Vorlesung

Stochastische Prozesse

Wintersemester 90/91

Prof. Dr. Manfred Schäl

Ausarbeitung von Michael Schmitz

<u>Inhaltsverzeichnis</u>

Kapitel 1: Zeitdiskrete stochastische Prozesse

Kapitel 2: Zeitstetige stochastische Prozesse

Literatur:

Ash, Robert: Real Analysis and Probability, Academic Press 1972 New York

Billingsley: Convergence of Probability Measures, Wiley New York 1968

Karatzas & Shreve: Brownian Motion and Stochastic Calculus, Springer New York 1988

von Querenburg, Boto: Mengentheoretische Topologie, Springer Berlin 1973

Die im Text angegebenen Verweise auf die Wahrscheinlichkeitstheorie-Vorlesung beziehen sich auf die von Prof. Schäl gehaltene Vorlesung im Wintersemester 89/90.

Bonn im Sommer 1992

Kapitel 1: Zeitdiskrete stochastische Prozesse

Zunächst die Definition dieses Begriffes:

Definition: Ein *stochastischer Prozeß* ist eine Familie $(X_t)_{t\in I}$ von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in einem Meßraum $(S, \mathcal{S})$. Dabei ist t der Zeitparameter, I die Parametermenge und S der Zustandsraum. Ist $I \subset \mathbb{Z}$ (z.B. $I = \mathbb{N}_0$), so spricht man von einem *zeitdiskreten Prozeß*, für $I \subset \mathbb{R}$ = Intervall wird $(X_t)_{t\in I}$ als *zeitstetiger Prozeß* bezeichnet. Die Abbildung $[\,t \mapsto X_t(\omega)\,]$ heißt *Pfad*.

Anschaulich gesehen modelliert $(X_t)_{t\in I}$ ein zufallsabhängiges Geschehen. X_t gibt den Zustand des beschriebenen Systems zum Zeitpunkt t wieder.

Als Beispiele aus der Praxis kann man anführen:

Stauseepegel: $X_t \mathrel{\hat{=}}$ Wasserstand zum Zeitpunkt t

Erdbebenmessung: $X_t \mathrel{\hat{=}}$ Ausschlag des Zeigers der Richter-Skala zum Zeitpunkt t

Molekülbewegung: $X_t \mathrel{\hat{=}}$ Ort eines Moleküls zum Zeitpunkt t

Für die nun folgenden Abschnitte benötigt man den Maßfortsetzungssatz, daher, um die genauen Voraussetzungen/Zusammenhänge zu kennen:

Abschnitt 0: Bemerkungen zum Maßfortsetzungssatz

Ein Ziel der Mathematik ist es, Teilmengen von Räumen (z.B. des $\mathbb{R}^n$) Maßzahlen zuzuordnen. Diese sollen möglichst mit den Kennzahlen der Elementargeometrie (wie Länge im $\mathbb{R}^1$, Fläche im $\mathbb{R}^2$ oder Volumen im $\mathbb{R}^3$) übereinstimmen. Dabei sollen folgende Rechenregeln gelten:

(1) Sind A, B zwei zueinander kongruente Mengen, so sollen sie die gleiche Maßzahl zugeordnet bekommen.

(2) Sind A, B zwei disjunkte Mengen mit Maßzahlen a bzw. b, so soll die Menge $A \cup B$ die Maßzahl $a + b$ besitzen.

Mit diesen beiden einleuchtenden Rechenregeln stößt man jedoch in der Elementargeometrie schnell an Grenzen, wenn man z.B. die Fläche einer offenen Kreisscheibe K im $\mathbb{R}^2$ bestimmen will. Dies geschieht durch Ausschöpfung von K durch eine Folge von offenen, paarweise disjunkten Dreiecken. Dies motiviert die Forderung

(3) Ist $(A_n)_{n\in\mathbb{N}}$ eine Folge von Mengen mit $A_i \cap A_j = \emptyset,\ i \neq j$, die jeweils die Maßzahlen a_n besitzen, so soll $\bigcup_{n=1}^{\infty} A_n$ die Maßzahl $\sum_{n=1}^{\infty} a_n$ zugeordnet werden.

Es ist ferner wünschenswert, daß das System solcher "meßbarer Mengen" durchschnittstabil ist, denn

(4) Sind A, B zwei Mengen mit $A \cap B \neq \emptyset$, so soll gelten:
Maßzahl $(A \cup B)$ = Maßzahl (A) + Maßzahl (B) - Maßzahl $(A \cap B)$

(Dies in Konsistenz mit der Anschauung.)

