Reinhard Höpfner

Vorlesung Stochastik I

Kapitel I: Grundbegriffe

Sommersemester 2016

Institut für Mathematik, Johannes Gutenberg Universität Mainz

May 11, 2016

Übersicht zu Kapitel I :

A. Einführung

Modellbildung in der Stochastik

Corde de Bertrand

B. σ-Algebren, Dynkin-Systeme, Masse

Definition σ-Algebra und Beispiele 1.1–1.4

Erzeuger und Beispiele (insbesondere für $\mathcal{B}(\mathbb{R}^d)$) 1.5–1.6'

Dynkin-Systeme 1.7–1.11

Masse 1.12

Eindeutigkeitssatz 1.13

Beispiele (Verteilungsfunktionen auf $\mathbb{R}$ und $\mathbb{R}^d$) 1.14–1.14'

C. Algebren, Inhalte, Prämasse, Konstruktion von Massen

Algebren und Prämasse 1.15–1.17'

Formulierung des Fortsetzungssatzes 1.18

Beispiele (Wahrscheinlichkeitsmasse auf $\mathbb{R}$ und $\mathbb{R}^d$) 1.19–1.20

Lebesgue-Mass auf $\mathbb{R}^d$ 1.21

Diracmass, Zählmass 1.21'

D. Beweis des Fortsetzungssatzes

äusseres Mass 1.22ff

Beweis des Fortsetzungssatzes 1.30

eindeutige Fortsetzung σ-endlicher Prämasse 1.31

E. Messbare Abbildungen

Definition, Grundeigenschaften, Beispiele 1.32–1.36

Transport von Massen, Bildmasse 1.37

messbare Abbildungen mit Werten in $\mathbb{R}$, $\overline{\mathbb{R}}$, $\mathbb{R}^d$ 1.38–1.42

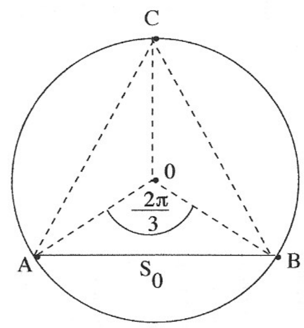
Faktorisierungslemma 1.43

A. Einführung

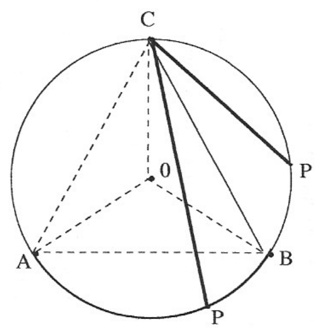
In der Alltagssprache steht das Wort 'Zufall' in Verbindung mit Ereignissen, von denen man noch nicht weiss, ob sie eintreten werden oder nicht. Die Anfänge der Stochastik als mathematischer Wissenschaft –im Frankreich des 17. Jahrhunderts– bestehen in dem Versuch, rationale Aussagen über den Ausgang von Glücksspielen zu machen: man wollte gewissen Spielausgängen 'Wahrscheinlichkeiten' zuordenen können. Dabei ist jede Quantifizierung von 'Wahrscheinlichkeit' auf klare Begriffe angewiesen, sonst führen verschiedene unausgesprochene Modellvorstellungen zu verschiedenen Antworten auf dieselbe Frage, wobei jede dieser Antworten 'in sich stimmig' ist. Dies nennen manche die 'Paradoxe der Stochastik'. Es gibt jedoch keine solchen 'Paradoxe', sondern nur nicht hinreichend klar formulierte Begriffe. Das folgende Beispiel ist berühmt.

1.0 Beispiel: (Bertrand 1907) Man betrachte einen Kreis mit Radius r, darin ein gleichseitiges Dreieck ABC, und schreibe S_0 für die Seite $\overline{AB}$.

Die Frage lautet: Wie wahrscheinlich ist es, dass die Länge $\ell = \ell(S)$ einer 'zufällig über den Kreis geworfenen Sehne S' grösser ist als $\ell(S_0)$?



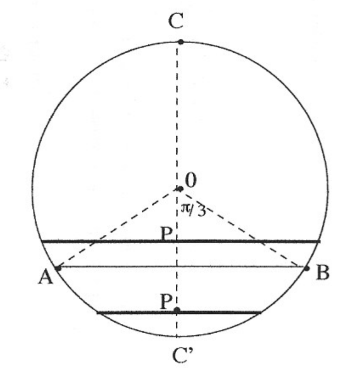
Ansatz 1: Wähle einen Punkt P 'zufällig auf dem Kreisbogen' und konstruiere S als $\overline{CP}$. Genau die Punkte P, die in das Kreisbogensegment zwischen A und B fallen, werden hierbei zu einer Sehnenlänge $\ell(\overline{CP}) > \ell(S_0)$ führen. Wenn 'Wahrscheinlichkeit' als Verhältnis zwischen einer 'Gesamtheit günstiger Möglichkeiten' zu einer 'Gesamtheit aller Möglichkeiten' aufgefasst wird, kommt man mit Ansatz 1 zur Antwort



$$\text{WS} = \frac{\ell(\text{Kreisbogensegment zwischen } A \text{ und } B)}{\text{Kreisumfang}} = \frac{\frac{1}{3}(2\pi r)}{2\pi r} = \frac{1}{3}\,.$$

Ansatz 2: Man setze den Radius $\overline{OC}$ fort zu einem Durchmesser $\overline{CC'}$, wähle einen Punkt P 'zufällig auf $\overline{OC'}$ ', und konstruiere die Sehne S als Senkrechte auf $\overline{OC'}$ im Punkte P.

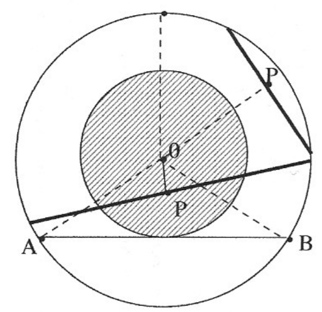
Jetzt führen genau die Punkte P, die zwischen O und dem Schnittpunkt von $\overline{OC'}$ mit $\overline{AB}$ liegen, zu einer Sehnenlänge $\ell(S) > \ell(S_0)$. Dies führt auf einen Längenvergleich



$$\mathrm{WS} \;=\; \frac{r\cdot\cos\frac{\pi}{3}}{r} \;=\; \cos\frac{\pi}{3} \;=\; \frac{1}{2}\,.$$

Ansatz 3: Wähle einen Punkt P 'zufällig auf der Kreisfläche', und konstruiere S als Senkrechte auf $\overline{OP}$ in P.

Dabei führen alle Punkte P im Inkreis des Dreiecks ABC zu einer Sehnenlänge $\ell(S) > \ell(S_0)$. Durch Flächenvergleich zwischen Inkreis und Umkreis erhält man



$$\mathrm{WS} \;=\; \frac{\pi(r\cos\frac{\pi}{3})^2}{\pi r^2} \;=\; \frac{1}{4}\,.$$

Alle 3 Lösungen des Problems sind in sich richtig, beruhen aber auf grundsätzlich verschiedenen Modellvorstellungen, was 'zufällig' bedeuten könne, im Sinne einer Gleichverteilung einmal auf dem Kreisbogen (Ansatz 1), dann auf der Geraden $\overline{OC'}$ (Ansatz 2), und schliesslich auf der Kreisfläche (Ansatz 3). □

Es ist also nötig, saubere Begriffe zu entwickeln, bevor man darangeht, Wahrscheinlichkeiten berechnen zu wollen. Seit Kolmogorov ($\approx$ 1930) legt man zuerst einen Raum Ω fest

$$\Omega = \text{Gesamtheit der möglichen } \textit{Realisierungen } \;\omega\;,$$

reichhaltig genug, um mögliche Ausgänge einer zu machenden Beobachtung durch Elemente $\omega\in\Omega$ kodieren zu können. Dann zeichnet man eine Klasse $\mathcal{A}$ interessanter Teilmengen von Ω aus

$$\text{Teilmengen } A\subset\Omega \text{ mit } A\in\mathcal{A} \text{ heissen } \textit{Ereignis}$$

wobei Elemente $\omega \in A$ Realisierung des Ereignisses A genannt werden. Die mathematischen Eigenschaften der Klasse $\mathcal{A}$, die man braucht, um Ereignissen eine Wahrscheinlichkeit zusprechen zu können, werden die einer σ-*Algebra* sein. Schliesslich erklärt man 'Wahrscheinlichkeit' als Mengenfunktion

$$P \,:\, \mathcal{A} \ni A \longrightarrow P(A) \in [0,1]$$

auf der σ-Algebra $\mathcal{A}$, deren wesentliche Eigenschaft die σ-*Additivität* sein wird. Das Tripel

$$(\,\Omega\,,\,\mathcal{A}\,,\,P\,)$$

wird damit ein Begriffssystem liefern, mit dem Mathematik 'Zufall' analysieren und zum Beispiel die Wahrscheinlichkeit komplizierterer Ereignisse –ausgehend von einer geeigneten Vorgabe auf einer kleinen Klasse sehr einfach strukturierter Ereignisse– berechnen kann.

B. σ-Algebren, Dynkin-Systeme, Masse

1.1 Definition: Ein System $\mathcal{A}$ von Teilmengen einer Menge Ω heisst *σ-Algebra* falls gilt

i) $\Omega \in \mathcal{A}$

ii) $A \in \mathcal{A} \implies A^c = \Omega \setminus A \in \mathcal{A}$

iii) $(A_n)_{n\geq 1} \subset \mathcal{A} \implies \bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$.

Ist $\mathcal{A}$ eine σ-Algebra auf Ω, nennt man das Paar $(\Omega, \mathcal{A})$ einen *messbaren Raum.*

Wir setzen stets voraus, dass die Menge Ω nichtleer ist, ohne dass dies eigens erwähnt würde.

1.2 Folgerung: Ist $\mathcal{A}$ eine σ-Algebra auf Ω, so gilt auch:

i) $\emptyset \in \mathcal{A}$, wegen $\emptyset = \Omega \backslash \Omega$ und 1.1 i) und ii);

ii) $(A_n)_{n\geq 1} \subset \mathcal{A} \implies \bigcap\limits_{n\geq 1} A_n \in \mathcal{A}$, wegen $\bigcap\limits_{n\geq 1} A_n \;=\; \Omega \setminus \left(\bigcup\limits_{n\geq 1} A_n^c \right)$ und 1.1 ii) und iii);

iii) $A_1, \ldots, A_n \in \mathcal{A} \implies A_1 \cap \ldots \cap A_n \in \mathcal{A}$ und $A_1 \cup \ldots \cup A_n \in \mathcal{A}$.

Dies gilt wegen 1.1 iii) und 1.2 iii) (schreibe $A_1 \cap \ldots \cap A_n \;=\; A_1 \cap \ldots \cap A_n \cap \Omega \cap \Omega \cap \ldots$ und $A_1 \cup \ldots \cup A_n = A_1 \cup \ldots \cup A_n \cup \emptyset \cup \emptyset \cup \ldots$).

iv) $A_1, A_2 \in \mathcal{A} \implies A_1 \backslash A_2 \in \mathcal{A}$, wegen $A_1 \backslash A_2 = A_1 \cap A_2^c$. $\square$

1.3 Bemerkungen: a) Auf jeder Menge Ω hat man zwei triviale σ-Algebren, nämlich

$$\{\emptyset, \Omega\}$$

(diese ist i.a. uninteressant) sowie die Potenzmenge

$$\mathcal{P}(\Omega) = \{A : A \subset \Omega\}$$

(diese ist aber i.a. viel zu gross, als dass wir auf ihr Wahrscheinlichkeiten erklären könnten).

b) Sind $\mathcal{A}_i$, $i \in I$, σ-Algebren auf derselben Menge Ω, so ist

$$\mathcal{A} \;:=\; \bigcap_{i \in I} \mathcal{A}_i \;:=\; \{A \subset \Omega \;:\; A \in \mathcal{A}_i \;\; \forall\, i \in I\}$$

eine σ-Algebra auf Ω. Dabei darf I eine beliebige Indexmenge sein.

c) Die Vereinigung von σ-Algebren ist i.a. keine σ-Algebra mehr.

d) Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, sei $\Omega' \subset \Omega$. Dann ist die *Spur von $\mathcal{A}$ auf Ω'*

$$\mathcal{A}\,|_{\Omega'} := \{A \cap \Omega' \ : \ A \in \mathcal{A}\}$$

eine σ-Algebra auf Ω'. Eine beliebige Teilmenge Ω' von Ω wird – versehen mit der Spur von $\mathcal{A}$ auf Ω'– selbst zu einem messbaren Raum.

Beweis: Die Aussagen a) und d) sind einfach. Für b) verifizieren wir die Eigenschaften 1.1 einer σ-Algebra:
Da jedes $\mathcal{A}_i$ eine σ-Algebra ist, hat man nach 1.1 i) $\Omega \in \mathcal{A}_i$ für jedes $i \in I$, damit $\Omega \in \mathcal{A}$.
Ist $A \in \mathcal{A}$, so ist $A \in \mathcal{A}_i$ für jedes $i \in I$, also nach 1.1 ii) $A^c = \Omega \setminus A \in \mathcal{A}_i$ für jedes $i \in I$; damit gilt $A^c \in \mathcal{A}$.
Seien $A_n, n \geq 1$ in $\mathcal{A}$. Fixiere $i \in I$. Dann $(A_n)_n \subset \mathcal{A}_i$; da $\mathcal{A}_i$ eine σ-Algebra ist, gilt $\bigcup\limits_{n\geq 1} A_n \in \mathcal{A}_i$ nach 1.1 iii). Dabei ist $i \in I$ beliebig, also ist gezeigt $\bigcup\limits_{n\geq 1} A_n \in \mathcal{A}$.
Zu c): das letzte Argument hat keine Entsprechung mehr, wenn man Vereinigungen von σ-Algebren betrachtet (wir werden in 1.5 unten sehen, wie man das *Mengensystem* $\bigcup\limits_{i\in I} \mathcal{A}_i$ zu einer σ-Algebra *ergänzen* kann). $\square$

1.4 Satz: Seien Ω_1, Ω_2 beliebige Mengen, sei $T : \Omega_1 \to \Omega_2$ eine Abbildung. Für $A_2 \subset \Omega_2$ heisst

$$T^{-1}(A_2) := \{\omega_1 \in \Omega_1 \ : \ T(\omega_1) \in A_2\}$$

Urbild der Menge A_2 unter T. Ist $\mathcal{A}_2$ eine σ-Algebra auf Ω_2, so ist das *Urbild von $\mathcal{A}_2$ unter T*

$$T^{-1}(\mathcal{A}_2) := \{T^{-1}(A_2) \ : \ A_2 \in \mathcal{A}_2\}$$

eine σ-Algebra auf Ω_1.

Beweis: Wir zeigen, dass das Mengensystem $\mathcal{A}_1 := T^{-1}(\mathcal{A}_2)$ eine σ-Algebra auf Ω_1 ist.
Klar gilt $\Omega_1 = T^{-1}(\Omega_2) \in \mathcal{A}_1$. Ist $A_1 \in \mathcal{A}_1$, so existiert ein $A_2 \in \mathcal{A}_2$ mit $A_1 = T^{-1}(A_2)$, also

$$\Omega_1 \setminus A_1 \ = \ \Omega_1 \setminus T^{-1}(A_2) \ = \ T^{-1}(\underbrace{\Omega_2 \setminus A_2}_{\in \mathcal{A}_2}) \ \in \mathcal{A}_1 \ :$$

damit ist $\mathcal{A}_1$ stabil bezüglich Komplementbildung. Ist nun $(A_{1,n})_{n\geq 1} \subset \mathcal{A}_1$, so existiert eine Mengenfolge $(A_{2,n})_n \subset \mathcal{A}_2$ so dass $A_{1,n} = T^{-1}(A_{2,n})$, und es gilt

$$\bigcup_{n\geq 1} A_{1,n} \ = \ T^{-1}(\underbrace{\bigcup_{n\geq 1} A_{2,n}}_{\in \mathcal{A}_2}) \ \in \mathcal{A}_1 \ .$$

Damit sind die definierenden Eigenschaften 1.1 einer σ-Algebra verifiziert. $\square$

Oft sucht man zu einer vorgegebenen Klasse $\mathcal{C}$ von Teilmengen einer Menge Ω eine (möglichst kleine) σ-Algebra, die alle Mengen aus $\mathcal{C}$ enthält.

