbb{R}.$$

(b) *Allgemeiner ist für $t_0 \in \mathbb{R}$ das Anfangswertproblem*

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{für alle } t \in \mathbb{R}, \\ x(t_0) = x_0, \end{cases}$$

global eindeutig lösbar und seine globale Lösung $x : \mathbb{R} \to \mathbb{C}^d$ ist gegeben durch

$$x(t) = e^{(t-t_0)A}x_0 \quad \text{für alle } t \in \mathbb{R}.$$

(c) *Die Abbildung $\mathbb{R} \to \mathbb{C}^{d\times d}$, $t \mapsto e^{tA}$ ist stetigdifferenzierbar mit Ableitung $\frac{\mathrm{d}}{\mathrm{d}t}e^{tA} = Ae^{tA} = e^{tA}A$ für alle $t \in \mathbb{R}$.*

(d) *Es ist $e^{0\cdot A} = \mathrm{id}$ und für alle $s, t \in \mathbb{R}$ gilt $e^{(s+t)A} = e^{sA}e^{tA}$.*

Beweis. (a) Das beweisen Sie in Aufgabe 14.2(c) auf Hausaufgabenblatt 14.

(b) Das folgt aus (a) mit Hilfe der Kettenregel.

(c) Sei $k \in \{1, \dots d\}$, sei $e_k \in \mathbb{C}^d$ der k-te kanonische Einheitsvektor, und sei $x : \mathbb{R} \to \mathbb{C}^d$, $x(t) = e^{tA}e_k$. Für jedes $t \in \mathbb{R}$ ist dann $x(t)$ die k-te Spalte von e^{tA}, und laut (a) ist x stetig differenzierbar mit

$$\frac{\mathrm{d}}{\mathrm{d}t}e^{tA}e_k = \frac{\mathrm{d}}{\mathrm{d}t}x(t) = Ax(t) = Ae^{tA}e_k.$$

[34] Es ist wichtig, dass die Matrix A quadratisch ist, weil wir in dieser Definition Potenzen A^k für $k \in \mathbb{N}_0$ verwenden. Wenn A nicht quadratisch ist, ist zum Beispiel das Produkt $A^2 = A \cdot A$ gar nicht definiert.

[35] Bitte bearbeiten Sie diese Aufgabe unbedingt. Sie ist nicht besonders schwer, aber wichtig um ein gutes Verständnis der Matrix-Exponentialfunktion zu bekommen.

Also ist die k-te Spalte von e^{tA} bezüglich t stetig differenzierbar und die Ableitung ist die k-te Spalte von Ae^{tA}).

Das zeigt, dass $t \mapsto e^{tA}$ stetig differenzierbar ist mit Ableitung Ae^{tA}. Dass A für jedes $t \in \mathbb{R}$ mit e^{tA} kommutiert, folgt aus der Definition der Matrix-Exponentialfunktion, weil A mit jeder seiner Potenzen kommutiert.

(d) Die Eigenschaft $e^{0\cdot A} = \mathrm{id}_{\mathbb{C}^d}$ folgt sofort aus der Definition der Matrix-e-Funktion (mit $A^0 = \mathrm{id}$). Um die zweite Gleichung zu zeigen, fixieren wir $s \in \mathbb{R}$ und $x_0 \in \mathbb{C}^d$ und wählen $x : \mathbb{R} \to \mathbb{C}^d$, $x(t) := e^{(t+s)A}x_0$. Dann folgt aus (b), dass x die Differentialgleichung $\dot{x}(t) = Ax(t)$ für $t \in \mathbb{R}$ löst. Außerdem gilt $x(0) = e^{sA}x_0$, das heißt x die die globale Lösung des Anfangswertproblems

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{für alle } t \in \mathbb{R}, \\ x(0) = e^{sA}x_0. \end{cases}$$

Andererseits ist dieses Anfangswertproblem laut (a) global eindeutig lösbar mit $x(t) = e^{tA}e^{sA}x_0$ für alle $t \in \mathbb{R}$. Damit ist $e^{(t+s)A}x_0 = e^{tA}e^{sA}x_0$ für alle $t \in \mathbb{R}$ gezeigt. Weil $s \in \mathbb{R}$ und $x_0 \in \mathbb{C}^d$ beliebig waren, folgt $e^{(t+s)A} = e^{tA}e^{sA}$ für alle $s, t \in \mathbb{R}$, wie behauptet. □

Man kann die Eigenschaften der Matrix-e-Funktion aus Theorem 4.3.2 übrigens stattdessen auch mit denselben Argumenten zeigen, mit denen man sie in der Analysis 1 für die skalare e-Funktion beweist. Aber es ist ja ganz nett zu sehen, dass man die Eigenschaften alternativ auch aus der Theorie der gewöhnlichen Differentialgleichungen erhält.