Zur Definition einer solchen Mengenfunktion mit den gewünschten Eigenschaften auf einer möglichst großen Anzahl von Mengen geht man in mehreren Schritten vor:

$$\text{Ring} \longrightarrow \text{Algebra} \longrightarrow \sigma\text{-Algebra}$$

bzw.

$$\text{Inhalt} \longrightarrow \text{Prämaß} \longrightarrow \text{Maß}$$

Daher zunächst die Definition eines Ringes:

Definition 1: $\Re \subset \text{Pot}(\Omega)$ heißt ***Ring*** (in Ω), wenn gilt:

(1) $\emptyset \in \Re$

(2) $A, B \in \Re \Longrightarrow A \backslash B \in \Re$

(3) $A, B \in \Re \Longrightarrow A \cup B \in \Re$

Zur Erinnerung: $\mathcal{R} \subset \text{Pot}(\Omega)$ heißt ***Algebra***, wenn gilt:

(1) $\emptyset \in \mathcal{R}$

(2) $A \in \mathcal{R} \Longrightarrow A^C \in \mathcal{R}$

(3) $A, B \in \mathcal{R} \Longrightarrow A \cup B \in \mathcal{R}$

Der folgende Satz liefert eine Charakterisierung eines Ringes:

Satz 0.1: a.) $\Re$ ist eine Algebra $\Longleftrightarrow$ $\Re$ ist ein Ring und enthält Ω.

b.) Jeder Ring ist durchschnittstabil.

Beweis: a.) "$\Rightarrow$": z.z.: $\Omega \in \Re$, $A \backslash B \in \Re$ für $A, B \in \Re$.

$\Omega = \emptyset^C$, $A \backslash B = (A^C \cup B)^C$

"$\Leftarrow$": z.z.: $A^C \in \Re$ für $A \in \Re$

$A^C = \Omega \backslash A$

b.) Seien $A, B \in \Re \Longrightarrow A \cap B = A \backslash (A \backslash B) \in \Re$

q.e.d.

Im Hinblick auf obige Motivation/Konstuktion definiert man folgende wichtige Mengenfamilie:

Definition 2: $\sum_{i \in I} (a_i, b_i]$, wo $(a_i, b_i] \in J_d := \{ (a, b] : a, b \in \mathbb{R}^d \}$, heißt eine ***d-dimensionale Figur***, falls $|I| < \infty$ und $\{ (a_i, b_i], i \in I \}$ paarweise disjunkt. Mit $\Im_d$ werde die Menge aller solcher Figuren bezeichnet.

$\Im_d$ besitzt die obigen Eigenschaften, denn man hat:

Beispiel 0.2: $\Im_d$ ist ein Ring (in $\mathbb{R}^d$)

Beweis: vgl. Bauer (1978, S.27)

□

Nun zur exakten Definition eines Inhalts bzw. eines Prämaßes:

Definition 3: Es sei ein Ring $\Re$ in Ω gegeben. Eine Abbildung $\mu : \Re \to \overline{\mathbb{R}}$ heißt *Inhalt* bzw. *Prämaß*, falls (1) und (2) bzw. (1) und (3) gelten:

(1) $\mu(A) \geq 0 \quad \forall A \in \Re, \ \mu(\emptyset) = 0$ (Positivität und Nulltreue)

(2) $A_1, A_2 \in \Re, \ A_1 \cap A_2 = \emptyset \implies \mu(A_1 + A_2) = \mu(A_1) + \mu(A_2)$ (Additivität)

(3) $(A_n)_{n\in\mathbb{N}} \subset \Re$ paarweise disjunkt und $\sum_{n=1}^{\infty} A_n \in \Re$

$$\implies \mu(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n) \quad (\ \sigma\text{-Additivität}\ )$$

Bemerkung: Jedes Prämaß ist ein Inhalt.

Beweis: z.z.: $(3) \Longrightarrow (2)$

Setze dazu $A_i := \emptyset, \ i \geq 3$. Dann ist $\sum_{n=1}^{\infty} A_n = A_1 + A_2$.

q.e.d.

Beachte: Der Unterschied zwischen einem Maß und einem Prämaß ist der Definitionsbereich:

Ein Maß ist auf einer σ-Algebra $\mathcal{F}$ definiert, dies sichert die Eigenschaft $\sum_{n=1}^{\infty} A_n \in \mathcal{F}$ für $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$.

Analog zum Maß kann man für den Inhalt folgenden Satz formulieren:

Satz 0.3: Sei μ ein Inhalt auf einem Ring $\Re$. Dann gelten für $A, B, A_i \in \Re$:

(1) $A \subset B \implies \mu(A) \leq \mu(B)$ (Isotonie)

(2) $A \subset B, \mu(A) < \infty \implies \mu(B - A) = \mu(B) - \mu(A)$ (Subtraktivität)

(3) $\mu(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} \mu(A_i)$ (Subadditivität)

Beweis: Analog zu Maßen. (vgl. W-Theorie 7.1)

□

Um den Maßfortsetzungssatz anwenden zu können, benötigt man das Vorliegen eines Prämaßes auf einem Ring $\Re$. Da der Nachweis, daß es sich bei einer Abbildung μ auf $\Re$ um ein Prämaß handelt, wegen der Bedingung (3) aus Definition 3 i.a. schwer zu führen ist, liefert der nachstehende Satz ein praktisches Kriterium:

Satz 0.4: Sei μ ein Inhalt auf dem Ring $\Re$. Ist μ endlich auf $\Re$ (d.h. $\mu(A) < \infty \quad \forall A \in \Re$), so sind äquivalent:

(1) μ ist ein Prämaß

(2) μ ist $\emptyset$-stetig $\big($ d.h. $\mu(A_n) \downarrow 0 \quad \forall (A_n)_{n\in\mathbb{N}} \subset \Re$ mit $A_n \downarrow \emptyset, \ (n \to \infty)$ $\big)$

Beweis: "(1) $\Rightarrow$ (2)": Es sei eine Folge $(A_n)_{n\in\mathbb{N}} \subset \Re$ mit $A_n \downarrow \emptyset\,,\ (n \to \infty)$, gegeben.