1.5 Satz: Sei $\mathcal{C}$ ein beliebiges System von Teilmengen einer Menge Ω. Stets existiert eine kleinste σ-Algebra auf Ω, welche $\mathcal{C}$ umfasst: für diese schreibt man $\sigma(\mathcal{C})$.
Man nennt $\sigma(\mathcal{C})$ die *von $\mathcal{C}$ erzeugte σ-Algebra*.

Beweis: Schreibe $\mathcal{J}$ für die Klasse aller σ-Algebren $\mathcal{A}$ auf Ω mit $\mathcal{C} \subset \mathcal{A}$. Nach 1.3 b) ist $\mathcal{A}_0 := \bigcap_{\mathcal{J}} \mathcal{A}$ eine σ-Algebra auf Ω; klar gilt $\mathcal{C} \subset \mathcal{A}_0$. Per Definition ist $\bigcap_{\mathcal{J}} \mathcal{A}$ die kleinste σ-Algebra auf Ω, welche $\mathcal{C}$ umfasst (denn: $\bigcap_{\mathcal{J}} \mathcal{A}$ is enhalten in jeder σ-Algebra, welche $\mathcal{C}$ unfasst). Die Aussage des Satzes gilt also mit $\sigma(\mathcal{C}) := \mathcal{A}_0$. $\square$

1.5' Definition: Sei $\mathcal{A}$ eine σ-Algebra von Teilmengen einer Menge Ω. Jedes Teilsystem $\mathcal{E} \subset \mathcal{A}$ mit der Eigenschaft $\sigma(\mathcal{E}) = \mathcal{A}$ nennt man einen *Erzeuger von $\mathcal{A}$*.

Für dieselbe σ-Algebra $\mathcal{A}$ kann man i.a. durchaus verschiedene Systeme $\mathcal{E}, \mathcal{E}', ...$ angeben, die die Eigenschaft $\sigma(\mathcal{E}) = \sigma(\mathcal{E}') = ... = \mathcal{A}$ besitzen; je nach Zweck wird man mit je anderen Erzeugern argumentieren.

1.5" Bemerkung: Ein System $\mathcal{O}$ von Teilmengen einer Menge E heisst Topologie, wenn gilt:

i) $\emptyset \in \mathcal{O}$, $E \in \mathcal{O}$;

ii) $E_1, \ldots, E_n \in \mathcal{O} \Rightarrow E_1 \cap \ldots \cap E_n \in \mathcal{O}$;

iii) für eine beliebige Indexmenge I und Mengen $E_i \in \mathcal{O}$, $i \in I$, gilt: $\bigcup_{i \in I} E_i \in \mathcal{O}$.

Die Elemente von $\mathcal{O}$ nennt man offen in E, und das Paar $(E, \mathcal{O})$ einen topologischen Raum.

Ein wichtiger Spezialfall: in einem metrischen Raum (E, d) bildet das System $\mathcal{O}$ aller im Sinne der Metrik d offenen Teilmengen von E eine Topologie.

1.6 Definition: Sei $\mathcal{O}$ eine Topologie auf E. Die von $\mathcal{O}$ erzeugte σ-Algebra von Teilmengen von E heisst *Borel-σ-Algebra*. Die übliche Schreibweise ist $\mathcal{B}(E)$ anstelle von $\sigma(\mathcal{O})$.

Wir diskutieren ausführlich das Beispiel des Euklidischen Raumes.

1.6' Beispiel: Betrachte $\Omega := \mathbb{R}^d$ mit der euklidischen Metrik $d(\cdot,\cdot)$

$$d(x,y) = \left(\sum_{i=1}^{d}(x_i - y_i)^2\right)^{1/2}, \quad x = (x_1,\dots,x_d),\ y = (y_1,\dots,y_d)\ .$$

Bezeichne $\mathcal{O}$ das System aller bezüglich d offenen Mengen in $\mathbb{R}^d$. Die kleinste σ-Algebra, welche alle offenen Teilmengen von $\mathbb{R}^d$ enthält, ist die Borelsche σ-Algebra $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{O})$. Die offene Kugel um $x \in \mathbb{R}^d$ mit Radius $\varepsilon > 0$ bezeichnene wir mit $B_\varepsilon(x) := \{y \in \mathbb{R}^d : d(y,x) < \varepsilon\}$.

a) Eine wichtige Eigenschaft des Euklidischen Raumes ist die *Separabilität*, d.h. die Existenz einer abzählbar dichten Teilmenge im Sinne der Metrik $d(\cdot,\cdot)$. In $\mathbb{R}^d$ ist eine solche z.B. durch die Gesamtheit aller Punkte mit rationalen Koordinaten $\mathbb{Q}^d := \{(x_1,\dots,x_d) : x_i \in \mathbb{Q},\ 1 \le i \le d\}$ gegeben. Aufgrund der Separabilität kann jede offene Menge in $\mathbb{R}^d$ als abzählbare Vereinigung offener Kugeln mit Mittelpunkt $x \in \mathbb{Q}^d$ und rationalem Radius $\varepsilon > 0$ geschrieben werden: also ist auch

$$\mathcal{C} \ := \ \{\, B_\varepsilon(x) \ : \ x \in \mathbb{Q}^d,\ \varepsilon > 0 \text{ rational } \}$$

ein Erzeuger der Borel-σ-Algebra $\mathcal{B}(\mathbb{R}^d)$.

b) Ein Erzeuger von $\mathcal{B}(\mathbb{R}^d)$ ist auch jede der folgenden Klassen $\mathcal{C}_i$ d-dimensionaler Intervalle:

$$\begin{aligned}
\mathcal{C}_1 &:= \{(-\infty, b] \ : \ -\infty < b < \infty\} \\
\mathcal{C}_2 &:= \{(a, b] \ : \ -\infty < a < b < \infty\} \\
\mathcal{C}_3 &:= \{[a, b] \ : \ -\infty < a < b < \infty\} \\
\mathcal{C}_4 &:= \{(a, b) \ : \ -\infty < a < b < \infty\}
\end{aligned}$$

(wir schreiben $a = (a_1,\dots,a_d)$, $b = (b_1,\dots,b_d)$ mit Notation $a < b$ falls $a_i < b_i$ für alle $i = 1,\dots,d$; $a \le b$ bedeutet $a_i \le b_i$ für alle $i = 1,\dots,d$; d-dimensionale Intervalle werden als $(a,b] := \mathop{\times}_{i=1}^{d}(a_i,b_i]$, $[a,b) := \mathop{\times}_{i=1}^{d}[a_i,b_i)$ etc. geschrieben).

Beweis: 1) Das folgende Argument wird wiederholt benutzt werden: ist $\mathcal{A}'$ eine σ-Algebra in Ω', ist $\mathcal{C}'$ ein System von Teilmengen von $\mathcal{A}'$, so gilt $\sigma(\mathcal{C}') \subset \mathcal{A}'$. Dies folgt sofort aus 1.5.

2) Stets gilt $\mathcal{C} \subset \mathcal{O} \subset \sigma(\mathcal{O})$, also $\sigma(\mathcal{C}) \subset \mathcal{B}(\mathbb{R}^d)$. Separabilität liefert $\mathcal{O} \subset \sigma(\mathcal{C})$ und damit die umgekehrte Inklusion $\mathcal{B}(\mathbb{R}^d) \subset \sigma(\mathcal{C})$. Dies zeigt a).

Dasselbe Argument mit offenen Quadern statt offenen Kugeln liefert $\mathcal{B}(\mathbb{R}^d) = \sigma(\mathcal{C}_4)$.

3) Mit Darstellungen der Art

$$\begin{aligned} (a,b] &= \bigcap_{n\geq 1} (a, b + (\frac{1}{n}, \ldots, \frac{1}{n})) \\ (a,b) &= \bigcup_{n \geq n_0(a,b)} (a, b - (\frac{1}{n}, \ldots, \frac{1}{n})] \end{aligned}$$

sieht man $\mathcal{C}_2 \subset \sigma(\mathcal{C}_4)$ und $\mathcal{C}_4 \subset \sigma(\mathcal{C}_2)$. Beweisschritt 1) liefert also $\sigma(\mathcal{C}_2) = \sigma(\mathcal{C}_4)$. Analog sieht man $\sigma(\mathcal{C}_3) = \sigma(\mathcal{C}_2)$.

4) Wegen

$$(-\infty, b] = \bigcup_{n\geq 1} (b - (n, \ldots, n), b]$$

gilt $\mathcal{C}_1 \subset \sigma(\mathcal{C}_2)$. Weiter gilt $\mathcal{C}_2 \subset \sigma(\mathcal{C}_1)$ wegen

$$(a,b] = (-\infty, b] \setminus \left(\bigcup_{i=1}^{d} (-\infty, b^{i,a}] \right)$$

mit Bezeichnung $b^{i,a} := (b_1, \ldots, b_{i-1}, a_i, b_{i+1}, \ldots, b_d)$; dies folgt aus

$$(a,b] = \{\, x = (x_1, \ldots, x_d) \in (-\infty, b] \ : \ x_i > a_i \text{ für alle } 1 \leq i \leq d \,\}$$

$$(-\infty, b^{i,a}] = \{\, x = (x_1, \ldots, x_d) \in (-\infty, b] \ : \ x_i \leq a_i \} \ .$$

Also gilt $\sigma(\mathcal{C}_1) = \sigma(\mathcal{C}_2)$. $\square$

Die bisher benutzten Begriffe wurden schon in der 'Einführung in die Stochastik' zumindest skizzenhaft präsentiert; der folgende wichtige Begriff ist neu.

1.7 Definition: Ein System $\mathcal{D}$ von Teilmengen von Ω heisst *Dynkin-System*, falls gilt:

i) $\Omega \in \mathcal{D}$

ii) $D \in \mathcal{D} \implies D^c = \Omega \backslash D \in \mathcal{D}$

iii) D_n, $n \geq 1$, $\in \mathcal{D}$ paarweise disjunkt $\implies \dot{\bigcup}_{n\geq 1} D_n \in \mathcal{D}$.

Wir werden sehen, dass man in 1.7 iii) essentiell weniger fordert als an der entsprechenden Stelle 1.1 iii) der Definition einer σ-Algebra.

1.8 Bemerkung: a) Jede σ-Algebra ist insbesondere ein Dynkin-System.

b) Ist $\mathcal{D}$ Dynkin, so gilt auch

i) $\emptyset = \Omega \setminus \Omega \in \mathcal{D}$;

ii) $D_1, \ldots, D_n$ paarweise disjunkt in $\mathcal{D} \implies D_1 \dot{\cup} \ldots \dot{\cup} D_n \in \mathcal{D}$

(wegen $\emptyset \in \mathcal{D}$ schreibt man $D_1 \dot{\cup} \ldots \dot{\cup} D_n = D_1 \dot{\cup} \ldots \dot{\cup} D_n \dot{\cup} \emptyset \dot{\cup} \emptyset \ldots$).

c) Vorsicht: Dynkin-Systeme sind i.a. *nicht* stabil bezüglich Bildung von Durchschnitten oder von beliebigen Vereinigungen (insofern wird in 1.7 iii) essentiell weniger als 1.1 iii) gefordert):
Betrachte als Beispiel eine endliche Menge Ω mit gerader Mächtigkeit $|\Omega| = 2n$, $n \geq 2$. Dann ist

$$\mathcal{D} := \{A \subset \Omega \ : \ |A| = 2\ell, \ 0 \leq \ell \leq n\}$$

sicher ein Dynkin-System. Für verschieden gewählte $\omega_1, \ldots, \omega_{2\ell-1}, \omega', \omega''$ aus Ω und für Mengen

$$A' := \{\omega_1, \ldots, \omega_{2\ell-1}, \omega'\} \in \mathcal{D}, \quad A'' := \{\omega_1, \ldots, \omega_{2\ell-1}, \omega''\} \in \mathcal{D}$$

haben aber Durchschnitt $A' \cup A''$ und Vereinigung $A' \cap A''$ ungerade Mächtigkeit, folglich

$$A' \cup A'' \notin \mathcal{D} \ , \quad A' \cap A'' \notin \mathcal{D} \ .$$

Insbesondere ist $\mathcal{D}$ keine σ-Algebra. □

1.8' Definition: Nenne ein Mengensystem $\mathcal{C}$ $\cap$-*stabil* falls

$$C_1, C_2 \in \mathcal{C} \quad \Longrightarrow \quad C_1 \cap C_2 \in \mathcal{C} \ .$$

1.9 Satz: Sei $\mathcal{D}$ ein Dynkin-System auf einer Menge Ω. Dann gilt:

$$\mathcal{D} \text{ ist } \sigma\text{-Algebra} \quad \Longleftrightarrow \quad \mathcal{D} \text{ ist } \cap\text{-stabil} \ .$$

Beweis: Die Richtung '$\Longrightarrow$' ist bekannt aus 1.2 iii) und 1.8 a). Wir zeigen '$\Longleftarrow$'.