Beispiel 4.3.3 (Exponentialfunktion einiger Matrizen)**.** Betrachten wir die Matrix

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2\times 2}.$$

Dann gilt

$$e^{tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \quad \text{für alle } t \in \mathbb{R}.$$

Beweis. Wenn man die Potenzen $A^0, \ldots, A^4$ von A ausrechnet, erhält man $A^0 = \mathrm{id}$, $A^1 = A$, $A^2 = -\mathrm{id}$, $A^3 = -A$ und $A^4 = \mathrm{id} = A^0$. Somit folgt für alle $t \in \mathbb{R}$

$$\begin{aligned} e^{tA} &= \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \sum_{k=0}^{\infty} \frac{t^{2k+1}A^{2k+1}}{(2k+1)!} + \sum_{k=0}^{\infty} \frac{t^{2k}A^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{(2k+1)!} A + \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{(2k)!} \mathrm{id} \\ &= \sin(t)A + \cos(t)\,\mathrm{id} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

□

Die Wirkung der Matrix e^{tA} aus dem vorangehenden Beispiel auf reelle Vektoren – das heißt, auf Elemente von $\mathbb{R}^2$ – kann man besonders gut sehen, wenn man e^{tA} auf die beiden kanonischen Einheitsvektoren e_1, e_2 anwendet. Es ist

$$e^{tA}e_1 = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \quad \text{und} \quad e^{tA}e_2 = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

für alle $t \in \mathbb{R}$ – das heißt e^{tA} dreht Vektoren im $\mathbb{R}^2$ um den Winkel t (im Bogenmaß angegeben) gegen den Uhrzeigersinn. Insbesondere bedeutet das, dass die Lösung des Einfangswertproblems

$$\begin{cases} \dot{x}(t) = Ax(t) & \text{für alle } t \in \mathbb{R}, \\ x(0) = x_0, \end{cases}$$

die laut Theorem 4.3.2(a) durch $x(t) = e^{tA}x_0$ für alle $t \in \mathbb{R}$ gegeben ist, im Kreis verläuft.

Beispiel 4.3.4 (Exponentialfunktion eines $\mathbf{2} \times \mathbf{2}$-Jordanblocks)**.** Es sei

$$J = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Dann gilt für alle $t \in \mathbb{R}$

$$e^{tJ} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Beweis. Es gilt $J^2 = 0$ und somit $J^k = 0$ für alle $k \geq 2$. Somit folgt aus der Definition der Matrix-Exponentialfunktion

$$e^{tJ} = \sum_{k=0}^{1} \frac{(tJ)^k}{k!} = \mathrm{id} + tJ = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

für alle $t \in \mathbb{R}$. □

Um für möglichst viele Matrizen die Matrixexponentialfunktion ausrechnen zu können, kann man die folgenden beiden Propositionen mit der Diagonalisierung beziehungsweise der Jordan-Normalform von Matrizen kombinieren, die Sie aus der Linearen Algebra kennen.

Vorlesung 29 (Mi, 04.02.) beginnt mit Proposition 4.3.5.

Proposition 4.3.5 (Matrix-e-Funktion für (Block)Diagonalmatrizen)**.**

(a) *Seien $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$. Dann gilt*

$$\exp\left(t \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_d \end{pmatrix} \right) = \begin{pmatrix} e^{t\lambda_1} & & \\ & \ddots & \\ & & e^{t\lambda_d} \end{pmatrix}$$

für alle $t \in \mathbb{R}$.

(b) *Allgemeiner seien nun $d_1, \dots, d_\ell \in \mathbb{N}$ und $d := d_1 + \cdots + d_\ell$. Seien $B_1 \in \mathbb{C}^{d_1 \times d_1}$, …, $B_\ell \in \mathbb{C}^{d_\ell \times d_\ell}$ Wenn wir die $d \times d$-Blockdiagonalmatrix mit Blockdiagonaleinträgen $B_1, \dots, B_\ell$ in die Matrix-e-Exponential einsetzen, erhalten wir*

$$\exp\left(t\begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_d \end{pmatrix}\right) = \begin{pmatrix} e^{tB_1} & & \\ & \ddots & \\ & & e^{tB_d} \end{pmatrix}$$

Beweis. (a) Für 2×2-Diagonalmatrizen rechnen Sie das in Aufgabe 14.2(d) auf Hausaufgabenblatt 14 nach. Für größere Dimension ist die Rechnung im Wesentlichen dieselbe.