Es ist $A_1 = \sum_{n=1}^{\infty}(A_n - A_{n+1})$ (Reduktionssatz)

$$
\begin{aligned}
\Longrightarrow \mu(A_1) &= \mu\Big(\sum_{n=1}^{\infty}(A_n - A_{n+1})\Big)\\
&= \sum_{n=1}^{\infty} \mu(A_n - A_{n+1}) \quad (\ \sigma\text{-Additivität}\ )\\
&= \sum_{n=1}^{\infty} \big(\mu(A_n) - \mu(A_{n+1})\big) \quad (\ \text{Additivität}\ )\\
&\equiv \lim_{m\to\infty} \sum_{n=1}^{m} \big(\mu(A_n) - \mu(A_{n+1})\big)\\
&= \lim_{m\to\infty} \big(\mu(A_1) - \mu(A_{m+1})\big)\\
&= \mu(A_1) - \lim_{m\to\infty} \mu(A_m)
\end{aligned}
$$

$$\Longrightarrow \lim_{m\to\infty} \mu(A_m) = 0$$

"(2) $\Rightarrow$ (1)": z.z.: μ ist σ-additiv.

Gegeben sei eine Folge $(A_n)_{n\in\mathbb{N}} \subset \Re$ mit $A_i \cap A_j = \emptyset$ für $i \neq j$ sowie $\sum_{n=1}^{\infty} A_n \in \Re$.

Es gilt: $\sum_{n=1}^{\infty} A_n \setminus \sum_{n=1}^{m} A_n \downarrow \emptyset\,,\ (m \to \infty)$

$$
\begin{aligned}
&\Longrightarrow \mu\Big(\sum_{n=1}^{\infty} A_n \setminus \sum_{n=1}^{m} A_n\Big) \downarrow 0\,,\ (m \to \infty)\\
&\Longrightarrow \mu\Big(\sum_{n=1}^{\infty} A_n\Big) - \mu\Big(\sum_{n=1}^{m} A_n\Big) \downarrow 0\,,\ (m \to \infty) \quad (\ \text{Subtraktivität}\ )\\
&\Longrightarrow \mu\Big(\sum_{n=1}^{\infty} A_n\Big) - \sum_{n=1}^{m} \mu(A_n) \downarrow 0\,,\ (m \to \infty) \quad (\ \text{Additivität}\ )
\end{aligned}
$$

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

q.e.d.

Wie dieser Satz auf die Menge der Figuren im $\mathbb{R}^d$ anwendbar ist, zeigen die beiden Bemerkungen nach der folgenden Definition:

Definition 4: Für eine Funktion $G : \mathbb{R}^d \to \mathbb{R}$ definiere man für $a, b \in \mathbb{R}^d$

$$\triangle_a^b G := \sum_{y \in \bigtimes_{j=1}^{d} \{a_j, b_j\}} (-1)^{|\{i:\, y_i = a_i\}|}\, G(y)$$

Dann heißt G <u>*$\triangle$-isoton*</u>, falls $\triangle_a^b G \geq 0$ für alle $a, b \in \mathbb{R}^d$ mit $a \leq b$.
($a \leq b$ komponentenweise)

G heißt <u>*maßdefinierende Funktion*</u>, falls gilt:

(1) G ist stetig von oben: $\forall\, (a_n)_{n\in\mathbb{N}} \subset \mathbb{R}^d$ mit $a_n \downarrow a \in \mathbb{R}^d$ gilt: $G(a_n) \downarrow G(a)$

(2) G ist $\triangle$-isoton.

Bemerkung 0.5: Sei $G : \mathbb{R}^d \to \mathbb{R}$ $\triangle$-isoton. Dann existiert genau ein Inhalt μ auf $\Im_d$ mit der Eigenschaft $\quad (*) \quad \mu\big((a,b]\big) \;=\; \triangle_a^b G \quad \forall\, a,b \in \mathbb{R}^d$ mit $a \leq b$.

Beweis: Billingsley (1978), Beweis von Theorem 12.5

□

Bemerkung 0.6: Sei $G : \mathbb{R}^d \to \mathbb{R}$ eine maßdefinierende Funktion. Dann ist der durch (0.5) gegebene Inhalt auf $\Im_d$ ein Prämaß auf $\Im_d$.

□

Es folgt nun der Hauptsatz des Abschnitts, der den Vollzug des Schritts Prämaß $\to$ Maß regelt:

Satz 0.7: (Fortsetzungssatz von Carathéodory)

Jedes Prämaß μ auf einem Ring $\Re$ in Ω kann zu einem Maß $\tilde{\mu}$ auf $\sigma(\Re)$ fortgesetzt werden.

$\big($ gemäß $\tilde{\mu}(A) := \inf \{ \sum_{n=1}^{\infty} \mu(A_n) : (A_n)_{n \in \mathbb{N}} \subset \Re,\ \bigcup_{n=1}^{\infty} A_n \supset A \}$ für $A \in \sigma(\Re)$.