1) Für Dynkin-Systeme $\mathcal{D}$ gilt die folgende (im Vergleich zu 1.2 iv) schwächere) Eigenschaft

$$D_1, D_2 \in \mathcal{D}, \quad D_1 \subset D_2 \implies D_2 \backslash D_1 \in \mathcal{D} \ :$$

wegen $D_1 \subset D_2$ sind $D_1 \in \mathcal{D}$ und $D_2^c \in \mathcal{D}$ disjunkt, also gilt nach 1.7 ii) und iii)

$$D_2 \backslash D_1 = (\underbrace{D_1 \dot{\cup} D_2^c}_{\in \mathcal{D}})^c \in \mathcal{D} \ .$$

2) Sei nun $\mathcal{D}$ Dynkin und $\cap$-stabil. Wir zeigen 1.1 iii), dann ist $\mathcal{D}$ sogar eine σ-Algebra. Betrachte eine beliebige Mengenfolge $(A_n)_n \subset \mathcal{D}$. Setze

$$D_1 \;:=\; A_1 \;\in \mathcal{D}\;,$$

dann mit der vorausgesetzten $\cap$-Stabilität von $\mathcal{D}$ und mit Schritt 1)

$$D_2 \;:=\; A_2 \setminus (A_1 \cap A_2) \;\in \mathcal{D}\;, \quad A_1 \cup A_2 = D_1 \,\dot\cup\, D_2 \in \mathcal{D}\;.$$

Sind schon $D_1, \ldots, D_n$ paarweise disjunkt in $\mathcal{D}$ gefunden mit $\bigcup_{i=1}^{n} A_i = \dot{\bigcup}_{i=1}^{n} D_i \in \mathcal{D}$, so bilde weiter

$$D_{n+1} := A_{n+1} \setminus \left((\dot{\bigcup_{i=1}^{n}} D_i) \cap A_{n+1} \right) \;\in \mathcal{D}$$

wegen $\cap$-Stabilität von $\mathcal{D}$ und Schritt 1): damit sind $D_1, \ldots, D_{n+1}$ paarweise disjunkt, und

$$\bigcup_{i=1}^{n+1} A_i \;=\; \dot{\bigcup_{i=1}^{n+1}} D_i \;\in \mathcal{D}\;.$$

Sukzessiv in n erhält man eine Folge D_n, $n \geq 1$, paarweise disjunkter Mengen in $\mathcal{D}$ so dass

$$\bigcup_{n \geq 1} A_i \;=\; \dot{\bigcup_{n \geq 1}} D_i \quad \in \mathcal{D}\;.$$

Damit ist die Eigenschaft 1.1 iii) nachgewiesen. $\square$

1.10 Bemerkung: Genau analog zu σ-Algebren prüft man nach:

a) Sind $\mathcal{D}_i$, $i \in I$, Dynkin-Systeme auf derselben Menge Ω, so ist auch

$$\mathcal{D} \;:=\; \bigcap_{i \in I} \mathcal{D}_i$$

ein Dynkin-System. Dabei darf I eine beliebige Indexmenge sein.

b) Für jedes System $\mathcal{C}$ von Teilmengen von Ω existiert ein kleinstes Dynkin-System, welches $\mathcal{C}$ umfasst: dies bezeichnet man mit $\delta(\mathcal{C})$. Man nennt $\delta(\mathcal{C})$ das *von $\mathcal{C}$ erzeugte Dynkin-System.*

1.11 Satz: Ist $\mathcal{C}$ ein $\cap$-stabiles System von Teilmengen von Ω, so gilt

$$\delta(\mathcal{C}) \;=\; \sigma(\mathcal{C})\;.$$

Beweis: Als σ-Algebra ist $\sigma(\mathcal{C})$ insbesondere ein Dynkin-System, welches $\mathcal{C}$ umfasst: also gilt $\delta(\mathcal{C}) \subset \sigma(\mathcal{C})$ nach 1.10 b). Wir werden zeigen

$$(*) \qquad \mathcal{C} \;\; \cap\text{-stabil} \quad \Longrightarrow \quad \delta(\mathcal{C}) \;\; \cap\text{-stabil}\;.$$

Mit 1.9 und $(*)$ und der Voraussetzung über $\mathcal{C}$ wird $\delta(\mathcal{C})$ eine σ-Algebra, welche $\mathcal{C}$ umfasst. Damit gilt die Inklusion $\sigma(\mathcal{C}) \subset \delta(\mathcal{C})$. Damit hat man '=', und der Beweis ist abgeschlossen. Wir beweisen $(*)$ in zwei Schritten.

1) Sei $\mathcal{C}$ $\cap$-stabil, sei $D \in \delta(\mathcal{C})$ beliebig, betrachte

$$\mathcal{H}_D \;:=\; \{A \subset \Omega \;:\; D \cap A \in \delta(\mathcal{C})\} \;.$$

Dann ist $\mathcal{H}_D$ ein Dynkin-System:

Klar gilt $\Omega \in \mathcal{H}_D$: das ist 1.7 i). Sei $A \in \mathcal{H}_D$. Nach Definition der Klasse $\mathcal{H}_D$ dann $D \cap A \in \delta(\mathcal{C})$, also dank der Voraussetzung $D \in \delta(\mathcal{C})$ wie in Schritt 1) des Beweises von 1.9

$$D \cap A^c \;=\; D \setminus (D \cap A) \;\in\; \delta(\mathcal{C})$$

und damit $A^c \in \mathcal{H}_D$: das ist 1.7 ii). Sei $(A_n)_{n\geq 1} \subset \mathcal{H}_D$ eine Folge paarweise disjunkter Mengen: dann $(D \cap A_n)_n$ paarweise disjunkt und in $\delta(\mathcal{C})$ nach Definition von $\mathcal{H}_D$, also

$$D \cap (\dot{\bigcup_n} A_n) \;=\; \dot{\bigcup_n}(D \cap A_n) \;\in \delta(\mathcal{C})$$

da $\delta(\mathcal{C})$ Dynkin. Nach Definition von $\mathcal{H}_{\mathcal{D}}$ heisst das $(\dot{\bigcup_n} A_n) \in \mathcal{H}_D$: das ist 1.7 iii).

2) Für beliebiges festes $E \in \mathcal{C}$ betrachte nun

$$\mathcal{H}_E \;:=\; \{A \subset \Omega \;:\; E \cap A \in \delta(\mathcal{C})\} \;.$$

Da $\mathcal{C}$ $\cap$-stabil nach Voraussetzung, gilt $\mathcal{C} \subset \mathcal{H}_E$. Wie in Schritt 1) ist $\mathcal{H}_E$ Dynkin. Damit ist $\mathcal{H}_E$ ein Dynkin-System, welches $\mathcal{C}$ umfasst: nach 1.10 b) folgt $\delta(\mathcal{C}) \subset \mathcal{H}_E$. Das aber bedeutet

$$E \in \mathcal{C} \text{ beliebig}, \quad D \in \delta(\mathcal{C}) \text{ beliebig} \quad \Longrightarrow \quad E \cap D \in \delta(\mathcal{C}) \;.$$

Nun vertauschen wir die Rollen von E und D und lesen die letzte Formelzeile als

$$D \in \delta(\mathcal{C}) \text{ beliebig} \quad \Longrightarrow \quad \mathcal{C} \subset \mathcal{H}_D \;.$$

Da aber $\mathcal{H}_D$ Dynkin, folgt hieraus

$$D \in \delta(\mathcal{C}) \text{ beliebig} \quad \Longrightarrow \quad \delta(\mathcal{C}) \subset \mathcal{H}_D \;.$$

Damit ist die $\cap$-Stabilität von $\delta(\mathcal{C})$ gezeigt: das ist $(*)$, und Satz 1.11 ist vollständig bewiesen. $\square$

Mit Hilfe von Dynkin-Systemen wird das erste Hauptergebnis dieses Kapitels (der Eindeutigkeitssatz 1.13 für Masse) bewiesen werden.

1.12 Definition: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Eine Mengenfunktion $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$ heisst *Mass auf* $\mathcal{A}$, falls für jede Folge $(A_n)_{n \geq 1} \subset \mathcal{A}$ paarweise disjunkter Mengen gilt

$$\mu(\dot{\bigcup_{n\geq 1}} A_n) \;=\; \sum_{n=1}^{\infty} \mu(A_n)$$

(σ-*Additivität*). Ein Mass μ auf $\mathcal{A}$ heisst *endlich*, falls $\mu(\Omega) < \infty$. Ein *Wahrscheinlichkeitsmass* auf $\mathcal{A}$ ist ein endliches Mass mit Normierung $\mu(\Omega) = 1$.

Bemerkung: Mit einer Darstellung $\dot{\bigcup\limits_{i=1}^{n}} A_n \;=\; A_1 \dot\cup \ldots \dot\cup A_n \dot\cup \emptyset \dot\cup \emptyset \dot\cup \ldots$ endlicher disjunkter Vereinigungen ist ein Mass auf $\mathcal{A}$ nach 1.12 insbesondere *additiv* (d.h.: endlich additiv) auf $\mathcal{A}$. Aus der Additivität folgt die Eigenschaft der *Monotonie*: für Mengen A, B in $\mathcal{A}$ mit $A \subset B$ gilt stets $\mu(A) \leq \mu(B)$.

1.12' Definition: (*aufsteigende/absteigende Mengenfolgen*) Für $(A_n)_n \subset \mathcal{P}(\Omega)$ und $A \subset \Omega$ schreibe $A_n \uparrow A$ falls

$$A_n \subset A_{n+1} \quad \text{für alle } n, \text{ und} \quad A \;=\; \bigcup_n A_n \;;$$

schreibe $A_n \downarrow A$ falls

$$A_n \supset A_{n+1} \quad \text{für alle } n, \text{ und} \quad A \;=\; \bigcap_n A_n \;.$$

Der folgende erste Hauptsatz dieses Kapitels erlaubt es, Masse auf einer σ-Algebra allein anhand ihrer Werte auf *geeigneten* Erzeugern der σ-Algebra zu identifizieren.

1.13 Eindeutigkeitssatz: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, sei $\mathcal{C}$ ein Erzeuger von $\mathcal{A}$, seien μ_1 und μ_2 Masse auf $\mathcal{A}$ mit

$$\mu_1(C) \;=\; \mu_2(C) \quad \forall\, C \in \mathcal{C} \,.$$

Ist $\mathcal{C}$ $\cap$-stabil *und* gibt es in $\mathcal{C}$ eine Mengenfolge $(C_n)_{n\in I\!N}$ mit

$$C_n \uparrow \Omega \,, \quad \mu_i(C_n) \;<\; \infty \quad \text{für alle } n \in I\!N, \;\; i = 1, 2 \,,$$

so stimmen μ_1 und μ_2 als Wahrscheinlichkeitsmasse auf $\mathcal{A}$ überein:

$$\mu_1(A) \;=\; \mu_2(A) \quad \text{für alle } A \in \mathcal{A}\,.$$

Beweis: 1) Sei $E \in \mathcal{C}$ mit $(\mu_1(E) =)\, \mu_2(E) < \infty$ gegeben; wir betrachten die Restriktionen $\mu_i(\cdot \cap E)$ der Masse μ_i auf E. Definiere

$$\mathcal{H}_E \;:=\; \{A \in \mathcal{A} \,:\, \mu_1(A \cap E) = \mu_2(A \cap E)\}\,.$$

i) Dann ist $\mathcal{H}_E$ Dynkin:

Klar ist $\Omega \in \mathcal{H}_E$. Mit $A \in \mathcal{H}_E$ gilt auch $A^c \in \mathcal{H}_E$, denn man hat

$$\mu_1(A^c \cap E) = \mu_1(E) - \mu_1(A \cap E) = \mu_2(E) - \mu_2(A \cap E) = \mu_2(A^c \cap E)$$

wegen $\mu_1(E) = \mu_2(E) < \infty$ (nach Wahl von E) und $\mu_1(A \cap E) = \mu_2(A \cap E)$ (da $A \in \mathcal{H}_E$). Seien nun A_n, $n \geq 1$, in $\mathcal{H}_E$ paarweise disjunkt. Dann sind auch $A_n \cap E$, $n \geq 1$, paarweise disjunkt in $\mathcal{A}$, und σ-Additivität von μ_i kombiniert mit der definierenden Eigenschaft von $\mathcal{H}_E$ zeigt

$$\begin{aligned}
\mu_1\left((\dot{\bigcup_{n\geq 1}} A_n) \cap E \right) \;&=\; \mu_1\left(\dot{\bigcup_{n\geq 1}} (A_n \cap E) \right) \;=\; \sum_{n\geq 1} \mu_1(A_n \cap E) \\
&=\; \sum_{n\geq 1} \mu_2(A_n \cap E) \;=\; \ldots \;=\; \mu_2\left((\dot{\bigcup_{n\geq 1}} A_n) \cap E \right);
\end{aligned}$$

also ist $\dot{\bigcup\limits_{n\geq 1}} A_n \in \mathcal{H}_E$. Damit sind die Eigenschaften 1.7 eines Dynkin-Systems nachgewiesen.

ii) Wir zeigen $\mathcal{H}_E = \mathcal{A}$:

Da $\mathcal{C}$ $\cap$-stabil und $\mu_1 = \mu_2$ auf $\mathcal{C}$, gilt

$$\mathcal{C} \subset \mathcal{H}_E\;;$$

da $\mathcal{H}_E$ Dynkin, folgt mit 1.10 b)

$$\delta(\mathcal{C}) \subset \mathcal{H}_E\,.$$

$\cap$-Stabilität von $\mathcal{C}$ impliziert aber auch nach 1.11

$$\delta(\mathcal{C}) \;=\; \sigma(\mathcal{C}) \;=\; \mathcal{A}$$

und damit $\mathcal{H}_E = \mathcal{A}$. Das ist ii).

Für beliebiges $E \in \mathcal{C}$ mit $\mu_1(E) = \mu_2(E) < \infty$ ist wegen i) und ii) gezeigt:

$$(+) \qquad \mu_1(A \cap E) \;=\; \mu_2(A \cap E) \quad \text{für alle } A \in \mathcal{A}\,.$$

Bis hierhin wurde nur die $\cap$-Stabilität des Erzeugers $\mathcal{C}$ ausgenutzt.

2) Wir zeigen

$$\mu_1(A) \;=\; \mu_2(A) \quad \text{für alle } A \in \mathcal{A} \;:$$

Wähle mit der zweiten Voraussetzung des Satzes eine Folge $(C_n)_n \subset \mathcal{C}$ mit $C_n \uparrow \Omega$ und $\mu_1(C_n) = \mu_2(C_n) < \infty$ für alle $n \geq 1$. Setze $C_0 := \emptyset$ und definiere $B_n := C_n \setminus C_{n-1}$, $n \geq 1$ (diese liegen in $\mathcal{A}$, und i.a. nicht mehr im Erzeuger $\mathcal{C}$). Die B_n, $n \geq 1$ sind paarweise disjunkt in $\mathcal{A}$, und erfüllen $\Omega = \dot{\bigcup\limits_{n\geq 1}} B_n$ sowie $C_n = B_1 \dot\cup \ldots \dot\cup B_n$. Dann impliziert die σ-Additivität der Masse μ_i auf $\mathcal{A}$

$$\begin{aligned}
\infty \;\geq\; \mu_i(A) \;&=\; \mu_i\left(A \cap (\dot{\bigcup_{n\geq 1}} B_n) \right) \;=\; \mu_i\left(\dot{\bigcup_{n\geq 1}} (A \cap B_n) \right) \\
&=\; \sum_{n\geq 1} \mu_i(A \cap B_n) \;=\; \lim_{N\to\infty} \sum_{n=1}^{N} \mu_i(A \cap B_n) \;=\; \lim_{N\to\infty} \mu_i(A \cap C_N)
\end{aligned}$$

für beliebiges $A \in \mathcal{A}$. Nach (+) mit $E := C_N$ hat man $\mu_1(A \cap C_N) = \mu_2(A \cap C_N)$ für jedes $N \geq 1$, und mit $N \to \infty$ folgt die Behauptung. $\square$

Die letzte Formelzeile dieses Beweises war nichts anderes als die aus der Einführungsvorlesung bekannte *aufsteigende Stetigkeit von Wahrscheinlichkeitsmassen*, die wir in 1.17 unten in etwas allgemeinerem Rahmen neu formulieren werden.

1.14 Beispiel: Sei μ ein Wahrscheinlichkeitsmass auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. Bilde zu μ die Funktion

$$F_\mu : \; \mathbb{R}^d \ni x \;\longrightarrow\; \mu((-\infty, x]) \;=:\; F_\mu(x) \quad \in [0,1]$$

(mit Notation: $x = (x_1, \ldots, x_d)$, $(-\infty, x] = \mathop{\times}\limits_{i=1}^{d} (-\infty, x_i]$, ... wie in 1.6'); F_μ heisst *Verteilungsfunktion zu* μ. Dann gilt: das Wahrscheinlichkeitsmass μ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ist durch seine Verteilungsfunktion eindeutig bestimmt.