(b) Der Beweis ist derselbe wie für Diagonalmatrizen; man muss lediglich verwenden, dass man Blockmatrizen blockweise multiplizieren kann. □

Die Matrix-Exponentialfunktion ist verträglich mit einem Basiswechsel im $\mathbb{C}^d$ – das heißt, anders ausgedrückt, mit Ähnlichkeitstransformationen von Matrizen:

Proposition 4.3.6 (Matrix-Exponentialfunktion unter Ähnlichkeit)**.** *Seien $D, T \in \mathbb{C}^{d\times d}$ und sei T invertierbar. Dann gilt*

$$e^{tTDT^{-1}} = Te^{tD}T^{-1}$$

für alle $t \in \mathbb{R}$.

Beweis. Das beweisen Sie auf Präsenzblatt 15 in Aufgabe 15.1(b). □

Beispiel 4.3.7 (Blockform durch Umsortieren)**.** Betrachten wir die Matrix

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{pmatrix} \in \mathbb{C}^{3\times 3}.$$

Wenn man die zweite und die dritte Zeile von A vertauscht, dann befindet sich die Matrix in Blockdiagonalgestalt mit einem 2×2-Block und einem 1×1-Block.

Genauer: Wir betrachten die Permutationematrix

$$P := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Sie ist invertierbar und gleich ihrer eigenen Inversen, das heißt $P^{-1} = P$. Eine Matrix C von links mit P zu multiplizieren vertauscht die zweite und die dritte Zeile von C; und C von rechts mit P zu multiplizieren, vertauscht die zweite und die dritte Spalte von C. Somit gilt

$$P^{-1}AP = PAP = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 0 & -1 & \\ 1 & 0 & \\ & & -2 \end{pmatrix} =: B.$$

Laut Propositionen 4.3.6 und 4.3.5 sowie Beispiel 4.3.3 folgt für alle $t \in \mathbb{R}$

$$e^{tA} = e^{tPBP^{-1}} = Pe^{tB}P^{-1} = P\begin{pmatrix} \cos t & -\sin t & \\ \sin t & \cos t & \\ & & e^{-2t} \end{pmatrix}\underbrace{P^{-1}}_{=P} = \begin{pmatrix} \cot t & 0 & -\sin t \\ 0 & e^{-2t} & 0 \\ \sin t & 0 & \cos t \end{pmatrix}.$$

Ebenso kann man mit Hilfe der Propositionen 4.3.6 und 4.3.5(a) die Matrix-Exponentialfunktion einer diagonalisierbaren Matrix ausrechnen (sofern man die Eigenwerte und Eigenvektoren ausrechnen kann). Hier ist ein Beispiel für diese Vorgehen; dasselbe Beispiel behandeln Sie auch auf Hausaufgabenblatt 14, aber mit einer anderen Methode.

Beispiel 4.3.8 (Ein Matrix-e-Funktion via Diagonalisierung)**.** Wir betrachten die Matrix

$$B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Dann gilt für alle $t \in \mathbb{R}$

$$e^{tB} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

Beweis. In Aufgabe 14.2(e) auf Hausaufgabenblatt 14 rechnen Sie diese Formel direkt mit Hilfe der Definition der Matrix-Exponentialfunktion nach.[36] Hier besprechen wir, wie man stattdessen mit Hilfe von Diagonalisierung zumselben Ziel kommt.

Man überprüft leicht, dass B die beiden Eigenwerte -1 und 1 hat mit zugehörigen Eigenvektoren[37]

$$v_{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \qquad v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Also gilt $B = TDT^{-1}$ mit

$$D = \begin{pmatrix} -1 & \\ & 1 \end{pmatrix} \qquad \text{und} \qquad T = \begin{pmatrix} v_{-1} & v_1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Weil die beiden Spalten von T Euklidische Norm 1 haben und senkrecht aufeinander stehen, ist T orthogonal, das heißt es gilt $T^{-1} = T^{\mathrm{T}}$. Somit erhalten wir für alle $t \in \mathbb{R}$

$$\begin{aligned} e^{tB} &= T\exp\left(t\begin{pmatrix} -1 & \\ & 1 \end{pmatrix}\right)T^{-1} = T\begin{pmatrix} e^{-t} & \\ & e^{t} \end{pmatrix}T^{\mathrm{T}} \\ &= \frac{1}{2}\begin{pmatrix} e^t + e^{-t} & e^t - e^{-t} \\ e^t - e^{-t} & e^t + e^{-t} \end{pmatrix} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}. \end{aligned}$$

□

[36] Die Rechnung ist ganz ähnlich wie in Beispiel 4.3.3; man muss nur bemerken, dass $B^2 = \mathrm{id}$ gilt.