Setze dabei $\inf\{\emptyset\} \;:=\; \infty \,\big)$.

Beweis: Bauer (1978), S.31 ff.

□

Da nun $\mathcal{B}_d \;=\; \sigma(J_d) \;=\; \sigma(\Im_d)$ $\quad$ ($\mathcal{B}_d$ = System der borel'schen Mengen in $\mathbb{R}^d$ bzgl. euklid'scher Metrik) gilt, kann man als Korollar zu (0.6) und (0.7) formulieren:

Korollar 0.8: Zu jeder maßdefinierenden Funktion $G : \mathbb{R}^d \to \mathbb{R}$ gibt es genau ein Maß μ auf $\mathcal{B}_d$ mit der Eigenschaft $(*)$.

$\big($ Die Eindeutigkeit folgt aus dem Eindeutigkeitssatz für Maße (vgl. W-Theorie 7.5) $\big)$

Soviel zum Maßfortsetzungssatz und seinen Voraussetzungen/Anwendungen. Im nächsten Abschnitt wird im Hinblick auf die Martingal- und Markoffketten-Theorie der schon aus dem elementaren Fall bekannte Begriff der bedingten Wahrscheinlichkeit verallgemeinert. Im zweiten Teil des Abschnittes folgt dann, aufbauend auf dem Begriff der bedingten Verteilung, die Definition des bedingten Erwartungswertes.

Abschnitt 1: Bedingte Verteilung und bedingter Erwartungswert

Zur Verdeutlichung der Problemstellung ein Beispiel aus der Warteschlangentheorie:

Ein Ankunftsstrom von Kunden eines Geschäfts werde durch einen Zählprozeß $(N_t)_{t \geq 0}$ eines Erneuerungsprozesses $(T_n)_{n \in \mathbb{N}}$ modelliert (vgl. §32 W-Theorie). $(N_t)_{t \geq 0}$ sei ein Poissonprozeß mit Intensität λ. N_t steht also für die Anzahl der bis zum Zeitpunkt t (einschließlich) eingetroffenen Kunden, T_n ist die Ankunftszeit des n-ten Kunden.

Eine Zufallsvariable X beschreibe die Servicelänge, d.h. die Zeitdauer einer Bedienung. Diese werde unabhängig von $(N_t)_{t \geq 0}$ angenommen (d.h. in diesem Beispiel, daß beim Personal weder Ermüdungserscheinungen noch Lerneffekte bei fortschreitender Zeit auftreten).

Mit Y werde ferner die Anzahl der Kunden bezeichnet, die während der ersten Bedienung ankommen, man definiert daher $Y(\omega) \;:=\; N_{X(\omega)}(\omega)$.

Fragt man nun nach der bedingten Wahrscheinlichkeit, daß bei einer Bedienungszeit x des ersten Kunden k neue Kunden eintreffen, so ist dieses Problem in der Anschauung leicht lösbar: (vgl. §32 W-Theorie)

$\mathrm{P}[\,Y = k \mid X = x] = \mathrm{P}[\,N_x = k] = \Pi(\lambda x, k) \equiv e^{-\lambda x}\dfrac{(\lambda x)^k}{k!}$

Problematisch wird die Fragestellung nach der bedingten Verteilung jedoch in dem Fall, wo X eine Zufallsvariable mit stetiger Verteilung ist. Dann ist nämlich $\mathrm{P}[\,X = x\,] = 0 \quad \forall x$ und die elementare

Definition $\quad \mathrm{P}[\,Y = k \mid X = x\,] := \begin{cases} \dfrac{\mathrm{P}[\,Y = k, X = x\,]}{\mathrm{P}[\,X = x\,]} & \text{falls } \mathrm{P}[\,X = x\,] > 0 \\ 0 & \text{sonst} \end{cases}$

versagt hier ihren Dienst, d.h. liefert i.a. falsche Ergebnisse.
Durch eine Setzung $\mathrm{P}[\,Y = k \mid X = x\,] = 0 \quad \forall x$, d.h. $\mathrm{P}[\,Y = k \mid X = \cdot\,]$ ist Nullfunktion, würden nämlich alle Informationen gelöscht.

Ein Ansatz zur Lösung dieses Problems ist die Einführung von Übergangswahrscheinlichkeiten:

<u>Definition 1:</u> Seien $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$ Meßräume. Eine Abbildung $Q : \Omega_1 \times \mathcal{F}_2 \longrightarrow \mathbb{R}$ heißt <u>*Übergangswahrscheinlichkeit von Ω_1 nach Ω_2*</u> (oder <u>*Markoff-Kern*</u>), falls gilt:

(1) $\quad Q(x, \cdot\,)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_2 \quad \forall x \in \Omega_1$

(2) $\quad Q(\,\cdot\,, C)$ ist $\mathcal{F}_1 - \mathcal{B}$-meßbar $\quad \forall\, C \in \mathcal{F}_2$