Beweis: Das System $\mathcal{C}_1 = \{(-\infty, x] : x \in \mathbb{R}^d\}$ aus 1.6' ist ein $\cap$-stabiler Erzeuger von $\mathcal{B}(\mathbb{R}^d)$, es gibt Mengenfolgen $(C_n)_n \subset \mathcal{C}$ mit $C_n \uparrow \mathbb{R}^d$ und $\mu(C_n) \leq \mu(\mathbb{R}^d) = 1$ für alle n, etwa $C_n := (-\infty, (n, ..., n)]$, $n \geq 1$. Also sind alle Voraussetzungen des Eindeutigkeitssatzes 1.13 erfüllt. Folglich sind Wahrscheinlichkeitsmasse auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ bereits durch ihre Werte auf $\mathcal{C}_1$ eindeutig festgelegt. Genau diese gehen in die Verteilungsfunktion ein. $\square$

1.14' Bemerkung: Mit Voraussetzungen und Bezeichnungen aus 1.14: für d-dimensionale halboffene Rechtecke $(a, b]$, $-\infty < a < b < +\infty$, gilt stets $\mu((a, b]) \geq 0$. Dies übersetzt sich in eine

'Rechteckbedingung', die die Verteilungsfunktion $F := F_\mu$ notwendig erfüllen muss:

$$\text{in Dimension } d=1: \quad F(b) - F(a) \ \geq 0 \quad \text{(d.h.: } F(\cdot) \text{ ist nichtfallend)}$$

$$\text{in Dimension } d=2: \quad F((b_1,b_2)) - F((a_1,b_2)) - F((b_1,a_2)) + F((a_1,a_2)) \ \geq 0$$

sowie allgemein in Dimension $d > 2$ (Chow and Teicher 1978, S. 178; Elstrodt 2002, S. 44):

$$(*) \qquad \left\{ \begin{array}{l} F((b_1,...,b_d)) \ - \ \sum_{j=1}^d F((b_1,...,b_{j-1},a_j,b_{j+1},...,b_d)) \\ + \ \sum_{1\leq j_1<j_2\leq d} F((b_1,...,b_{j_1-1},a_{j_1},b_{j_1+1},...,b_{j_2-1},a_{j_2},b_{j_2+1},...,b_d)) \\ \pm \quad \ldots \quad + \ (-1)^d F((a_1,...,a_d)) \qquad \geq \ 0 \ . \end{array} \right.$$

C. Algebren, Inhalte, Prämasse, Konstruktion von Massen

In Beispiel 1.6' wurden Klassen $\mathcal{C}_i$ sehr einfacher Mengen betrachtet, welche die Borelsche σ-Algebra auf $\mathbb{R}^d$ erzeugen. Will man $\mathbb{R}^d$-wertige Zufallsphänomene modellieren, so hat man in der Regel eine präzise Vorstellung davon, welche Masszahl den (wenigen und einfachen) Mengen aus $\mathcal{C}_i$ zugesprochen werden sollte. Wie kann man sicherstellen, dass solche 'Startvorgaben' durch (genau) ein Wahrscheinlichkeitsmass auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ erfüllt werden ? Ganz allgemein wird diese Frage durch den Fortsetzungssatz von Carathéodory beantwortet (1.18 unten).

1.15 Definition: Ein System $\mathcal{G}$ von Teilmengen einer Menge Ω heisst *Algebra*, falls gilt

i) $\Omega \in \mathcal{G}$

ii) $A \in \mathcal{G} \implies A^c = \Omega \backslash A \in \mathcal{G}$.

iii) $A_1, \ldots, A_n \in \mathcal{G} \implies \bigcup_{i=1}^n A_i \in \mathcal{G}$.

Der Unterschied zur σ-Algebra besteht darin, dass in 1.15 iii) nur Stabilität unter endlichen Vereinigungen gefordert wird; eine Algebra ist stets $\cap$-stabil, enthält $\emptyset$, und mit zwei Mengen A_1, A_2 auch $A_1 \setminus A_2$ (analog 1.2).

1.16 Definition: Sei $\mathcal{G}$ eine Algebra von Teilmengen einer Menge Ω.
a) Eine Mengenfunktion $\mu : \mathcal{G} \to [0, \infty]$ mit $\mu(\emptyset) = 0$ heisst *Inhalt auf* $\mathcal{G}$, falls für paarweise disjunkte Mengen $A_1, \ldots, A_n$ in $\mathcal{G}$ gilt

$$\mu\left(\dot{\bigcup_{i=1}^n} A_i\right) = \sum_{i=1}^n \mu(A_i)$$

(*Additivität*). Ein Inhalt μ auf $\mathcal{G}$ heisst *endlich* falls $\mu(\Omega) < \infty$.

b) Ein Inhalt μ auf einer Algebra $\mathcal{G}$ heisst *σ-endlich* falls

$$\text{es existieren } B_n,\ n \geq 1, \text{ in } \mathcal{G} \text{ mit } B_n \uparrow \Omega \text{ und } \mu(B_n) < \infty \text{ für jedes } n\ .$$

c) Ein *Prämass* μ auf einer Algebra $\mathcal{G}$ ist ein Inhalt, welcher sogar *σ-additiv* ist: sind A_n, $n \geq 1$, paarweise disjunkt in $\mathcal{G}$, und ist $\dot{\bigcup}_{n=1}^{\infty} A_n \in \mathcal{G}$, dann gilt

$$\mu\left(\dot{\bigcup_{n=1}^{\infty}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

1.16' Bemerkung: Vergleich von 1.12 und 1.16 c) zeigt: ein Mass ist ein Prämass, welches auf einer σ-Algebra definiert ist.

1.17 Satz: Sei $\mathcal{G}$ eine Algebra auf Ω, sei μ ein Inhalt auf $\mathcal{G}$. Betrachte Eigenschaften i) - iv):

i) μ ist σ-additiv.

ii) μ ist *aufsteigend stetig*: für $(A_n)_n \subset \mathcal{G}$ und $A \in \mathcal{G}$ gilt: aus $A_n \uparrow A$ folgt $\lim\limits_{n\to\infty} \mu(A_n) = \mu(A)$.

iii) μ ist *absteigend stetig*: für $A \in \mathcal{G}$ und $(A_n)_n \subset \mathcal{G}$ mit $\mu(A_n) < \infty$ für $n \geq n_0$ gilt: aus $A_n \downarrow A$ folgt $\lim\limits_{n\to\infty} \mu(A_n) = \mu(A)$.

iv) μ ist *stetig in $\emptyset$*: für $(A_n)_n \subset \mathcal{G}$ mit $\mu(A_n) < \infty$ für $n \geq n_0$ gilt: aus $A_n \downarrow \emptyset$ folgt $\lim\limits_{n\to\infty} \mu(A_n) = 0$.

Dann gelten die folgenden Inklusionen:

$$\text{i)} \iff \text{ii)}, \quad \text{iii)} \iff \text{iv)}, \quad \text{ii)} \Longrightarrow \text{iii)}\ .$$

Ist $\mu(\Omega) < \infty$, so gilt auch iii) $\Longrightarrow$ ii): in diesem Fall sind alle Eigenschaften i) - iv) äquivalent.

Beweis: i) $\iff$ ii): Betrachte eine Folge $(A_n)_{n\geq 1} \subset \mathcal{G}$, $A_n \uparrow A$, setze $A_0 := \emptyset \in \mathcal{G}$, dann sind $B_n := A_n \setminus A_{n-1}$, $n \geq 1$, paarweise disjunkt, $(B_n)_n \subset \mathcal{G}$, und $\dot{\bigcup}_{n\geq 1} B_n = \bigcup_{n\geq 1} A_n = A$. Sind umgekehrt B_n, $n \geq 1$, in $\mathcal{G}$ paarweise disjunkt vorgegeben, dann ist $A_n := \dot{\bigcup}_{i=1}^{n} B_i$ aufsteigend gegen $A := \bigcup_{n\geq 1} A_n = \dot{\bigcup}_{n\geq 1} B_n$. Unter der Voraussetzung $A \in \mathcal{G}$ ist folglich σ-Additivität von μ entlang $(B_n)_n$ äquivalent zu aufsteigender Stetigkeit von μ entlang $(A_n)_n$.

iii) $\Longleftrightarrow$ iv): Klar ist iv) ein Spezialfall von iii); zeige also iv) $\Longrightarrow$ iii): Sei $A_n \downarrow A$ in $\mathcal{G}$, sei $\mu(A_n) < \infty$ für $n \geq n_0$. Zerlege für $n \geq n_0$ $A_n = B_n \dot\cup A$ mit $B_n := A_n \backslash A \in \mathcal{G}$. Dann liefert Stetigkeit von μ in $\emptyset$ entlang der Folge $B_n \downarrow \emptyset$ zusammen mit Additivität von μ die Aussage iii).
ii) $\Longrightarrow$ iii): Betrachte $A_n \downarrow A$ in $\mathcal{G}$, sei $\mu(A_n) < \infty$ für $n \geq n_0$. Setze $B_m := A_{n_0} \backslash A_{n_0+m} \in \mathcal{G}$, $m \geq 1$, und $B := A_{n_0} \setminus A \in \mathcal{G}$. Dann gilt $B_n \uparrow B$ in $\mathcal{G}$, und aufsteigende Stetigkeit von μ entlang $(B_m)_m$ kombiniert mit der Additivität von μ zeigt

$$\mu(A_{n_0+m}) = \mu(A_{n_0}) - \mu(B_m) \downarrow \mu(A_{n_0}) - \mu(B) = \mu(A)$$

und damit iii).
iii) $\Longrightarrow$ ii) unter der Zusatzvoraussetzung $\mu(\Omega) < \infty$: Betrachte zusammen mit einer aufsteigenden Mengenfolge $A_n \uparrow A$ in $\mathcal{G}$ die absteigende Mengenfolge $A_n^c \downarrow A^c$ in $\mathcal{G}$. Ist $\mu(\Omega)$ endlich, so

$$\mu(A_n) = \mu(\Omega) - \mu(A_n^c) , \quad \mu(A) = \mu(\Omega) - \mu(A^c) ,$$

und Additivität von μ transformiert absteigende Stetigkeit von μ entlang $(A_n^c)_n$ in aufsteigende Stetigkeit von μ entlang $(A_n)_n$. □

In der Einführungsvorlesung wurde die Äquivalenz aller vier Aussagen aus 1.17 für Wahrscheinlichkeitsmasse formuliert.

1.17' Beispiel: Sei μ ein Wahrscheinlichkeitsmass auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$; betrachte wie in 1.14 die zu μ assoziierte Verteilungsfunktion $F := F_\mu : x \to \mu((-\infty, x])$ auf $\mathbb{R}^d$. Als Folgerung aus aufsteigender und absteigender Stetigkeit von μ nach 1.17 erhält man folgende Stetigkeitseigenschaften:

$$(+) \qquad \begin{cases} \lim_{y \to -\infty} F(x_1, ..., x_{i-1}, y, x_{i+1}, ..., x_d) = 0 \quad \text{für beliebiges } x \in \mathbb{R}^d,\ 1 \leq i \leq d \\ \lim_{y \to +\infty} F((y, ..., y)) = 1 \\ \text{für } (x^{(n)})_n \text{ komponentenweise nichtwachsend mit } x^{(n)} \to x \text{ in } \mathbb{R}^d\text{:} \quad F(x^{(n)}) \to F(x) . \end{cases}$$

In Dimension $d = 1$ reduziert sich (+) zu

$$\lim_{t \downarrow -\infty} F(t) = 0 , \ \lim_{t \uparrow +\infty} F(t) = 1 , \quad F \text{ ist rechtsstetig} . \qquad \square$$

Zur Bedeutung der dritten Aussage in (+): betrachtet man im Erzeuger $\mathcal{C}_2$ der σ-Algebra $\mathcal{B}(\mathbb{R}^d)$ Mengen $A := (a, b]$, $-\infty < a < b < \infty$, $A_m := \big((a_1 + \frac{1}{m}, \ldots, a_d + \frac{1}{m}), b\big]$, $m \in \mathbb{N}$ hinreichend gross, und drückt man $\mu(A_m)$ sowie $\mu(A)$ mit Hilfe der Rechteckbedingung $(*)$ aus 1.14' durch die Verteilungsfunktion F aus, so liefert die dritte Aussage in (+) –anzuwenden auf jeden der Summanden in der Rechteckbedingung $(*)$– die absteigende Stetigkeit $\mu(A_m) \downarrow \mu(A)$.

Der zweite Hauptsatz in diesem Kapitel ist

1.18 Fortsetzungssatz (Carathéodory) Jedes Prämass μ auf einer Algebra $\mathcal{G}$ kann auf mindestens eine Weise zu einem Mass $\widetilde{\mu}$ auf der von $\mathcal{G}$ erzeugten σ-Algebra $\mathcal{A} := \sigma(\mathcal{G})$ fortgesetzt werden.

Der Beweis dieses Satzes wird im Teilkapitel D gegeben werden. Zuerst zeigen wir in Fortsetzung von 1.6', 1.14 und 1.17' an Beispielen, was dieser Satz typischerweise leistet.

1.19 Beispiel: In Dimension $d = 1$ schreibe $\mathcal{F}$ für die Klasse aller Funktionen $F : \mathbb{R} \to [0, 1]$ mit der Eigenschaft

$$(*) \qquad F(\cdot) \text{ ist rechtsstetig, nichtfallend, } \quad \lim_{x\to-\infty} F(x) = 0, \quad \lim_{x\to+\infty} F(x) = 1 \,.$$

Dann gilt: Zu jedem $F \in \mathcal{F}$ existiert genau ein Wahrscheinlichkeitsmass $P = P_F$ auf $\mathcal{B}(\mathbb{R})$ mit

$$(\diamond) \qquad P_F((-\infty, b]) = F(b) \quad \forall\, b \in \mathbb{R} \,.$$

Damit gibt es eine *1-1-Zuordnung* zwischen der Klasse $\mathcal{F}$ und der Klasse aller Wahrscheinlichkeitsmasse auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. $\mathcal{F}$ ist daher die *Klasse aller Verteilungsfunktionen auf* $\mathbb{R}$.

Beweis : 1) Der Eindeutigkeitssatz 1.13 angewandt auf den Erzeuger $\mathcal{C}_1 = \{(-\infty, b] : b \in \mathbb{R}\}$ der Borel-σ-Algebra $\mathcal{B}(\mathbb{R})$ zeigt: für festes $F \in \mathcal{F}$ gibt es *höchstens* ein Wahrscheinlichkeitsmass P_F auf $\mathcal{B}(\mathbb{R})$ mit der Eigenschaft $(\diamond)$. Es ist also hier nur die *Existenz* von Wahrscheinlichkeitsmassen mit $(\diamond)$ zu vorgegebenem $F \in \mathcal{F}$ aus dem Fortsetzungssatz 1.18 herzuleiten.
2) Die kleinste $\mathcal{C}_1$ enthaltende *Algebra* von Teilmengen von $\mathbb{R}$ ist

$$\mathcal{G} \;:=\; \{\emptyset, \mathbb{R}\} \;\cup\; \{\dot{\bigcup_{\ell=1}^{r}} I_\ell \;:\; I_\ell \in \widetilde{\mathcal{C}},\; r \geq 1\} \;;$$

dabei bezeichnet $\widetilde{\mathcal{C}}$ das System von Intervallen

$$\widetilde{\mathcal{C}} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \{(b, \infty) \;:\; b \in \mathbb{R}\}$$

mit $\mathcal{C}_2 := \{(a, b] : -\infty < a < b < \infty\}$ wie in 1.6'. Beachte dabei, dass sowohl der volle Raum $\Omega = \mathbb{R}$ als auch jede endliche Vereinigung von Intervallen aus $\widetilde{\mathcal{C}}$ stets in Form einer endlichen *disjunkten* Vereinigung geeigneter Intervalle aus $\widetilde{\mathcal{C}}$ geschrieben werden kann, und dass $\mathcal{G}$ stabil unter Bildung endlicher Vereinigungen und stabil bezüglich Komplementbildung ist.