[37] Den Vorfaktor $\frac{1}{\sqrt{2}}$ braucht man natürlich nicht unbedingt; aber sie sorgen dafür, dass die zugehörige Transformationsmatrix orthogonal wird, was die Rechnung ein wenig erleichter.

4.4 Inhomogenitäten und Faltung

Einstiegsfragen. (a) Was versteht man in der linearen Algebra unter einer **Inhomogenität** bei einer linearen Gleichung?

(b) Woran denken Sie bei dem Wort "Resonanzkatastrophe"?

In diesem letzten Abschnitt der Vorlesung behandeln wir lineare Differentialgleichungen der Form

$$\dot{x}(t) = Ax(t) + f(t) \quad \text{für } t \geq 0,$$

wobei $A \in \mathbb{C}^{d\times d}$ eine Matrix und $f : [0,\infty) \to \mathbb{C}^d$ eine gegebene stetige Funktion ist.

Lassen Sie uns zuerst ein Beispiel besprechen um zu sehen, woher solch eine **Inhomogenität** f in einem Model stammen kann.

Beispiel 4.4.1 (Federpendel mit externer Kraft)**.** In Aufgabe 15.2 auf Präsenzblatt 15 betrachten Sie ein Federpendel mit Masse $m > 0$ und Federhärte $D > 0$. Für die Position $x(t)$ und die Geschwindigkeit $v(t) = \dot{x}(t)$ des Pendels zum Zeitpunkt t leiten Sie in dieser Aufgabe die Differentialgleichung

$$\begin{pmatrix}\dot{x}(t)\\ \dot{v}(t)\end{pmatrix} = \begin{pmatrix}0 & 1\\ -\frac{D}{m} & 0\end{pmatrix}\begin{pmatrix}x(t)\\ v(t)\end{pmatrix}$$

für alle t her.

Nun ändern wir das Model folgendermaßen ab: Wir nehmen nun an, das Federpendel hat zum Zeitpunkt 0 die Position x_0 und die Geschwindigkeit v_0. Wir wollen das Federpendel erst vom Zeitpunkt 0 an betrachten und nehmen an, dass zu jedem Zeitpunkt t eine zusätzliche äußere Kraft $F(t)$ auf die Masse wirkt (also zusätzliche zur Rückstellkraft des Pendels). Somit erhält man aus dem zweiten Newtonschen Axiom eine andere Bewegungsgleichung; sie lautet nun

$$\ddot{x}(t) = -\frac{D}{m}x(t) + \frac{F(t)}{m}$$

für $t \geq 0$. Auch diese Gleichung kann man wieder in eine Differentialgleichung erster Ordnung umschreiben. Wegen $\ddot{x}(t) = \dot{v}(t)$ lautet Sie nun aber

$$\begin{pmatrix}\dot{x}(t)\\ \dot{v}(t)\end{pmatrix} = \begin{pmatrix}0 & 1\\ -\frac{D}{m} & 0\end{pmatrix}\begin{pmatrix}x(t)\\ v(t)\end{pmatrix} + \underbrace{\begin{pmatrix}0\\ \frac{F(t)}{m}\end{pmatrix}}_{=:f(t)} \qquad \text{für alle } t \in [0,\infty).$$

Um zunächst eine Intuition dafür zu erhalten, wie man linearen Differentialgleichungen mit Inhomogenitäten lösen kann, sehen wir uns zunächst ein Beispiel einer ähnlichen, aber einfacheren Gleichung an – nämlich einer linearen **Differenzengleichung**:

Beispiel 4.4.2 (Kontostand bei gleichbleibender Verzinsung und jährlicher Einzahlung)**.** Betrachten Sie ein Konto mit anfänglichem Kontostand $x_0 \in \mathbb{R}$. Am Ende des n-ten Jahres wird der Kontostand x_n vom Anfang des n-ten Jahres mit einem Zinssatz q verzinst und weiteres Guthaben eingezahlt, in Höhe von b_n. Somit beträgt der Kontostand am Anfang von Jahr $n+1$

$$x_{n+1} = x_n + qx_n + b_n = (1+q)x_n + b_n$$

für jedes $n \in \mathbb{N}_0$. Solche eine Gleichung bezeichnet man als **Differenzengleichung**, weil die Differenz $x_{n+1} - x_n$ zwischen den Werten zu den Zeitpunkten $n+1$ und n durch den Wert zum Zeitpunkt n ausgedrückt wird.