Zu dieser Definition einige

<u>Bemerkungen:</u>

1. Um Punkt (2) von Definition 1 zu verifizieren, hat man folgendes Kriterium:
   Sei $\mathcal{F}_2 = \sigma(\mathcal{C}_2)$, wo $\mathcal{C}_2$ ein durchschnittstabiles Mengensystem ist. Dann gilt:
   $Q(\,\cdot\,, C)$ ist $\mathcal{F}_1 - \mathcal{B}$-meßbar $\forall\, C \in \mathcal{F}_2 \iff Q(\,\cdot\,, C)$ ist $\mathcal{F}_1 - \mathcal{B}$-meßbar $\forall\, C \in \mathcal{C}_2$
2. Ist μ ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_2$, so ist die Abbildung
   $Q(x, C) := \mu(C)$ auf $\Omega_1 \times \mathcal{F}_2$ eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2.
3. Ist $f : \Omega_1 \longrightarrow \Omega_2$ meßbar, so ist $Q(x, C) := \delta_{f(x)}(C)$ eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2.
   Dabei sei $\delta_y(M) := \begin{cases} 1 & \text{falls } y \in M \\ 0 & \text{sonst} \end{cases}$ <u>*Einpunktmaß*</u> auf y.
   Diese Übergangswahrscheinlichkeit wird auch als <u>*degenerierte Übergangswahrscheinlichkeit*</u> bezeichnet.

<u>Beweise:</u>

1. "$\Rightarrow$": klar
   "$\Leftarrow$": z.z.: $\mathcal{M} := \{\, A \in \mathcal{F}_2 : Q(\,\cdot\,, A)$ ist $\mathcal{F}_1 - \mathcal{B}$-meßbar$\,\} \supset \mathcal{C}_2$ ist ein Dynkinsystem.
   (vgl. §14 W-Theorie)
   Es ist $\emptyset \in \mathcal{M}$, da $Q(\,\cdot\,, \emptyset) = 0$, also konstant und daher meßbar.
   Ist $A \in \mathcal{M} \implies Q(\,\cdot\,, A)$ ist meßbar
   $\implies 1 - Q(\,\cdot\,, A) \equiv Q(\,\cdot\,, A^C)$ ist meßbar, d.h. $A^C \in \mathcal{M}$.
   Sind $(A_i)_{i \in \mathbb{N}} \subset \mathcal{M}$ paarweise disjunkt $\quad \implies Q(\,\cdot\,, A_i)$ ist meßbar $\forall\, i \in \mathbb{N}$
   $\implies \displaystyle\sum_{i=1}^{\infty} Q(\,\cdot\,, A_i) \equiv Q(\,\cdot\,, \sum_{i=1}^{\infty} A_i)$ ist meßbar, d.h. $\displaystyle\sum_{i=1}^{\infty} A_i \in \mathcal{M}$.
   $\implies \mathcal{M}$ ist ein Dynkinsystem $\quad \implies \mathcal{F}_2 = \sigma(\mathcal{C}_2) = \mathcal{M}$.
2. Der Beweis ist trivial, da $Q(x, \cdot\,) = \mu$ ein Wahrscheinlichkeitsmaß und $Q(\,\cdot\,, C)$ konstant, mithin meßbar.
3. $Q(x, \cdot\,) = \delta_{f(x)}(\,\cdot\,)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_2$.
   $Q(\,\cdot\,, C)$ ist meßbar $\forall\, C \in \mathcal{F}_2$, da $Q(\,\cdot\,, C)^{-1}(1) \equiv f^{-1}(C) \in \mathcal{F}_1$.

q.e.d.

Im folgenden wird nun mit Hilfe einer Übergangswahrscheinlichkeit ein Maß auf $\mathcal{F}_1 \otimes \mathcal{F}_2$ entwickelt, dazu zwei vorbereitende Lemmata:

Lemma 1.1: Sei $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ und Q Übergangswahrscheinlichkeit von Ω_1 nach Ω_2. Dann ist die Abbildung $[\, x \mapsto Q(x, A_x)]$ $\mathcal{F}_1 - \mathcal{B}$-meßbar.
Zur Erinnerung: $A_x := \{\, y \in \Omega_2 : (x,y) \in A \,\} = x$-Schnitt von A.

Beweis: Der Beweis verläuft völlig analog zu dem von Lemma 20.3 der W-Theorie. Man zeigt, daß das System $\mathcal{M} := \{\, A \in \mathcal{F}_1 \otimes \mathcal{F}_2 : x \mapsto Q(x, A_x)$ ist meßbar $\}$ ein Dynkinsystem ist. (leicht)
Da $\mathcal{F}_1 \times \mathcal{F}_2$ durchschnittstabil und $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{M}$
$\big(A = A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2 \Longrightarrow Q\big(x, (A_1 \times A_2)_x\big) = 1_{A_1}(x)\, Q(x, A_2) \big)$
folgt: $\mathcal{F}_1 \otimes \mathcal{F}_2 = \mathcal{M}$.

q.e.d.

Lemma 1.2: Sei $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ $\mathcal{F}_1 \otimes \mathcal{F}_2$-meßbar. Ist Q eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2 und existiert $h(x) := \int\limits_{\Omega_2} Q(x, dy)\, f(x,y) \quad \forall x \in \Omega_1$, so ist h $\mathcal{F}_1 - \overline{\mathcal{B}}$-meßbar.