3) Zu gegebenem $F \in \mathcal{F}$ definieren wir zuerst einen *Inhalt* $\mu = \mu_F$ auf der Algebra $\mathcal{G}$.

Ausgehend von ($\diamond$) setzen wir

$$\mu(\mathbb{R}) := 1 \;, \quad \mu((-\infty, b]) := F(b) \;, \quad \mu((b, \infty)) = 1 - F(b) \;, \quad b \in \mathbb{R}$$

und erzwingen die gewünschte Additivität durch

$$\mu((a, b]) \;:=\; F(b) - F(a) \;, \quad -\infty < a < b < +\infty \;.$$

Es bleibt nur zu überlegen, dass μ auf $\mathcal{G}$ wohldefiniert ist: für verschiedene Darstellungen

$$\dot{\bigcup_{k=1}^{n}} J_k \;=\; B \;= \dot{\bigcup_{\ell=1}^{m}} I_\ell, \quad I_\ell \,, J_k \in \widetilde{\mathcal{C}}$$

derselben Menge $\emptyset \neq B \neq \Omega$ in $\mathcal{G}$ muss gelten

$$\sum_{k=1}^{n} \mu(J_k) = \sum_{\ell=1}^{m} \mu(I_\ell) \;.$$

Das sieht man so. Es genügt, $J := J_1$ und damit den Fall $n = 1$ zu betrachten. Ist $J = (a, b]$ ein halboffenes Intervall, schreibe $J = \dot{\bigcup}_{\ell=1}^{m} (J \cap I_\ell)$ in Form einer disjunkten Vereinigung

$$\dot{\bigcup_{i=1}^{m}} (a_i, b_i] \;\text{ mit }\; a = a_1 < b_1 = a_2 < \ldots < b_{m-1} = a_m < b_m = b \;\;;$$

dann hat man

$$\mu\left(\dot{\bigcup_{\ell=1}^{m}} (J \cap I_\ell)\right) = \mu\left(\dot{\bigcup_{i=1}^{m}} (a_i, b_i]\right) = \sum_{i=1}^{m} (F(b_i) - F(a_i)) = F(b) - F(a) = \mu(J).$$

Falls J eine Halbgerade $(-\infty, b]$ oder (b, ∞) ist, formuliert man das Argument analog.

4) Wir zeigen: der Inhalt μ auf $\mathcal{G}$ ist sogar σ-additiv.

Damit wird μ ein *Prämass* auf $\mathcal{G}$ (in diesem Beweisschritt steckt üblicherweise die Hauptschwierigkeit bei der Konstruktion von Massen aus einem vorgegebenen Inhalt). Wegen 1.17 und $\mu(\mathbb{R}) = 1 < \infty$ reicht es, die Stetigkeit von μ in $\emptyset$ nachzuweisen.

i) Zeige: zu jedem $I \in \widetilde{\mathcal{C}}$ (I ist nichtleer nach Definition) und zu jedem $\varepsilon > 0$ gibt es ein halboffenes Intervall $J_\varepsilon \in \mathcal{C}_2$ mit

$$K_\varepsilon := \overline{J_\varepsilon} \text{ kompakt}, \quad J_\varepsilon \subset K_\varepsilon \subset I, \quad \mu(J_\varepsilon) \geq \mu(I) - \varepsilon.$$

Beweis: Für $I = (a, b] \in \widetilde{\mathcal{C}}$ wähle $J_\varepsilon = (a + \frac{1}{m}, b]$ für hinreichend grosses m, damit $K_\varepsilon = [a + \frac{1}{m}, b]$. Die Behauptung folgt aus ($*$) und ($\diamond$) für $m \to \infty$, da F rechtsstetig und nichtfallend ist. Für $I = (-\infty, b]$ wähle $J_\varepsilon = (-m, b]$ für geeignetes m, wegen $\lim_{x \to -\infty} F(x) = 0$ in ($*$); für $I = (b, \infty)$ wähle $J_\varepsilon = (b + \frac{1}{m}, m]$ geeignet und benutze zusätzlich $\lim_{x \to \infty} F(x) = 1$.

ii) Zeige: für jede Mengenfolge $(A_n)_n \subset \mathcal{G}$ mit $A_n \downarrow \emptyset$ gilt $\lim_n \mu(A_n) = 0$.

Beweis (indirekt): Wir nehmen an $\lim_n \mu(A_n) =: \varepsilon$ mit $\varepsilon > 0$. $A_n \in \mathcal{G}$ hat eine Darstellung

$$A_n = \dot{\bigcup_{\ell=1}^{r_n}} I_{n\ell}, \quad I_{n\ell} \in \widetilde{\mathcal{C}} \, .$$

Zu $I_{n\ell}$ findet man nach i) $J_{n\ell} \in \mathcal{C}_2$ so dass für jedes $n \geq 1$ gilt:

$$(+) \qquad \begin{cases} J_{n\ell} \subset K_{n\ell} = \overline{J_{n\ell}} \subset I_{n\ell} \\ \text{für } B_n := \dot{\bigcup\limits_{\ell=1}^{r_n}} J_{n\ell} \text{ gilt } \mu(A_n \backslash B_n) < \frac{\varepsilon}{4} \cdot 2^{-n} \, . \end{cases}$$

Betrachte nun die fallende Folge $\bigcap_{i=1}^{n} B_i \in \mathcal{G}$ (und benutze wieder, dass $(A_n)_n$ absteigend ist):

$$\mu\left(A_n \backslash (\bigcap_{i=1}^{n} B_i)\right) = \mu\left(\bigcup_{i=1}^{n} (A_n \backslash B_i)\right) \leq \sum_{i=1}^{n} \mu(A_n \backslash B_i) \leq \sum_{i=1}^{n} \mu(A_i \backslash B_i) \leq \frac{\varepsilon}{2}$$

für alle $n \geq 1$, wegen (+). Damit ist der 'Unterschied' zwischen $(A_n)_n$ und $\left(\bigcap_{i=1}^{n} B_i\right)_n$ nicht sehr gross: aus $\lim_n \mu(A_n) = \varepsilon > 0$ und der letzten Ungleichung folgt

$$(++) \qquad \lim_n \mu(\bigcap_{i=1}^{n} B_i) \geq \frac{\varepsilon}{2} > 0.$$

'Zwischen' $(A_n)_n$ und $\left(\bigcap_{i=1}^{n} B_i\right)_n$ liegt aber die fallende Folge von Kompakta

$$K_n := \bigcap_{i=1}^{n} \left[\dot{\bigcup_{\ell=1}^{r_n}} K_{n\ell} \right], \quad n \geq 1.$$

Die Voraussetzung $A_n \downarrow \emptyset$ impliziert $K_n \downarrow \emptyset$. Wir wenden nun den Cantorschen Durchschnittssatz (siehe z.B. Rudin 1998 S. 44) auf die fallende Folge der Kompakta $(K_n)_n$ an:

$$\bigcap_{n=1}^{\infty} K_n = \emptyset \quad \Longrightarrow \quad \text{es existiert } n_0 < \infty \text{ so dass bereits } K_{n_0} = \emptyset \, .$$

Wegen $K_n \supset \bigcap_{i=1}^{n} B_i$ impliziert dies

$$\bigcap_{i=1}^{n} B_i = \emptyset \quad \text{für alle } n \geq n_0$$

im Widerspruch zu (++). Die Annahme $\lim_n \mu(A_n) =: \varepsilon > 0$ war also absurd, und die behauptete Stetigkeit von μ in $\emptyset$ ist bewiesen. Damit ist μ ein Prämass auf $\mathcal{G}$.

5) Der Fortsetzungssatz 1.18 erlaubt nun, μ (auf mindestens eine Weise) zu einem Mass $\widetilde{\mu}$ auf $\sigma(\mathcal{G}) = \mathcal{B}(I\!R)$ fortzusetzen; per Definition einer *Fortsetzung* erfüllt $\widetilde{\mu}$ notwendigerweise

$$\widetilde{\mu}(A) = \mu(A) \quad \text{für alle } A \in \mathcal{G}$$

und insbesondere wegen $\mathcal{C}_2 \subset \mathcal{G}$

$$\widetilde{\mu}((-\infty, b]) = F(b) \quad \forall\, b \in \mathbb{R} .$$

Das ist ($\diamond$), und die Aussage von 1.19 ist bewiesen. $\square$

1.20 Beispiel: a) In Dimension $d = 2$ definiert man $\mathcal{F}$ als Klasse aller Funktionen $F : \mathbb{R}^2 \to [0, 1]$, für die gilt (vgl. ($*$) in 1.14' und (+) in 1.17')

$$F((b_1, b_2)) - F((a_1, b_2)) - F((b_1, a_2)) + F((a_1, b_1)) \geq 0 , \quad -\infty < a < b < +\infty ,$$

$$\lim_{x_1 \to -\infty} F((x_1, x_2)) = \lim_{x_2 \to -\infty} F((x_1, x_2)) = 0 , \quad \lim_{y \to +\infty} F((y, y)) = 1$$

für $(x^{(n)})_n$ komponentenweise nichtwachsend mit $x^{(n)} \to x$ in $\mathbb{R}^2$: $F(x^{(n)}) \to F(x)$.

Die kleinste das Mengensystem $\mathcal{C}_1$ enthaltende Algebra ist für $d = 2$

$$\mathcal{G} := \{\emptyset, \mathbb{R}^2\} \cup \left\{ \dot{\bigcup_{\ell=1}^{r}} (I_\ell \times J_\ell) \ : \ I_\ell, J_\ell \in \mathcal{J},\ r \geq 1 \right\} ,$$

wobei $\mathcal{J}$ das System aller (eindimensionalen) Intervalle $(-\infty, t]$, $(t, t']$, (t', ∞) bezeichnet, t, t' in $\mathbb{R}$; einige rein notationelle Änderungen des oben gegebenen Beweises zeigen die 1-1-Zuordnung zwischen $\mathcal{F}$ in Dimension $d = 2$ und der Klasse aller Wahrscheinlichkeitsmasse auf $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.
b) Dasselbe funktioniert analog in Dimension $d > 2$: zur Definition der Klasse $\mathcal{F}$ aller Verteilungsfunktionen braucht man die Rechteckbedingung ($*$) aus 1.14' und die Stetigkeitsbedingung (+) aus 1.17', und der Beweis läuft (mit länglichen Notationen) entlang derselben Argumentationslinie. Insbesondere benutzt man ($*$) und (+), um absteigende Stetigkeit $\mu\left(((a_1 + \frac{1}{m}, \ldots, a_d + \frac{1}{m}), b]\right) \downarrow \mu\left((a, b]\right)$ in $\mathcal{C}_2$ und damit die Approximation durch Kompakta zu erhalten. Die Stetigkeitsbedingungen in (+) aus 1.17' implizieren somit, dass der auf der von $\mathcal{C}_2$ erzeugten Algebra $\mathcal{G}$ definierte Inhalt ein Prämass wird. $\square$

Gezeigt ist mit 1.19 und 1.20: für beliebiges $d \geq 1$ stehen Wahrscheinlichkeitsmasse auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ und Funktionen aus $\mathcal{F}$ in eineindeutiger Zuordnung.

Das in Beweisschritt 4) von 1.19 benutzte Argument funktioniert mutatis mutandis in vollständigen separablen metrischen Räumen (siehe auch Bauer 1978, S. 199-201).

1.21 Beispiel: Es gibt genau ein σ-endliches Mass $\lambda\!\!\lambda$ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit der Eigenschaft

$$(*) \qquad \lambda\!\!\lambda\,(\,(a, b]\,) = \prod_{i=1}^{d} |b_i - a_i| \,, \quad -\infty < a < b < \infty$$

(Notationen wie in 1.6'); $\lambda\!\!\lambda$ heisst *Lebesguemass* auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Beweis: 0) Wir pflastern $\mathbb{R}^d =: \dot{\bigcup}_{M \in \mathbb{Z}^d} E_M$ mit halboffenen Würfeln E_M der Seitenlänge 1

$$E_M \ := \ ((M_1, \ldots, M_d), (M_1{+}1, \ldots, M_d{+}1) \,] \ \ , \ \ M = (M_1, \ldots, M_d) \in \mathbb{Z}^d \ ,$$

und schreiben U_N für den Würfel $((-N, ..., -N), (N, ..., N)]$ mit Seitenlänge $2N$, für beliebiges $N \in \mathbb{N}$.

1) Wir betrachten zuerst den Einheitswürfel E_0. Benutzt man nach 1.6' die Klasse $\mathcal{C}_2$ der halboffenen Intervalle als Erzeuger von $\mathcal{B}(\mathbb{R}^d))$, so liefert ein Beweis analog zu 1.19 und 1.20 –als Anwendung des Fortsetzungssatzes 1.18 und des Eindeutigkeitssatzes 1.13– die Existenz genau eines Wahrscheinlichkeitsmasses μ_0 auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit

$$\text{für alle } C = (a,b] \subset E_0 \text{ gilt} \quad \mu(C) = \prod_{i=1}^d |b_i - a_i| \ , \quad \text{und} \ \ \mu(\mathbb{R}^d \setminus E_0) = 0 \, .$$

Dieses heisst *Gleichverteilung auf dem Einheitswürfel* E_0; die zugehörige Verteilungsfunktion ist

$$x \ \longrightarrow \ \prod_{i=1}^d ([0 \vee x_i] \wedge 1) \ , \quad x \in \mathbb{R}^d \ .$$

2) In derselben Weise definiert man für jedes $M \in \mathbb{Z}^d$ die Gleichverteilung μ_M auf dem um $M \in \mathbb{Z}^d$ verschobenen Einheitswürfel E_M, mit Verteilungsfunktion

$$x \ \longrightarrow \ \prod_{i=1}^d ([0 \vee (x_i - M_i)] \wedge 1) \ , \quad x \in \mathbb{R}^d \ .$$

3) Da für $M \in \mathbb{Z}^d$ die halboffenen Würfel E_M paarweise disjunkt sind, erhält man auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ein σ-endliches Mass, indem man definiert:

$$(+) \qquad \lambda\!\!\lambda \ := \ \sum_{M \in \mathbb{Z}^d} \mu_M \ .$$

Die Aussage (+) kann dann in der Form $\lambda\!\!\lambda(\cdot \cap E_M) = \mu_M$ für alle $M \in \mathbb{Z}^d$ geschrieben werden.

4) Betrachte nun die halboffenen Würfel $U_N = ((-N, ..., -N), (N, ..., N)]$ mit Seitenlänge $2N$. Wie in Schritt 1) erhält man auf U_N konzentrierte endliche Masse ν_N mit der Eigenschaft

$$(\times) \qquad \text{für alle } C = (a,b] \subset U_N \text{ gilt} \quad \nu_N(C) = \prod_{i=1}^d |b_i - a_i| \ , \quad \text{und} \ \ \nu_N(\mathbb{R}^d \setminus U_N) = 0 \, .$$

durch die Festsetzung

$$\nu_N \ \ := \sum_{M \in \mathbb{Z}^d \ , \ E_M \subset U_N} \mu_M \ \ = \ \ \lambda\!\!\lambda(\cdot \cap U_N) \ .$$

Ein solches endliches Mass ν_N hat man für jedes $N \in \mathbb{N}$.