Wenn man die Werte $x_0, x_1, x_2, \ldots$ sukzessive ausrechnet, erhält man

$$\begin{aligned}
x_1 &= (1+q)x_0 + b_0;\\
x_2 &= (1+q)^2 x_0 + (1+q)b_0 + b_1;\\
x_3 &= (1+q)^3 x_0 + (1+q)^2 b_0 + (1+q)b_1 + b_2;\\
x_4 &= (1+q)^4 x_0 + (1+q)^3 b_0 + (1+q)^2 b_1 + (1+q)b_2 + b_3;\\
&\vdots
\end{aligned}$$

Allgemein gilt für alle $n \in \mathbb{N}_0$ die Formel

$$x_{n+1} = (1+q)^{n+1} x_0 + \sum_{k=0}^{n} (1+q)^{n-k} b_k,$$

wie man leicht per Induktion zeigt.

Vorlesung 30 (Fr, 06.02.) beginnt mit Bemerkung 4.4.3.

Bemerkung 4.4.3 (Faltung von Folgen)**.** Für zwei Folgen $y = (y_n)_{n\in\mathbb{N}_0}$ und $z = (z_n)_{n\in\mathbb{N}_0}$ in $\mathbb{R}$ ist die **Faltung** definiert als die Folge $y \star z$ mit Indexmenge $\mathbb{N}_0$, deren n-ter Eintrage für jedes $n \in \mathbb{N}_0$ gegeben ist als

$$(y \star z)_n := \sum_{k=0}^{n} y_{n-k} z_k;$$

Das kennen Sie bereits von Aufgabe 9.2 auf Präsenzblatt 9.

Betrachten wir nun die inhomogene lineare Differenzengleichung

$$x_{n+1} = (1+q)x_n + b_n \qquad \text{für alle } n \in \mathbb{N}_0$$

aus Beispiel 4.4.2 sowie die zugehörige homogene Differenzengleichung

$$x_{n+1} = (1+q)x_n \qquad \text{für alle } n \in \mathbb{N}_0.$$

Die Lösung der homogenen Gleichung ist offensichtlich die Folge $\left((1+q)^n x_0\right)_{n\in\mathbb{N}_0} = \left((1+q)^n\right)_{n\in\mathbb{N}_0} x_0$ für alle $n \in \mathbb{N}_0$. Damit kann man die Lösungsformel der inhomogenen Gleichung aus Beispiel 4.4.2 folgendermaßen interpretieren: Man nimmt die Lösung der homogenen Gleichung und addiert (bis auf eine Verschiebung um den Index 1) die Faltung der Folge $\left((1+q)^n\right)_{n\in\mathbb{N}_0}$ mit der Inhomogenität $(b_n)_{n\in\mathbb{N}_0}$.

Zum Lösen linearer Differentialgleichungen mit Inhomogenität gehen wir nun ganz ähnlich vor. Wir definieren dafür die Faltung von Funktionen, die auf $[0,\infty)$ definiert sind.

Definition 4.4.4 (Faltung)**.** (a) Seien $g:[0,\infty)\to\mathbb{C}$ und $f:[0,\infty)\to\mathbb{C}$ stetig. Man nennt die Abbildung $g\star f:[0,\infty)\to\mathbb{C}$, die durch

$$(g\star f)(t):=\int_0^t g(t-s)f(s)\,\mathrm{d}s=\int_0^t g(s)f(t-s)\,\mathrm{d}s$$

für alle $t\in[0,\infty)$ gegeben ist, die **Faltung** von g und f.

(b) Allgemeiner seien nun $G:[0,\infty)\to\mathbb{C}^{d\times d}$ und $f:[0,\infty)\to\mathbb{C}^d$ stetig. Man nennt die Abbildung $G\star f:[0,\infty)\to\mathbb{C}^d$, die durch

$$(G\star f)(t):=\int_0^t G(t-s)f(s)\,\mathrm{d}s=\int_0^t G(s)f(t-s)\,\mathrm{d}s$$

für alle $t\in[0,\infty)$ gegeben ist, die **Faltung** von G und f.

Beispiel 4.4.5 (Faltung von zwei Schwingungen)**.** Seien $\omega,\tilde\omega\in(0,\infty)$ und seien $f,g:[0,\infty)\to\mathbb{R}$, $f(t)=\sin(\omega t)$ und $g(t)=\sin(\tilde\omega t)$ für alle $t\in[0,\infty)$. Dann gilt

$$(f\star g)(t)=\begin{cases}\frac{\omega\sin(\tilde\omega t)-\tilde\omega\sin(\omega t)}{\omega^2-\tilde\omega^2} & \text{falls } \tilde\omega\neq\omega,\\ \frac{-t\omega\cos(\omega t)+\sin(\omega t)}{2\omega} & \text{falls } \tilde\omega=\omega.\end{cases}$$

Beweis. Sei $t\in[0,\infty)$ fest.