Beweis: (vgl. Übung, dort sogar Q σ-endliches Übergangsmaß)
Man verwendet das Beweisprinzip für Integrale:
Ist $f = 1_A$ für ein $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$, so folgt: $h(x) = Q(x, A_x)$. h ist somit nach (1.1) meßbar.
Der weitere Verlauf des Beweises ($f = \sum\limits_{i=1}^{m} a_i \cdot 1_{A_i}$, $f \geq 0$, f beliebig) ist einfach.

□

Nun die schon angekündigte Definition eines Maßes auf $\mathcal{F}_1 \otimes \mathcal{F}_2$ mittels Übergangswahrscheinlichkeiten:

Satz 1.3: Sei P_1 ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_1$ und Q Übergangswahrscheinlichkeit von Ω_1 nach Ω_2. Dann ist P definiert durch

$$\mathrm{P}[\, A\,] \equiv (\mathrm{P}_1 \otimes Q)[\, A\,] := \int\limits_{\Omega_1} \mathrm{P}_1(dx)\, Q(x, A_x) \quad \text{für } A \in \mathcal{F}_1 \otimes \mathcal{F}_2$$

das einzige Wahrscheinlichkeitsmaß P auf $\mathcal{F}_1 \otimes \mathcal{F}_2$ mit der Eigenschaft

$$\mathrm{P}[\, A_1 \times A_2\,] = \int\limits_{A_1} \mathrm{P}_1(dx)\, Q(x, A_2) \quad \text{für } A_i \in \mathcal{F}_i \,,\ i = 1,2.$$

Beweis: Man bemerke, daß die Definition von P nach (1.1) sinnvoll ist, da $Q(x, A_x)$ meßbar ist.
Das Integral existiert, da gilt: $0 \leq Q(x, A_x) \leq 1 \quad \forall x \in \Omega_1 \,,\ A \in \mathcal{F}_2$.
Der Rest des Beweises verläuft völlig analog zu dem von Produktmaßen auf $\mathcal{F}_1 \otimes \mathcal{F}_2$,
(vgl. W-Theorie 20.7 a.) bzw. Übung)

□

Es liegt nahe, für diese Art von Maßen den Satz von Fubini zu formulieren:

Satz 1.4: (Satz von Fubini)
Sei P_1 ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_1$, Q Übergangswahrscheinlichkeit von Ω_1 nach Ω_2 und $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ $\mathcal{F}_1 \otimes \mathcal{F}_2$-meßbar. Dann gelten die folgenden Aussagen:

a.) Ist $f \geq 0$, so gilt $(**)$ $\quad \int\limits_{\Omega_1 \times \Omega_2} f\, d(\mathrm{P}_1 \otimes Q) = \int\limits_{\Omega_1} \mathrm{P}_1(dx) \int\limits_{\Omega_2} Q(x, dy)\, f(x,y)$

b.) Ist $f \in \mathcal{L}'_{\mathrm{P}_1 \otimes Q}$ (d.h. f $\mathrm{P}_1 \otimes Q$-quasiintegrierbar), so ist die Abbildung $g(x) := \int\limits_{\Omega_2} Q(x, dy)\, f(x,y)$ P_1-fast sicher erklärt. g ist aus $\mathcal{L}'_{\mathrm{P}_1}$ und für f gilt die Gleichung $(**)$.

Beweis: Das Integral in a.) ist definiert, da nach (1.2) $h(x) := \int\limits_{\Omega_2} Q(x,dy)\, f(x,y)$ $\mathcal{F}_1 - \overline{\mathcal{B}}$-meßbar, außerdem ist $h \geq 0$ wegen $f \geq 0$. Daher ist $\int\limits_{\Omega_1} \mathrm{P}_1(dx)\, h(x)$ existent.
Der Rest des Beweises verläuft wieder analog zu dem von Produktmaßen.
(vgl. W-Theorie 20.7 b.) bzw. Übung)

$\square$

Bemerkung: Die Voraussetzung $f \in \mathcal{L}'_{\mathrm{P}_1 \otimes Q}$ in b.) überprüft man durch Berechnung der beiden iterierten Integrale für $f^{\pm}$ mit Hilfe von a.).

Spezialfall: Es gilt für $f \in \mathcal{L}'_{\mathrm{P}_1 \otimes Q}$ und $A_i \in \mathcal{F}_i\,, i = 1,2$:

$$\int\limits_{A_1 \times A_2} f\, d(\mathrm{P}_1 \otimes Q) = \int\limits_{A_1} \mathrm{P}_1\,(dx) \int\limits_{A_2} Q(x,dy)\, f(x,y)$$

Dies ist eine direkte Folgerung aus (1.4) b.), wenn man statt f $\;f \cdot 1_{A_1 \times A_2}$ einsetzt ($f \cdot 1_{A_1 \times A_2} \in \mathcal{L}'_{\mathrm{P}_1 \otimes Q}$) und faktorisiert.