5) Abschluss des Beweises: Betrachte das in (+) definierte σ-endliche Mass $\lambda\!\!\lambda$ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ und das System $\mathcal{C}_2$ der halboffenen Intervalle $(a,b] = \bigtimes_{i=1}^{d} (a_i, b_i]$ mit $-\infty < a < b < \infty$ in $\mathbb{R}^d$. Der Eindeutigkeitssatz angewandt auf $\mathcal{C}_2$ als $\cap$-stabilen Erzeuger von $\mathcal{B}(\mathbb{R}^d)$ zeigt, dass es höchstens ein Mass $\widetilde{\lambda\!\!\lambda}$ auf $\mathcal{B}(\mathbb{R}^d)$ gibt, welches den Bedingungen

$$(*) \qquad \widetilde{\lambda\!\!\lambda}\,((a,b]) \;=\; \prod_{i=1}^{d} |b_i - a_i| \;\text{ für alle }\; -\infty < a < b < \infty$$

genügt (beachte bei der Anwendung von 1.13, dass die Vorschrift $(*)$ allen Mengen aus $\mathcal{C}_2$ endliche Masszahlen zuordnet, weshalb Folgen $(V_N)_N$ in $\mathcal{C}_2$ mit $V_N \uparrow \mathbb{R}^d$ –solche existieren offensichtlich– stets die Bedingung $\widetilde{\lambda\!\!\lambda}(V_m) < \infty$ für $m \in \mathbb{N}$ erfüllen). Man sieht leicht, dass durch $\widetilde{\lambda\!\!\lambda} := \lambda\!\!\lambda$ aus (+) die Bedingung $(*)$ erfüllt wird: für jedes $(a,b]$ in $\mathcal{C}_2$ hat man $(a,b] \subset U_N$ für hinreichend grosse N; für diese gilt $\lambda\!\!\lambda((a,b]) = \nu_N((a,b]) = \prod_{i=1}^{d} |b_i - a_i|$ wegen $(\times)$. $\square$

1.21' Bemerkung: a) Auf einer abzählbaren Menge Ω versehen mit der Potenzmenge $\mathcal{P}(\Omega)$ als σ-Algebra definiert man das *Zählmass* durch

$$\mu \;:\; \mathcal{P}(\Omega) \;\ni\; A \;\longrightarrow\; |A| \;\in\; \mathbb{N}_0 \cup \{\infty\}\,.$$

b) Für einen festen Punkt a in einem messbaren Raum $(\Omega, \mathcal{A})$ definiert man das *Diracmass in* a durch

$$\mu \;:\; \mathcal{A} \;\ni\; A \;\longrightarrow\; 1_A(a) \;\in\; \{0,1\}\,.$$

Die übliche Schreibweise für das Diracmass in a ist ϵ_a anstelle von μ.

D. Beweis des Fortsetzungssatzes

In diesem Teilkapitel beweisen wir den Fortsetzungsatz 1.18. Wir starten mit

1.22 Definition: Ein *äusseres Mass* ist eine auf der Potenzmenge von Ω definierte Mengenfunktion $\mu^* : \mathcal{P}(\Omega) \to [0,\infty]$ mit $\mu(\emptyset) = 0$ und den Eigenschaften

$$(1.23) \qquad C_1 \subset C_2 \subset \Omega \text{ beliebig} : \quad \mu^*(C_1) \le \mu^*(C_2) \quad \textit{(Monotonie)}$$

$$(1.24) \qquad (C_n)_n \subset \mathcal{P}(\Omega) \text{ beliebig} : \quad \mu^*\left(\bigcup_{n=1}^{\infty} C_n\right) \le \sum_{n=1}^{\infty} \mu^*(C_n) \quad \textit{(Sub-}\sigma\textit{-Addivität).}$$

Bemerkung: Die Verwendung des Wortbestandteils '... Mass' an dieser Stelle ist unglücklich und irreführend, aber leider fest eingebürgert; ein *äusseres Mass* auf $\mathcal{P}(\Omega)$ ist i.a. *kein* Mass im Sinne der Definition 1.12.

Als Voraussetzung für das ganze Teilkapitel D fixieren wir

(V) eine Algebra $\mathcal{G}$ von Teilmengen von Ω und ein Prämass μ auf $\mathcal{G}$

und nennen für eine beliebige Menge $C \subset \Omega$ jede Familie

$$(A_n)_{n\geq 1} \subset \mathcal{G} \quad \text{mit} \quad \bigcup_{n\geq 1} A_n \supset C$$

eine $\mathcal{G}$-Überdeckung von C (solche gibt es stets, z.B. trivialerweise $A_1 = \Omega$, $A_2 = A_3 = ... = \emptyset$). Mit Hilfe von $\mathcal{G}$-Überdeckungen konstruieren wir zunächst auf $\mathcal{P}(\Omega)$ ein äusseres Mass.

1.25 Hilfssatz: Die durch

$$\mu^*(C) \;:=\; \inf\left\{\sum_{n=1}^{\infty}\mu(A_n) : \; (A_n)_n \subset \mathcal{G}, \; \bigcup_{n\geq 1} A_n \supset C\right\}, \quad C \subset \Omega$$

definierte Mengenfunktion $\mu^* : \mathcal{P}(\Omega) \to [0,\infty]$ ist ein äusseres Mass mit der Eigenschaft

$$\mu^*(A) \;=\; \mu(A) \quad \text{für alle } A \in \mathcal{G}\,.$$

Beweis: 1) Wir zeigen zuerst die zweite Behauptung. Sei $A \in \mathcal{G}$, betrachte irgendeine $\mathcal{G}$-Überdeckung $(A_n)_n$ von A, die das 'inf' in der Definition von $\mu^*(A)$ bis auf einen beliebig kleinen Fehler $\varepsilon > 0$ ausschöpft:

$$\mu^*(A) \leq \sum_{n=1}^{\infty}\mu(A_n) \leq \mu^*(A) + \varepsilon\,.$$

Da $\mathcal{G}$ eine Algebra ist und da $A \in \mathcal{G}$, kann diese Überdeckung so gewählt werden, dass i) die A_n, $n \geq 1$ paarweise disjunkt sind (sonst ersetze A_n durch $A'_n := A_n \setminus (\bigcup_{j=1}^{n-1} A'_j)$, $n \geq 2$, mit $A'_1 := A_1$) und ii) alle A_n als Teilmengen in A enthalten sind (sonst ersetze weiter A'_n durch $A''_n := A'_n \cap A$). Wegen $A \in \mathcal{G}$ gilt dann $\dot{\bigcup}_{n=1}^{\infty} A_n = A \in \mathcal{G}$; da μ ein Prämass auf $\mathcal{G}$ ist, folgt

$$\mu(A) = \sum_{n=1}^{\infty}\mu(A_n)\,.$$

Beide Aussagen zusammen ergeben

$$\mu^*(A) \leq \mu(A) \leq \mu^*(A) + \varepsilon$$

für beliebig kleines $\varepsilon > 0$.

2) Wir zeigen die erste Behauptung des Hilfssatzes: μ^* ist ein äusseres Mass.
Nach 1) gilt $\mu^*(\emptyset) = \mu(\emptyset) = 0$. Die Monotonie (1.23) ist klar, da für $C_1 \subset C_2 \subset \Omega$ jede $\mathcal{G}$-Überdeckung von C_2 insbesondere eine $\mathcal{G}$-Überdeckung von C_1 ist. Die Sub-σ-Additivität (1.24) zeigt man, indem man für $\varepsilon > 0$ beliebig klein zu jedem $C_n \subset \Omega$ eine $\mathcal{G}$-Überdeckung $(A_{n,i})_i$ wählt, die das 'inf' über die möglichen $\mathcal{G}$-Überdeckungen bis auf einen Fehler $< \varepsilon \cdot 2^{-n}$ ausschöpft

$$\mu^*(C_n) \; \leq \; \sum_i \mu(A_{n,i}) \; \leq \; \mu^*(C_n) + \varepsilon \, 2^{-n} \,, \quad n \geq 1$$

und dann alle diese zu einer $\mathcal{G}$-Überdeckung $(A_{n,i})_{n,i}$ für $\bigcup\limits_{n=1}^{\infty} C_n$ zusammenfasst. Für diese gilt

$$\mu^*(\bigcup_n C_n) \; \leq \; \sum_{n,i} \mu(A_{n,i}) \; \leq \; \sum_n \mu^*(C_n) \; + \; 2\varepsilon$$

für $\varepsilon > 0$ beliebig klein. □

In einem nächsten Schritt prüft man nun, ob es Mengen in $\mathcal{P}(\Omega)$ gibt, bezüglich der die generelle Sub-σ-Additivität des äusseren Masses μ^* durch eine ihr entgegenwirkende Zerlegungseigenschaft kompensiert wird. Betrachte dazu

$$\mathcal{H} \; := \; \{A \in \mathcal{P}(\Omega) : \forall \, C \subset \Omega \text{ gilt } \mu^*(C) \geq \mu^*(C \cap A) + \mu^*(C \cap A^c)\} \, .$$

Wegen Sub-σ-Additivität von μ^* auf $\mathcal{P}(\Omega)$ hat man stets "$\leq$", also

(1.26) $$A \in \mathcal{H} \quad \Longleftrightarrow \quad \mu^*(\cdot) = \mu^*(\cdot \cap A) + \mu^*(\cdot \cap A^c) \quad \text{auf } \mathcal{P}(\Omega) \, .$$

1.27 Satz: Das System $\mathcal{H}$ in (1.26) ist eine σ-Algebra auf Ω, welche $\mathcal{G}$ umfasst.
In Einschränkung auf $\mathcal{H}$ ist die Mengenfunktion $\mu^* : \mathcal{H} \to [0, \infty]$ ein Mass.

Beweis: 1) Zuerst zeigen wir $\mathcal{G} \subset \mathcal{H}$. Sei dazu $A \in \mathcal{G}$ beliebig aber fest. Fixiere ein beliebiges $C \subset \Omega$; wähle zu C eine $\mathcal{G}$-Überdeckung $(A_n)_n$ mit

$$\bigcup_n A_n \; \supset C \,, \quad \mu^*(C) \leq \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(C) + \varepsilon \, .$$

Zerlege nun jedes A_n in $A_n \cap A$ und $A_n \cap A^c$. Da $\mathcal{G}$ eine Algebra ist und da μ insbesondere additiv ist auf $\mathcal{G}$, gilt $\mu(A_n) = \mu(A_n \cap A) + \mu(A_n \cap A^c)$ für jedes $n \geq 1$. Nun ist $(A_n \cap A)_n$ eine $\mathcal{G}$-Überdeckung von $C \cap A \subset \Omega$, und $(A_n \cap A^c)_n$ eine $\mathcal{G}$-Überdeckung von $C \cap A^c \subset \Omega$, also nach Definition von μ^*:

$$\begin{aligned}
\mu^*(C) + \varepsilon &\geq \sum_{n=1}^{\infty} \mu(A_n) \\
&= \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \cap A^c) \\
&\geq \mu^*(C \cap A) + \mu^*(C \cap A^c)\ .
\end{aligned}$$

Hierbei war $\varepsilon > 0$ beliebig. Dies zeigt $A \in \mathcal{H}$, nach Definition der Klasse $\mathcal{H}$.

2) Als nächstes zeigen wir: $\mathcal{H}$ ist eine Algebra von Teilmengen von Ω.
Es gilt $\Omega \in \mathcal{H}$ wegen (1.26) mit $A := \Omega$ und $\mu^*(\emptyset) = 0$. Nach Definition ist $\mathcal{H}$ stabil unter Komplementbildung, denn (1.26) ist symmetrisch in A und A^c. Prüfe jetzt die $\cup$-Stabilität von $\mathcal{H}$. Seien $A, B \in \mathcal{H}$, sei $C \subset \Omega$ beliebig, dann

$$\begin{aligned}
\mu^*(C) &= \mu^*(C \cap A) + \mu^*(C \cap A^c) \qquad (\text{da } A \in \mathcal{H}) \\
&= \mu^*(C \cap A \cap B) + \mu^*(C \cap A \cap B^c) + \\
&\qquad + \mu^*(C \cap A^c \cap B) + \mu^*(C \cap A^c \cap B^c) \qquad (\text{da } B \in \mathcal{H})\ .
\end{aligned}$$

Ersetzt man hier nun C durch $\widetilde{C} := C \cap (A \cup B) = (C \cap A) \cup (C \cap B)$, so ergibt sich für $\widetilde{C}$

$$\mu^*(C \cap (A \cup B)) = \mu^*(C \cap A \cap B) + \mu^*(C \cap A \cap B^c) + \mu^*(C \cap A^c \cap B) + 0 \tag{1.28}$$

wegen $\mu^*(\emptyset) = 0$. Die Formel (1.28) vereinfacht die rechte Seite der vorletzten Gleichung zu

$$\mu^*(C) = \mu^*(C \cap (A \cup B)) + \mu^*(C \cap A^c \cap B^c) = \mu^*(C \cap (A \cup B)) + \mu^*(C \cap (A \cup B)^c)\ .$$

Dies zeigt $A \cup B \in \mathcal{H}$, also ist $\mathcal{H}$ eine Algebra.

3) Wir zeigen, dass $\mathcal{H}$ ein Dynkin-System ist. Da $\mathcal{H}$ nach Schritt 2) insbesondere als Algebra $\cap$-stabil ist, muss $\mathcal{H}$ dann nach 1.9 sogar eine σ-Algebra sein: dies beweist die erste Aussage des Satzes.
Seien A_n, $n \geq 1$, in $\mathcal{H}$ paarweise disjunkt, schreibe $\dot{\bigcup}_{n=1}^{\infty} A_n =: A \subset \Omega$; zu zeigen ist: $A \in \mathcal{H}$. Die Formel (1.28) vereinfacht sich für *disjunkte* Vereinigungen zu

$$\mu^*(C \cap (A \mathbin{\dot\cup} B)) = \mu^*(C \cap A) + \mu^*(C \cap B)\ , \quad C \subset \Omega \text{ beliebig}\ ;$$

damit gilt für $A_1, \ldots, A_n$

$$\mu^*\left(C \cap (\dot{\bigcup_{i=1}^{n}} A_i)\right) = \sum_{i=1}^{n} \mu^*(C \cap A_i), \quad C \subset \Omega \text{ beliebig}\ . \tag{1.29}$$

Da $(\dot{\bigcup}_{i=1}^{n} A_i) \in \mathcal{H}$ nach Schritt 2), gilt für $C \subset \Omega$ beliebig

$$\begin{aligned}
\mu^*(C) &= \mu^*\left(C \cap (\dot{\bigcup_{i=1}^{n}} A_i)\right) + \mu^*\left(C \cap (\dot{\bigcup_{i=1}^{n}} A_i)^c\right) \qquad \text{(nach (1.26))} \\
&\geq \mu^*\left(C \cap (\dot{\bigcup_{i=1}^{n}} A_i)\right) + \mu^*(C \cap A^c) \qquad \text{(Definition von } A \subset \Omega\text{, Monotonie von } \mu^*\text{)} \\
&= \sum_{i=1}^{n} \mu^*(C \cap A_i) + \mu^*(C \cap A^c)
\end{aligned}$$

nach (1.29). In der letzten Zeile ist $n \in I\!N$ beliebig. Also hat man auch

$$(+) \qquad \mu^*(C) \;\geq\; \sum_{i=1}^{\infty} \mu^*(C \cap A_i) + \mu^*(C \cap A^c)\;;$$

aber Sub-σ-Additivität (1.24) zeigt

$$(++) \qquad \mu^*(C) \;\leq\; \mu^*(C \cap A) + \mu^*(C \cap A^c) \;\leq\; \sum_{i=1}^{\infty} \mu^*(C \cap A_i) + \mu^*(C \cap A^c).$$

Also hat man überall "=" in (+) und (++), und diese Gleichungskette zeigt insbesondere

$$\mu^*(C) \;=\; \mu^*(C \cap A) + \mu^*(C \cap A^c) \quad \forall\, C \in \mathcal{P}(\Omega)\,.$$

Damit aber ist $A = \dot{\bigcup}_{i=1}^{\infty} A_i \in \mathcal{H}$, nach Definition der Klasse $\mathcal{H}$ in (1.26). Somit ist $\mathcal{H}$ Dynkin.