Erster Fall: $\tilde\omega\neq\omega$. Es gilt

$$\begin{aligned}(f\star g)(t)&=\int_0^t\sin(\omega(t-s))\sin(\tilde\omega s)\,\mathrm{d}s\\&=\frac{1}{\omega}\cos(\omega(t-s))\sin(\tilde\omega s)|_{s=0}^{s=t}-\frac{\tilde\omega}{\omega}\int_0^t\cos(\omega(t-s))\cos(\tilde\omega s)\,\mathrm{d}s\\&=\frac{\sin(\tilde\omega t)}{\omega}+\frac{\tilde\omega}{\omega^2}\sin(\omega(t-s))\cos(\tilde\omega s)|_{s=0}^{s=t}+\frac{\tilde\omega^2}{\omega^2}\int_0^t\sin(\omega(t-s))\sin(\tilde\omega s)\,\mathrm{d}s\\&=\frac{\sin(\tilde\omega t)}{\omega}-\frac{\tilde\omega\sin(\omega t)}{\omega^2}+\frac{\tilde\omega^2}{\omega^2}(f\star g)(t).\end{aligned}$$

Wegen $\omega\neq\tilde\omega$ und $\omega,\tilde\omega>0$ ist $\frac{\tilde\omega^2}{\omega^2}\neq1$. Somit kann man die soeben erhaltene Gleichung nach $(f\star g)(t)$ auflösen und erhält

$$(f\star g)(t)=\frac{1}{1-\frac{\tilde\omega^2}{\omega^2}}\Big(\frac{\sin(\tilde\omega t)}{\omega}-\frac{\tilde\omega\sin(\omega t)}{\omega^2}\Big)=\frac{\omega\sin(\tilde\omega t)-\tilde\omega\sin(\omega t)}{\omega^2-\tilde\omega^2}$$

Zweiter Fall: $\tilde{\omega} = \omega$. Lassen Sie uns eine Folge (ω_n) in $(0,\infty) \setminus \{\omega\}$ betrachten, die gegen ω konvergiert. Dann gilt

$$\begin{aligned}(f \star g)(t) &= \int_0^t \sin(\omega(t-s))\sin(\omega s)\,\mathrm{d}s \\ &= \lim_{n\to\infty} \int_0^t \sin(\omega(t-s))\sin(\omega_n s)\,\mathrm{d}s \\ &= \lim_{n\to\infty} \frac{\omega\sin(\omega_n t) - \omega_n \sin(\omega t)}{(\omega-\omega_n)(\omega+\omega_n)}\end{aligned}$$

wobei die zweite Gleichheit aus dem Satz von der dominierten Konvergenz (Theorem 2.3.8) folgt;[38] die dritte Gleichheit folgt aus dem ersten Fall, den wir bereits behandelt haben. Den Grenzwert am Ende kann man zum Beispiel mit Hilfe der Taylor-Entwicklung des Sinus oder mit Hilfe der Regel von l'Hôpital (siehe Analysis 1,) bestimmen und erhält damit die behauptete Formel. □

Proposition 4.4.6 (Eigenschaften der Faltung)**.** *Sei $G : [0,\infty) \to \mathbb{C}^{d\times d}$ stetig differenzierbar und $f : [0,\infty) \to \mathbb{C}^d$ stetig. Dann ist auch $G \star f : [0,\infty) \to \mathbb{C}^d$ stetig differenzierbar mit der Ableitung*

$$\frac{\mathrm{d}}{\mathrm{d}t}(G \star f)(t) = (\dot{G} \star f)(t) + G(0)f(t) \qquad \text{für alle } t \in [0,\infty).$$

wobei das $\dot{G}$ auf der rechten Seite wie üblich die Zeitableitung von G bezeichnet.

Beweis. Mit Hilfe der Stetigkeit von $\dot{G}$ kann man zeigen, dass die Faltung $\dot{G} \star f$ ebenfalls eine stetige Funktion von $[0,\infty)$ nach $\mathbb{C}^d$ ist; wir verzichten an dieser Stelle auf die Details.