Nun zurück zum Problem der bedingten Verteilung einer Zufallsvariablen Y bzgl. einer stetig verteilten Zufallsvariablen X: $\big($ X, Y, Z seien im folgenden, falls nichts anderes gesagt, Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$ bzw. $(\Omega_3, \mathcal{F}_3)$ $\big)$
Da die bedingte Verteilung im diskreten Fall erklärt ist, verwendet man diese als Vorbild für die Konstruktion der bedingten Verteilung im allgemeinen Fall:
Ist X diskret, so gilt: $\mathrm{P}[\,X \in A, Y \in B\,] = \sum\limits_{x \in A} \mathrm{P}[\,Y \in B \mid X = x\,] \cdot \mathrm{P}[\,X = x\,]$ für $A \in \mathcal{F}_1\,,\ B \in \mathcal{F}_2$.
Daraus wird im allgemeinen Fall: $(*)$ $\mathrm{P}[\,X \in A, Y \in B\,] = \int\limits_{A} \mathrm{P}X^{-1}\,(dx)\, Q(x,B)$ für $A \in \mathcal{F}_1\,,\ B \in \mathcal{F}_2$.
Man sucht daher als Analogon zu $\mathrm{P}[\,Y \in B \mid X = x\,]$ im diskreten Fall eine Funktion $Q : \Omega_1 \times \mathcal{F}_2 \longrightarrow \mathbb{R}$, welche folgende drei Bedingungen erfüllen muß:

1.) Q erfüllt die Gleichung $(*)$.
2.) Q ist bzgl. seiner ersten Komponente $\mathcal{F}_1 - \mathcal{B}$-meßbar, damit das Integral in $(*)$ definiert ist.
3.) Setzt man $B := \emptyset$, so erhält man: $0 = \mathrm{P}[\,X \in A, Y \in \emptyset\,] = \int\limits_{A} \mathrm{P}X^{-1}\,(dx)\, Q(x,\emptyset)$ $\;\forall A \in \mathcal{F}_1$,
   also $Q(\,\cdot\,,\emptyset) = 0$ $\mathrm{P}X^{-1}$-fast sicher. $\big($ W-Theorie 18.6 c.) $\big)$
   Analog: $B = \Omega \Longrightarrow Q(\,\cdot\,,\Omega) = 1$ $\mathrm{P}X^{-1}$-fast sicher.
   $$B = \sum_{i=1}^{\infty} B_i\,,\ B_i \in \mathcal{F}_2 \Longrightarrow Q(\,\cdot\,,\sum_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} Q(\,\cdot\,,B_i) \quad \mathrm{P}X^{-1}\text{-fast sicher.}$$

Die Bedingungen 2.) und 3.) werden durch eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2 erfüllt, daher wird die folgende Definition sinnvoll:

Definition 2: a.) Eine Abbildung $Q : \Omega_1 \times \mathcal{F}_2 \longrightarrow \mathbb{R}$ heißt <u>bedingte Verteilung von Y bzgl. X</u>, falls sie erfüllt:
1.) Q ist eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2.
2.) $\mathrm{P}(X,Y)^{-1} = \mathrm{P}X^{-1} \otimes Q$ $\big($ dies ist nach (1.3) äquivalent zu:
$$\mathrm{P}[\,X \in A, Y \in B\,] \equiv \mathrm{P}(X,Y)^{-1}[\,A \times B\,] = (\mathrm{P}X^{-1} \otimes Q)\,[\,A \times B\,]$$
$$\equiv \int\limits_{A} \mathrm{P}X^{-1}\,(dx)\, Q(x,B) \quad \forall A \in \mathcal{F}_1\,,\ B \in \mathcal{F}_2 \;\big)$$

Schreibweisen: $Q = \mathrm{P}Y^{-1} \mid X = \cdot\ ,\ Q(x,B) = \mathrm{P}[\,Y \in B \mid X = x\,]$. Mit $\mathrm{I\!P}Y^{-1} \mid X = \cdot$ bezeichnet man die Menge aller bedingten Verteilungen von Y bzgl. X.

Wie man Wahrscheinlichkeitsmaße mit Dichten beschreiben kann, so verfährt man auch mit der bedingten Verteilung:

Definition 2: b.) Sei ν ein σ-endliches Maß auf $\mathcal{F}_2$. Dann heißt eine $\mathcal{F}_1 \otimes \mathcal{F}_2$-meßbare Funktion $g : \Omega_1 \times \Omega_2 \to [\,0,\infty\,)$ eine *bedingte ν-Dichte von Y bzgl. X*, falls die Abbildung $(x,B) \mapsto \int\limits_B \nu\,(dy)\,g(x,y)$ ein Element aus $\mathrm{I\!P}Y^{-1} \mid X = \cdot$ ist.

Bemerkung: Um nachzuweisen, daß eine Funktion g eine bedingte ν-Dichte von Y bzgl. X ist, hilft das Kriterium:

Ist $g : \Omega_1 \times \Omega_2 \to [\,0,\infty\,)$ $\mathcal{F}_1 \otimes \mathcal{F}_2$-meßbar und $\int\limits_{\Omega_2} \nu\,(dy)\,g(x,y) = 1 \quad \forall x \in \Omega_1$,

so ist $Q(x,B) := \int\limits_B \nu\,(dy)\,g(x,y)$ Übergangswahrscheinlichkeit von Ω_1 nach Ω_2.

Um zu zeigen, daß g eine bedingte ν-Dichte von Y bzgl. X ist, müßte man nur noch Punkt 2.) der Definition 2 a.) nachweisen.