4) Wir beweisen die letzte Behauptung des Satzes. Seien A_i, $i \geq 1$, paarweise disjunkt in $\mathcal{H}$. Im Spezialfall $C := A = \dot{\bigcup}_{i=1}^{\infty} A_i \in \mathcal{H}$ – da $\mathcal{H}$ Dynkin – zeigt die in (+) und (++) entstandene Gleichungskette auch

$$\mu^*\left(\dot{\bigcup_{i=1}^{\infty}} A_i\right) \;=\; \sum_{i=1}^{\infty} \mu^*(A_i)$$

und damit die σ-Additivität der Mengenfunktion μ^* auf $\mathcal{H}$. Also entsteht durch Einschränkung des äusseren Masses μ^* auf die σ-Algebra $\mathcal{H}$ ein Mass. $\square$

Der Beweis des Fortsetzungssatzes ist nun so gut wie fertig:

1.30 Beweis von 1.18 : Wie in Voraussetzung (V) sei μ ein Prämass auf einer Algebra $\mathcal{G}$. Hilfssatz 1.25 liefert ein äusseres Mass $\mu^* : \mathcal{P}(\Omega) \to [0,\infty]$, welches in Einschränkung auf $\mathcal{G}$ mit μ übereinstimmt. Hilfssatz 1.27 liefert eine σ-Algebra $\mathcal{H}$ auf Ω mit $\mathcal{H} \supset \mathcal{G}$, so dass μ^* in Einschränkung auf $\mathcal{H}$ ein Mass ist. Aus $\mathcal{G} \subset \mathcal{H}$ folgt $\sigma(\mathcal{G}) \subset \mathcal{H}$; weitere Einschränkung

$$\mu^* : \sigma(\mathcal{G}) \to [0,\infty]$$

liefert ein Mass auf $\sigma(\mathcal{G})$ mit $\mu^*(A) = \mu(A)$ für alle $A \in \mathcal{G}$. $\square$

Beachte, dass mit dem gegebenen Beweisgang nichts über Eindeutigkeit einer solchen Fortsetzung gesagt werden kann. In Kombination mit dem Eindeutigkeitssatz 1.13 aber sieht man:

1.31 Korollar: Jedes σ-*endliche* Prämass μ auf einer Algebra $\mathcal{G}$ kann auf *genau eine* Weise zu einem Mass $\widetilde{\mu}$ auf $\sigma(\mathcal{G})$ fortgesetzt werden.

Beweis: Unter Voraussetzung (V) seien $\widetilde{\mu}_1$, $\widetilde{\mu}_2$ Masse auf $\sigma(\mathcal{G})$, welche μ gemäss 1.18 fortsetzen:

$$(+) \qquad \widetilde{\mu}_1(A) = \widetilde{\mu}_2(A) = \mu(A) \quad \forall\, A \in \mathcal{G}.$$

Als Algebra ist $\mathcal{G}$ ein $\cap$-stabiler Erzeuger von $\sigma(\mathcal{G})$. Weiter wird σ-Endlichkeit des Inhaltes μ auf $\mathcal{G}$ vorausgesetzt, also existiert eine Folge $(E_n)_n \subset \mathcal{G}$ mit

$$E_n \uparrow \Omega\ , \quad \mu(E_n) < \infty \quad \text{für jedes } n \in I\!N\ .$$

Nun sind beide Voraussetzungen von 1.13 erfüllt, und $\widetilde{\mu}_1 = \widetilde{\mu}_2$ auf $\sigma(\mathcal{G})$ folgt aus (+) mit 1.13. $\square$

E. Messbare Abbildungen

1.32 Definition: Seien $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ messbare Räume, sei $T : \Omega_1 \to \Omega_2$ eine Abbildung. T heisst $\mathcal{A}_1$–$\mathcal{A}_2$-*messbar*, falls das Urbild von $\mathcal{A}_2$ unter T (siehe 1.4) in $\mathcal{A}_1$ enthalten ist:

$$\forall\, A_2 \in \mathcal{A}_2 : \quad T^{-1}(A_2) = \{\omega \in \Omega_1 : T(\omega) \in A_2\} \ \in\ \mathcal{A}_1\ .$$

Die Messbarkeit von T drückt man auch durch die Schreibweise $T : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ aus.

1.33 Satz: Betrachte messbare Räume $(\Omega_i, \mathcal{A}_i)$, $i = 1, 2, 3$, und messbare Abbildungen

$$T_1 : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)\ , \quad T_2 : (\Omega_2, \mathcal{A}_2) \to (\Omega_3, \mathcal{A}_3)\ .$$

Dann gilt für die Komposition $T_2 \circ T_1$:

$$T_2 \circ T_1 : (\Omega_1, \mathcal{A}_1) \to (\Omega_3, \mathcal{A}_3) \quad \text{ist messbar}\ .$$

Beweis: Für alle $A_3 \in \mathcal{A}_3$ gilt $(T_2 \circ T_1)^{-1}(A_3) = T_1^{-1}(\underbrace{T_2^{-1}(A_3)}_{\in \mathcal{A}_2}) \in \mathcal{A}_1$. $\square$

Zum Nachweis der Messbarkeit einer Abbildung kann man sich auf Erzeuger der σ-Algebra auf dem Bildraum zurückziehen:

1.34 Satz: Seien $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ messbare Räume, sei $T : \Omega_1 \to \Omega_2$ eine Abbildung. Gilt für einen Erzeuger $\mathcal{E}_2$ der σ-Algebra $\mathcal{A}_2$

$$T^{-1}(C) \in \mathcal{A}_1 \quad \forall\, C \in \mathcal{E}_2\ ,$$

so ist T $\mathcal{A}_1$–$\mathcal{A}_2$-messbar.

Beweis: Schreibe $\mathcal{H} := \{C \subset \Omega_2 : T^{-1}(C) \in \mathcal{A}_1\}$. Es ist leicht zu sehen, dass $\mathcal{H}$ eine σ-Algebra in Ω_2 ist. Gilt also $\mathcal{E}_2 \subset \mathcal{H}$ für ein System $\mathcal{E}_2$ von Teilmengen von Ω_2, so folgt $\sigma(\mathcal{E}_2) \subset \mathcal{H}$. □

1.35 Beispiel: Für $i = 1, 2$ sei $\mathcal{O}_i$ eine Topologie auf einem Raum E_i, und sei E_i versehen mit seiner Borelschen σ-Algebra $\mathcal{B}(E_i)$. Jede stetige Abbildung $f : E_1 \to E_2$ ist dann $\mathcal{B}(E_1)$–$\mathcal{B}(E_2)$-messbar.

Beweis: Nach Definition der Stetigkeit in topologischen Räumen ist $f : E_1 \to E_2$ genau dann stetig, wenn Urbilder offener Mengen offen sind:

$$f^{-1}(O_2) \in \mathcal{O}_1 \quad \forall\, O_2 \in \mathcal{O}_2\ .$$

$\mathcal{O}_2$ erzeugt die σ-Algebra $\mathcal{B}(E_2)$, vgl. 1.6. Versieht man den Raum E_1 mit einer σ-Algebra $\mathcal{A}_1$, welche $\mathcal{O}_1$ umfasst, wird nach 1.34 jede stetige Abbildung $f : E_1 \to E_2$ zu einer messbaren Abbildung $f : (E_1, \mathcal{A}_1) \to (E_2, \mathcal{B}(E_2))$. Die kleinste solche σ-Algebra ist $\sigma(\mathcal{O}_1) = \mathcal{B}(E_1)$. □

Ist man noch frei, auf dem Grundraum Ω eine σ-Algebra zu wählen, so kann man jede Familie von Abbildungen mit Werten in messbaren Räumen mit Hilfe von Satz 1.5 messbar *machen*:

1.36 Bemerkung: Sei I eine beliebige Indexmenge, seien $(E_i, \mathcal{E}_i)$ messbare Räume, $i \in I$. Sei auf einem festen Grundraum Ω eine Familie von Abbildungen $T_i : \Omega \to E_i$ gegeben.

a) Für jedes $i \in I$ existiert eine kleinste σ-Algebra $\mathcal{A}$ auf Ω, welche T_i $\mathcal{A}$–$\mathcal{E}_i$-messbar macht, nämlich

$$\mathcal{A} \;=\; T_i^{-1}(\mathcal{E}_i) \;=\; \{T_i^{-1}(B) \;:\; B \in \mathcal{E}_i\}\ ;$$

nenne $\mathcal{A} =: \sigma(T_i)$ die von T_i erzeugte σ-Algebra.

b) Stets existiert eine kleinste σ-Algebra $\mathcal{A}'$ auf Ω, welche *alle* Abbildungen T_i $\mathcal{A}$–$\mathcal{E}_i$-messbar macht, $i \in I$, nämlich

$$\mathcal{A}' \;=\; \sigma\left(\bigcup_{i \in I} T_i^{-1}(\mathcal{E}_i)\right)\ ;$$

nenne diese die von den T_i, $i \in I$, erzeugte σ-Algebra und schreibe $\mathcal{A}' =: \sigma(T_i \ : \ i \in I)$.

Eine wichtige Eigenschaft messbarer Abbildungen besteht darin, dass sie Masse 'transportieren':

1.37 Satz: Sei $T : (\Omega_1, \mathcal{A}_1) \to (\Omega_2, \mathcal{A}_2)$ messbar; sei μ_1 ein Mass auf $(\Omega_1, \mathcal{A}_1)$. Dann definiert

$$\mathcal{A}_2 \ni \ A_2 \ \longrightarrow \ \mu_2(A_2) := \mu_1(T^{-1}(A_2)) = \mu_1(\{T \in A_2\}) \ \in \ [0, \infty]$$

ein Mass μ_2 auf $(\Omega_2, \mathcal{A}_2)$: nenne μ_2 *Bildmass zu* μ_1 *unter* T oder *Verteilung von* T *unter* μ_1; übliche Schreibweisen für das Bildmass sind μ_1^T oder $\mu_1 \circ T^{-1}$ oder $\mathcal{L}(T \mid \mu_1)$.

Beweis: μ_2 wie oben definiert ist ein Mass auf $\mathcal{A}_2$: Zuerst $\mu_2(\emptyset) = \mu_1(\emptyset) = 0$. Seien $A_{2,n}, n \geq 1$, paarweise disjunkt in $\mathcal{A}_2$. Dann sind auch die Urbilder $A_{1,n} := T^{-1}(A_{2,n})$ paarweise disjunkt in $\mathcal{A}_1$, $n \geq 1$, folglich

$$\mu_2\left(\dot{\bigcup_n} A_{2,n}\right) = \mu_1\left(T^{-1}(\dot{\bigcup_n} A_{2,n})\right) = \mu_1\left(\dot{\bigcup_n} T^{-1}(A_{2,n})\right) = \sum_{n=1}^{\infty} \mu_1(T^{-1}(A_{2,n})) = \sum_{n=1}^{\infty} \mu_2(A_{2,n}).$$

Damit ist μ_2 σ-additiv auf $\mathcal{A}_2$. □

Nun betrachten wir genauer messbare Abbildungen $(\Omega, \mathcal{A}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$; zuerst für $d = 1$:

1.38 Satz: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum, sei $f : \Omega \to \mathbb{R}$ eine Abbildung. Jede der folgenden Bedingungen i) - v) ist gleichwertig mit $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$–Messbarkeit von f:

i) $\{f > a\} \in \mathcal{A} \quad \forall\, a \in \mathbb{R}$;

ii) $\{f \leq b\} \in \mathcal{A} \quad \forall\, b \in \mathbb{R}$;

iii) $\{a < f \leq b\} \in \mathcal{A} \quad \forall\, -\infty < a < b < +\infty$;

iv) $\{f \geq a\} \in \mathcal{A} \quad \forall\, a \in \mathbb{R}$;

v) $\{f < b\} \in \mathcal{A} \quad \forall\, b \in \mathbb{R}$;

mit $\{f > a\} := \{\omega \in \Omega \ : \ f(\omega) > a\} = f^{-1}((a, \infty))$, und analogen Schreibweisen in ii)–v).

Beweis: Jede der Bedingungen i)–v) ist hinreichend für $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$–Messbarkeit von f: dazu wendet man 1.34 auf geeignete Erzeugendensysteme von $\mathcal{B}(\mathbb{R})$ an. Für die Aussage ii) zum Beispiel benutzt man wegen $\{f \leq b\} = f^{-1}((-\infty, b])$ das System $\mathcal{C}_1$ aus 1.6'. Jede der Bedingungen i)–v) ist nach Definition 1.32 auch notwendig: ist eine Abbildung $f : \Omega \to \mathbb{R}$ $\mathcal{A}$–$\mathcal{B}(\mathbb{R})$–messbar, so gehört etwa $\{a < f \leq b\}$ als Urbild von $(a, b] \in \mathcal{B}(\mathbb{R})$ zur σ-Algebra $\mathcal{A}$. □

1.38' Beispiel: Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Alle $\mathcal{A}$-*Elementarfunktionen* $\varrho : \Omega \to I\!R$

$$\varrho \;=\; \sum_{i=1}^{\ell} \alpha_i 1_{A_i} \;, \quad \alpha_i \in I\!R \;,\; A_i \in \mathcal{A} \;,\; \ell \in I\!N$$

sind $\mathcal{A}$–$\mathcal{B}(I\!R)$–messbar.

Beweis: Stets gibt es eine Darstellung von $\varrho : \Omega \to I\!R$ der Form

$$\varrho \;=\; \sum_{j=1}^{\ell'} \alpha'_j 1_{A'_j} \;, \quad \alpha'_j \in I\!R, \quad A'_j,\; 1 \le j \le \ell',\; \text{paarweise disjunkt in } \mathcal{A}\,;$$

mit dieser hat man für beliebiges $b \in I\!R$

$$\{\varrho \le b\} \;=\; \dot{\bigcup_{\substack{j \in \{1,\ldots,\ell'\} \\ \alpha'_j \le b}}} A'_j \quad \in \mathcal{A}\,.$$

Nach 1.38 ii) ist damit $\rho : (\Omega, \mathcal{A}) \to (I\!R, \mathcal{B}(I\!R))$ messbar. $\square$

Man hätte 1.38 gleich in beliebiger Dimension $d \ge 1$ formulieren können; viel nützlicher aber ist

1.39 Satz: Genau dann ist

$$f = (f_1, \ldots, f_d) \;:\; \Omega \;\longrightarrow\; I\!R^d$$

$\mathcal{A}$–$\mathcal{B}(I\!R^d)$–messbar, wenn alle Komponenten f_i $\mathcal{A}$–$\mathcal{B}(I\!R)$–messbar sind, $1 \le i \le d$.