Nun betrachten wir die Funktion $h : [0,\infty) \to \mathbb{C}^d$, $h(t) = \int_0^t (\dot{G} \star f)(r) + G(0)f(r)\,\mathrm{d}r$ für alle $t \in [0,\infty)$. Wegen des Hauptsatzes der Differential- und Integralrechnung ist h stetig differenzierbar mit Ableitung $\dot{h}(t) = (\dot{G} \star f)(t) + G(0)f(t)$ für alle $t \in [0,\infty)$. Somit genügt es zu zeigen, dass $h = G \star f$ ist. Sei also $t \in [0,\infty)$ fest. Dann gilt

$$\begin{aligned}\int_0^t \int_0^r \dot{G}(r-s)f(s)\,\mathrm{d}s\,\mathrm{d}r &= \int_0^t \int_s^t \dot{G}(r-s)f(s)\,\mathrm{d}r\,\mathrm{d}s = \int_0^t \int_0^{t-s} \dot{G}(r)\,\mathrm{d}r\,f(s)\,\mathrm{d}s \\ &= \int_0^t \big(G(t-s) - G(0)\big)f(s)\,\mathrm{d}s = (G \star f)(t) - G(0)\int_0^t f(s)\,\mathrm{d}s,\end{aligned}$$

wobei wir für die erste Gleichheit den Satz von Fubini verwendet haben. Indem wir den Term $G(0)\int_0^t f(s)\,\mathrm{d}s$ auf die linke Seite bringen, erhalten wir wie behauptet $h(t) = (G \star f)(t)$. □

Korollar 4.4.7 (Lineare Differentialgleichungen mit Inhomogenität)**.** *Sei $A \in \mathbb{C}^{d\times d}$ und $x_0 \in \mathbb{C}^d$. Außerdem sei $f : [0,\infty) \to \mathbb{C}^d$ stetig. Dann ist das Anfangswertproblem*

$$\begin{cases}\dot{x}(t) = Ax(t) + f(t) & \text{für alle } t \in [0,\infty), \\ x(0) = x_0\end{cases}$$

[38] Können Sie hier die Details ausführen?

global eindeutig lösbar und die globale Lösung ist gegeben durch

$$x(t) = e^{tA}x_0 + \big(e^{\cdot A} \star f\big)(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A} f(s)\,\mathrm{d}s$$

für alle $t \in [0,\infty)$.

Beweis. Wir betrachten die Abbildung $h : [0,\infty) \times \mathbb{C}^d \to \mathbb{C}^d$, $h(t,z) = Az + f(t)$ für alle $t \in [0,\infty)$, die die rechte Seite der Differentialgleichung beschreibt. Für alle $t \in [0,\infty)$ und alle $y, z \in \mathbb{C}^d$ gilt

$$\big\|h(t,z) - h(t,y)\big\|_2 = \big\|A(z-y)\big\|_2 \le \|A\|_{2\leftarrow 2}\,\big\|z-y\big\|_2 .$$

Somit ist die globale Lipschitzbedingung (4.1.2) aus dem globalen Existenz- und Eindeutigkeitssatz (Theorem 4.1.3) erfüllt (mit $I = [0,\infty)$). Also ist das Anfangswertproblem global eindeutig lösbar.

Dass die gegebene Funktion x tatsächlich die Differentialgleichung erfüllt, folgt aus Proposition 4.4.6: Laut dieser Proposition ist x stetig differenzierbar mit

$$\dot{x}(t) = Ae^{tA}x_0 + \big(Ae^{\cdot A} \star f\big)(t) + e^{0\cdot A}f(t) = A\big(e^{tA}x_0 + (e^{\cdot A} \star f)(t)\big) + f(t) = Ax(t) + f(t)$$

für alle $t \in [0,\infty)$. Außerdem erfüllt x die Anfangsbedingung, denn es ist $x(0) = e^{0\cdot A}x_0 + \big(e^{\cdot A} \star f\big)(0) = x_0$. □

Beispiel 4.4.8 (Resonanzkatastrophe für ein Federpendel)**.** Wir betrachten noch einmal das Federpendel mit externer Kraft aus Beispiel 4.4.1; seine Bewegung wird durch die Differentialgleichung

$$\begin{pmatrix}\dot{x}(t)\\ \dot{v}(t)\end{pmatrix} = \begin{pmatrix}0 & 1\\ -\frac{D}{m} & 0\end{pmatrix}\begin{pmatrix}x(t)\\ v(t)\end{pmatrix} + \underbrace{\begin{pmatrix}0\\ \frac{F(t)}{m}\end{pmatrix}}_{=:f(t)} \qquad \text{für alle } t \in [0,\infty).$$

beschrieben. Lassen Sie uns der Einfachkeit halber annehmen, dass das Pendel zum Zeitpunkt 0 im Gleichgewicht ruht, das heißt, das $x(0) = 0$ und $v(0) = 0$ gilt.