Beweis: Für alle $b = (b_1, \ldots, b_d) \in I\!R^d$ gilt mit komponentenweise definiertem '$\le$'

$$\{f_i \le b_i\} \;=\; \bigcup_n \{f \le (\underbrace{n, \ldots, n}_{\text{i}-1\text{ mal}}, b_i, \underbrace{n, \ldots, n}_{\text{d}-\text{i mal}})\} \quad , \quad \{f \le b\} \;=\; \bigcap_{i=1}^{d} \{f_i \le b_i\}\,.$$

Mit 1.38 ii) – angewandt in $(I\!R^d, \mathcal{B}(I\!R^d))$ und in $(I\!R, \mathcal{B}(I\!R))$ – folgt die Aussage. $\square$

1.40 Bemerkung: Die reellwertigen *Zufallsvariablen* X der Stochastik-Einführungsvorlesungen auf einem – häufig nicht explizit genannten – Wahrscheinlichkeitraum $(\Omega, \mathcal{A}, P)$ waren nichts anderes als *messbare Abbildungen* $(\Omega, \mathcal{A}) \to (I\!R, \mathcal{B}(I\!R))$; sonst hätte man z.B. nie einem Ereignis $\{X \le b\}$ eine Wahrscheinlichkeit $P(X \le b)$ zusprechen dürfen. Wichtig ist (in statistischen Problemen ist das die Standardsituation), dass auf demselben Grundraum $(\Omega, \mathcal{A})$ durchaus *verschiedene* Wahrscheinlichkeitsmasse $P, P', \ldots$ spielen können: folglich muss 'Zufallsvariable' notwendig ein Begriff 'ohne Wahrscheinlichkeit' sein, genau wie es in 1.32–36 oder in 1.38–1.39 angesetzt wurde; danach kann man in einem zweiten Schritt nach (Verteilungs-) Eigenschaften einer gegebenen Zufallsvariable unter Wahrscheinlichkeitsmassen $P, P', \ldots$ fragen, siehe 1.37.

Für einen beliebigen messbaren Raum $(E,\mathcal{E})$ sprechen wir im folgenden unterschiedslos von 'Zufallsvariablen auf $(\Omega,\mathcal{A})$ mit Werten in $(E,\mathcal{E})$' und von 'messbaren Abbildungen $(\Omega,\mathcal{A}) \to (E,\mathcal{E})$'.

1.41 Hilfssatz: Betrachte messbare Funktionen $f_n : (\Omega,\mathcal{A}) \to (I\!R, \mathcal{B}(I\!R))$, $n \geq 1$.

o) Für jedes $A \in \mathcal{A}$ ist die Abbildung $f_1 1_A$ $\mathcal{A}$-$\mathcal{B}(I\!R)$–messbar.

i) Unter der Voraussetzung $\sup\limits_{n\geq 1} f_n < \infty$ gilt

$$f \;:=\; \sup_{n\geq 1} f_n \quad \text{ist } \mathcal{A}\text{–}\mathcal{B}(I\!R)\text{–messbar} \;;$$

unter der Voraussetzung $\inf\limits_{n\geq 1} f_n > -\infty$ gilt

$$g \;:=\; \inf_{n\geq 1} f_n \quad \text{ist } \mathcal{A}\text{–}\mathcal{B}(I\!R)\text{–messbar} \,.$$

ii) Stets gilt

$$\{f_1 > f_2\} \in \mathcal{A} \,, \quad \{f_1 = f_2\} \in \mathcal{A} \,, \quad \{f_1 \geq f_2\} \in \mathcal{A} \,.$$

iii) Mit f_1, f_2 sind auch die folgenden Funktionen $\mathcal{A}$–$\mathcal{B}(I\!R)$–messbar:

$$f_1 \pm f_2, \quad f_1 \cdot f_2, \quad \text{sowie } f_1/f_2 \text{ falls } f_2 \neq 0$$

$$f_1 \vee f_2 := \max\{f_1, f_2\} \,, \quad f_1 \wedge f_2 := \min\{f_1, f_2\} \,.$$

Beweis: Wir geben eine Skizze, alle Details überlege man sich als Übungsaufgabe.

o) Schreibe $\{f_1 1_A \leq b\}$ als $\{f_1 \leq b\} \cap A$ falls $b < 0$, und als $A^c \,\dot\cup\, (\{f_1 \leq b\} \cap A)$ falls $b \geq 0$.

i) Für beliebiges $a \in I\!R$ gilt $\{f > a\} = \bigcup\limits_{n\geq 1} \{f_n > a\} \in \mathcal{A}$; analog sieht man $\{g < a\} \in \mathcal{A}$.

ii) Zuerst sieht man

$$\{f_1 > f_2\} \;=\; \bigcup_{r\in \mathbb{Q}} (\{f_1 > r\} \cap \{r > f_2\}) \;\in \mathcal{A}$$

und schreibt danach $\{f_1 \geq f_2\}$ in Form $\bigcap\limits_{n\geq 1} \{f_1 > f_2 - \frac{1}{n}\}$.

iii) Man schreibt für $a \in I\!R$ Mengen $\{f_1 + f_2 > a\}$ in Form $\{f_1 > \widetilde{f}\}$, $\widetilde{f} := a - f_2$. Dabei ist $\widetilde{f} : (\Omega,\mathcal{A}) \to (I\!R, \mathcal{B}(I\!R))$ messbar, wegen $\{\widetilde{f} \geq b\} = \{f_2 \leq a - b\}$, und ii) liefert $\{f_1 > \widetilde{f}\} \in \mathcal{A}$. Gezeigt ist $\{f_1 + f_2 > a\} \in \mathcal{A}$ für beliebiges $a \in I\!R$, und damit die $\mathcal{A}$–$\mathcal{B}(I\!R)$–Messbarkeit von $f_1 + f_2$. Mit $g : (\Omega,\mathcal{A}) \to (I\!R, \mathcal{B}(I\!R))$ ist g^2 messbar, folglich auch

$$f_1 \cdot f_2 \;=\; \frac{1}{4}\,(f_1 + f_2)^2 - \frac{1}{4}\,(f_1 - f_2)^2 \;.$$

Mit $F := \{f_1 \geq f_2\} \in \mathcal{A}$ schreibt man schliesslich $f_1 \vee f_2 = f_1\, 1_F + f_2\, 1_{F^c}$ und benutzt o). $\square$

Die Aussagen in 1.41 i) gehen nicht weit genug: als Limesobjekte für auf- oder absteigende Folgen reellwertiger Funktionen muss man $\overline{\mathbb{R}}$-wertige Zufallsvariable betrachten können.

1.42 Bemerkung: Versehe $\overline{\mathbb{R}}$ mit der kleinsten σ-Algebra, welche alle in $\overline{\mathbb{R}}$ offenen Mengen enthält (zusätzlich zu $\mathcal{O}$ braucht man offene Umgebungen $(a,+\infty]$ von $+\infty$ und $[-\infty,b)$ von $-\infty$, $a,b\in\mathbb{R}$):

$$\mathcal{B}(\overline{\mathbb{R}}) := \sigma\left(\mathcal{B}(\mathbb{R})\,,\,\{+\infty\}\,,\,\{-\infty\}\right)\ .$$

a) Es gilt analog zu 1.38 (Details als Übungsaufgabe)

$$\begin{aligned} f:(\Omega,\mathcal{A})\to(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))\quad \text{ist messbar} \quad &\iff \quad \{f>a\}\in\mathcal{A}\quad \forall\, a\in\mathbb{R}\\ &\iff \quad \{f\ge a\}\in\mathcal{A}\quad \forall\, a\in\mathbb{R}\\ &\iff \quad \{f<b\}\in\mathcal{A}\quad \forall\, b\in\mathbb{R}\\ &\iff \quad \{f\le b\}\in\mathcal{A}\quad \forall\, b\in\mathbb{R}\,. \end{aligned}$$

Die Eigenschaft iii) aus 1.38 dagegen ist jetzt nicht mehr hinreichend (warum?) für Messbarkeit einer Funktion $f:(\Omega,\mathcal{A})\to(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$.

b) Für beliebige Folgen messbarer Funktionen $f_n:(\Omega,\mathcal{A})\to(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$, $n\ge 1$, hat man nun ohne jede Zusatzvoraussetzung

$$\begin{aligned} f &:= \sup_{n\ge1} f_n \quad \text{ist } \mathcal{A}\text{–}\mathcal{B}(\overline{\mathbb{R}})\text{–messbar}\\ g &:= \inf_{n\ge1} f_n \quad \text{ist } \mathcal{A}\text{–}\mathcal{B}(\overline{\mathbb{R}})\text{–messbar}\,. \end{aligned}$$

1.41 ii) und iii) gelten analog für $f_i:(\Omega,\mathcal{A})\to(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$ messbar, $i=1,2$; zur Verallgemeinerung der ersten Zeile von 1.41 iii) muss man einschränkend voraussetzen, dass $f_1\pm f_2$, $f_1\cdot f_2$, f_1/f_2 auf Ω wohldefiniert sind. $\square$

Wir schliessen das Kapitel mit dem wichtigen

1.43 Faktorisierungslemma: Sei $(E,\mathcal{E})$ ein messbarer Raum, sei $T:\Omega\to E$ eine Abbildung. Versehe Ω mit der von T erzeugten σ-Algebra $\sigma(T)=T^{-1}(\mathcal{E})$. Für beliebige Abbildungen $f:\Omega\to\overline{\mathbb{R}}$ sind dann gleichwertig:

i) f ist $\sigma(T)$-$\mathcal{B}(\overline{\mathbb{R}})$-messbar.

ii) Es existiert eine messbare Abbildung $h:(E,\mathcal{E})\to(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$ so dass $f=h\circ T$.

Ist f insbesondere reellwertig, so kann auch h reellwertig gewählt werden: die Aussage gilt dann mit $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ anstelle von $(\overline{\mathbb{R}},\mathcal{B}(\overline{\mathbb{R}}))$.

Beweis: ii) $\Longrightarrow$ i) gilt nach 1.33; wir zeigen i) $\Longrightarrow$ ii).

1) Betrachte zuerst die Klasse der $\sigma(T)$-Elementarfunktionen

$$f \;=\; \sum_{i=1}^{\ell} \alpha_i 1_{A_i} \;, \quad A_i \in \sigma(T)\;,\; \alpha_i \in I\!R\;,\; \ell \geq 1\;.$$

Für $A_i \in \sigma(T) = T^{-1}(\mathcal{E})$ wähle ein $B_i \in \mathcal{E}$ so dass $A_i = T^{-1}(B_i) = \{T \in B_i\}$. Dann ist $1_{A_i} = 1_{B_i} \circ T$, also ist

$$h \;=\; \sum_{i=1}^{\ell} \alpha_i 1_{B_i}$$

eine $\mathcal{E}$-Elementarfunktion mit der Eigenschaft $f = h \circ T$.

2) Betrachte nun $f : (\Omega, \sigma(T)) \to (\overline{I\!R}, \mathcal{B}(\overline{I\!R}))$ messbar und *nichtnegativ*. Dann ist $(f_n)_n$

$$f_n(\omega) \;:=\; \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\, 1_{\{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\}}(\omega) \;+\; n 1_{\{f \geq n\}}(\omega)\;, \quad n \geq 1\;,\; \omega \in \Omega$$

eine Schar von $\sigma(T)$-Elementarfunktionen, welche f monoton aufsteigend approximiert:

$$(+) \qquad f_n(\omega) \leq f_{n+1}(\omega)\;,\; n \geq 1\;, \quad f(\omega) \;=\; \sup_{n\geq 1} f_n(\omega)\;, \quad \omega \in \Omega\;.$$

Schreibe nun kurz für $f : \Omega \to [0,\infty]$ $\sigma(T)$-messbar

$$A_{n,k} := \{\frac{k}{2^n} \leq f < \frac{k+1}{2^n}\} \;\in\; \sigma(T)\;, \quad \widetilde{A}_n := \{f \geq n\} \;\in\; \sigma(T)$$

und wähle $B_{n,k}$, $\widetilde{B}_n$ in $\mathcal{E}$ so dass $A_{n,k} = T^{-1}(B_{n,k})$, $\widetilde{A}_n = T^{-1}(\widetilde{B}_n)$. Wie in 1) ist dann

$$\widetilde{h}_n := \sum_{k=0}^{n2^n-1} \frac{k}{2^n}\, 1_{B_{n,k}} \;+\; n 1_{\widetilde{B}_n}\;, \quad n \geq 1$$

eine Schar nichtnegativer $\mathcal{E}$-Elementarfunktionen mit der Eigenschaft $f_n = \widetilde{h}_n \circ T$. Die Folge $(\widetilde{h}_n)_n$ ist i.a. nicht auf ganz E aufsteigend. Durch (+) wird dies nur auf dem Wertebereich von T sichergestellt. Ausserhalb der Menge $T(\Omega) \subset E$ können die Funktionen $\widetilde{h}_n$ beliebig umdefiniert werden, sofern die $\mathcal{E}$–$\mathcal{B}(I\!R)$–Messbarkeit erhalten bleibt: wir ersetzen also $\widetilde{h}_n$ durch $h_n := \max\{\widetilde{h}_1, ..., \widetilde{h}_n\}$. So ist für die aufsteigende Folge $(f_n)_n$ von $\sigma(T)$-Elementarfunktionen eine aufsteigende Folge $(h_n)_n$ von $\mathcal{E}$-Elementarfunktionen gefunden mit $f_n = h_n \circ T$ für alle n. Nach 1.41+1.42 leistet

$$h := \sup_n h_n \;: \quad (E, \mathcal{E}) \to (\overline{I\!R}, \mathcal{B}(\overline{I\!R}))\ \text{messbar}$$

das Gewünschte: es gilt $f = h \circ T$.

3) Zerlege nun eine beliebige messbare Funktion $f : (\Omega, \sigma(T)) \to (\overline{I\!R}, \mathcal{B}(\overline{I\!R}))$ in ihren Positivteil und ihren Negativteil: definiere für $\omega \in \Omega$

$$f^+(\omega) \;:=\; f(\omega) \vee 0 \;=\; \left\{ \begin{array}{ll} f(\omega) & \text{falls } f(\omega) \in (0,\infty] \\ 0 & \text{sonst} \end{array} \right\}\;,$$

$$f^-(\omega) \;:=\; -(f(\omega) \wedge 0) \;=\; \left\{ \begin{array}{ll} -f(\omega) & \text{falls } f(\omega) \in [-\infty,0) \\ 0 & \text{sonst} \end{array} \right\}\;.$$

Dann sind f^+, f^- nichtnegativ und $\sigma(T)$-$\mathcal{B}(\overline{I\!R})$-messbar (nach 1.42), und es gilt

$$f \;=\; f^+ - f^- \text{ auf } \Omega.$$

Schritt 2) liefert messbare Funktionen $h_1, h_2 : (E, \mathcal{E}) \to (\overline{I\!R}, \mathcal{B}(\overline{I\!R}))$ so dass

$$(++) \qquad f^+ = h_1 \circ T, \quad f^- = h_2 \circ T, \quad h_1, h_2 \geq 0\ .$$

Beachte: auf $N := \{h_1 = +\infty\} \cap \{h_2 = +\infty\} \in \mathcal{E}$ ist eine Differenz $h_1 - h_2$ nicht definiert. Aber N ist immer disjunkt zum Wertebereich von T (aus $T(\omega) \in N$ folgt der Widerspruch $\omega \in \{f^+ = +\infty\} \cap \{f^- = +\infty\} = \emptyset$). Folglich können h_1 und h_2 ohne Beeinträchtigung von $(++)$ durch die $\mathcal{E}$-messbaren Funktionen $h_1\, 1_{N^c}$ und $h_2\, 1_{N^c}$ ersetzt werden:

$$f \;=\; h \circ T \quad \text{mit} \quad h \;:=\; h_1\, 1_{N^c} - h_2\, 1_{N^c}\ : (E, \mathcal{E}) \to (\overline{I\!R}, \mathcal{B}(\overline{I\!R})) \text{ messbar}\ .$$

4) Die Äquivalenz der Behauptungen i) und ii) ist nun bewiesen. Ist f insbesondere reellwertig, so kann man jede Funktion h mit $f = h \circ T$ ohne Beeinträchtigung dieser Eigenschaft durch $h\, 1_{\{-\infty < h < +\infty\}}$ ersetzen. Dies zeigt die Zusatzbemerkung. $\Box$