Wir verwenden die Abkürzung $\omega := \sqrt{D/m}$. Dann gilt alle $t \in [0,\infty)$ laut Aufgabe 15.1(e) auf Präsenzblatt 15

$$\exp\left(t\begin{pmatrix}0 & 1\\ -\frac{D}{m} & 0\end{pmatrix}\right) = \begin{pmatrix}\cos(\omega t) & \frac{1}{\omega}\sin(\omega t)\\ -\omega\sin(\omega t) & \cos(\omega t)\end{pmatrix}$$

Nun betrachten wir die Situation, in der die externe Kraft F mit ebenfalls oszilliert: Konkret sei $F(t) = F_0\sin(\hat{\omega}t)$ ein $\hat{\omega} \in (0,\infty)$ und alle $t \in [0,\infty)$. Die Lösung unseres Anfangswertproblems ist somit laut Korollar 4.4.7 gegeben durch

$$\begin{aligned}\begin{pmatrix}x(t)\\ v(t)\end{pmatrix} &= \int_0^t \begin{pmatrix}\cos(\omega(t-s)) & \frac{1}{\omega}\sin(\omega(t-s))\\ -\omega\sin(\omega(t-s)) & \cos(\omega(t-s))\end{pmatrix}\begin{pmatrix}0\\ \frac{F_0}{m}\sin(\hat{\omega}s)\end{pmatrix}\mathrm{d}s\\ &= \begin{pmatrix}\int_0^t \frac{F_0}{\omega m}\sin(\omega(t-s))\sin(\hat{\omega}s)\,\mathrm{d}s\\ \int_0^t \frac{F_0}{m}\cos(\omega(t-s))\sin(\hat{\omega}s)\,\mathrm{d}s\end{pmatrix}.\end{aligned}$$

Die Formel für $x(t)$ ist das $\frac{F_0}{\omega m}$-fache einer Faltung zweier Sinus-Funktionen mit den Winkelgeschwindigkeiten ω und $\hat{\omega}$. In Beispiel 4.4.5 sieht man nun: Für $\hat{\omega} \neq \omega$ bleibt die Lösung für $t \to \infty$ beschränkt. Für $\hat{\omega} = \omega$ ist $x(t)$ hingegen unbeschränlt für $t \to \infty$ (wegen des Vorfaktors t, der vor dem Cosinus im zweiten Fall auftritt).

Wenn man das Pendel mit einer externen Kraft anregt, die mit derselben Frequenz schwingt wie das Pendel selbst (das heißt mit der sogenannte **Eigenfrequenz** des Pendels), schwingt das Pandel also im Laufe der Zeit unbeschränkt weit aus.[39] Das bezeichnet man als **Resonanzkatastrophe**.

[39] Das ist natürlich – wieder einmal – eine grobe Vereinfachung. In Wirklichkeit wird das Pendel nicht immer weiter ausschwingen, sondern der Apparat irgendwann zerstört werden, das heißt, die Differentialgleichung ist ab einem bestimmten Punkt keine gute Approximation an das reale Verhalten des Pendels. Außerdem kann man als $\hat{\omega}$ natürlich nie ganz genau die Eigenfrequenz des Pendels erwischen – aber wenn man nahe genuged dran ist, werden die Schwingungen halt schon sehr groß. Denken Sie übrigens auch daran, dass in der Realität auch Reibung eine wichtige Rolle spielt und das Pendel ausbremst; das haben wir in diesem sehr einfachen Modell völlig ignoriert.

Literaturverzeichnis

[1] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Heidelberg: Springer Spektrum, 8th enlarged and updated edition edition, 2018.

[2] Otto Forster. *Analysis 3. Maß- und Integrationstheorie, Integralsätze im $\mathbb{R}^n$ und Anwendungen*. Wiesbaden: Springer Spektrum, 8th revised edition edition, 2017.

[3] Jochen Glück. Gewöhnliche Differentialgleichungen. Vorlesungsmanuskript, Universität Passau, Sommersemester 2020.

[4] Jochen Glück. Analysis 2. Vorlesungsmanuskript, Bergische Universität Wuppertal, Sommersemester 2025.

[5] Jan W. Prüss and Mathias Wilke. *Gewöhnliche Differentialgleichungen und dynamische Systeme*. Grundstud. Math. Birkhäuser, 2010.

[6] Karina Schumacher. Analytische Betrachtung von Klimamodellen. Bachelorthesis, Bergische Universität Wuppertal, 2026, in Arbeit.

[7] Uwe Storch and Hartmut Wiebe. *Lehrbuch der Mathematik. Für Mathematiker, Informatiker und Physiker. Band III: Analysis mehrerer Veränderlicher - Integrationstheorie*. Mannheim: B.I. Wissenschaftsverlag, 1993.