Beweis: 1.) $Q(x,\cdot\,)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_2$.

($Q(x,\cdot\,) \geq 0$ wegen $g \geq 0$; $Q(x,\emptyset) = 0$, $Q(x,\Omega_2) = 1$ nach Voraussetzung,

$Q(x, \sum\limits_{i=1}^{\infty} A_i) = \sum\limits_{i=1}^{\infty} Q(x,A_i)$ nach monotoner Konvergenz)

2.) Es existiert $Q(x,B) \equiv \int\limits_B \nu\,(dy)\,g(x,y) \equiv \int\limits_{\Omega_2} 1_B(y)\,\nu\,(dy)\,g(x,y)$

für festes $B \in \mathcal{F}_2$ und alle $x \in \Omega_1$. Nach einer Vorstufe des Satzes von Fubini (W-Theorie 11.6) ist dann $Q(\,\cdot\,,B)$ $\mathcal{F}_1 - \mathcal{B}$-meßbar.

q.e.d.

Betrachte nun die Anwendung der vorigen Definition auf das Einführungsbeispiel:

Beispiel: Es war $T_n := \sum\limits_{i=1}^{n} Z_i$ mit unabhängigen $Z_i \sim \mathrm{Exp}(\lambda)$. X war von den Z_i unabhängig angenommen. Es gilt für den Zählprozeß $(N_t)_{t\geq 0}$ (vgl. §32 W-Theorie):

$$N_t = k \iff T_k \leq t < T_{k+1}$$

Somit ist $Y := N_X \equiv \sum\limits_{n=1}^{\infty} 1_{\{\,T_n \leq X\,\}}$ Zufallsvariable und es gilt für $A \in \mathcal{B}$:

$$\begin{aligned}
\mathrm{P}[\,X \in A, Y = k\,] &= \mathrm{P}[\,X \in A, T_k \leq X < T_{k+1}\,] \\
&= \mathrm{P}[\,\big(X,(T_k,T_{k+1})\big) \in \tilde{A}\,] \ \text{ wo } \tilde{A} = \big\{\big(x,(s,t)\big) : x \in A,\ s \leq x < t\big\} \in \mathcal{B}_3 \\
&= \int\limits_A \mathrm{P}X^{-1}\,(dx)\,\mathrm{P}[\,(T_k,T_{k+1}) \in \tilde{A}_x\,] \quad \text{nach W-Theorie 13.4 c.),} \\
&\qquad\qquad\qquad\qquad \text{da } X,(T_k,T_{k+1}) \text{ unabhängig} \\
&= \int\limits_A \mathrm{P}X^{-1}\,(dx)\,\mathrm{P}[\,T_k \leq x < T_{k+1}\,] \\
&\equiv \int\limits_A \mathrm{P}X^{-1}\,(dx)\,\mathrm{P}[\,N_x = k\,].
\end{aligned}$$

Vergleicht man dieses Resultat mit obiger Definition, so stellt man fest:
$Q(x,B) := \mathrm{P}[\,N_x \in B\,]$ ist bedingte Verteilung von Y bzgl. X.

Benutzt man dieses Ergebnis, so kann man z.B. die Wahrscheinlichkeit berechnen, daß k Kunden während der ersten Bedienung eintreffen:

$$\mathrm{P}[\,Y = k\,] \equiv \mathrm{P}[\,X \in \Omega_1, Y = k\,] = \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\,Q(x,k)$$

$$\equiv \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\,\mathrm{P}[\,N_x = k\,] \equiv \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\;\; e^{-\lambda x}\cdot\frac{(\lambda x)^k}{k!}$$

Damit errechnet sich leicht der Erwartungswert von Y:

$$\mathrm{E}[\,Y\,] \equiv \sum_{k\geq 0} k\cdot \mathrm{P}[\,Y = k\,] = \sum_{k\geq 0} k\cdot \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\,\mathrm{P}[\,N_x = k\,]$$

$$= \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx) \sum_{k\geq 0} k\cdot \mathrm{P}[\,N_x = k\,] \quad \text{wegen monotoner Konvergenz}$$

$$\equiv \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\,\mathrm{E}[\,N_x\,] = \int\limits_{\Omega_1} \mathrm{P}X^{-1}\,(dx)\;\;\lambda x \equiv \lambda\cdot \mathrm{E}[\,X\,]$$

$\mathrm{E}[\,X\,]$ ist die _*mittlere Servicedauer*_, $\lambda\cdot\mathrm{E}[\,X\,]$ wird als _*Verkehrsrate*_ bezeichnet.

So sinnvoll die obige Konstruktion der bedingten Verteilung auch erscheint, es stellt sich natürlich die Frage nach der Existenz. Diese Frage beantwortet Satz 1.5, jedoch zuvor noch

Lemma 1.5$'$: Sei $\mathcal{F}_2 = \sigma(\mathcal{C}_2)$, wo $\mathcal{C}_2$ ein durchschnittstabiles Mengensystem ist. Dann ist Q bedingte Verteilung von Y bzgl. $X \iff$

(i) $Q(x,\,\cdot\,)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_2 \quad \forall x \in \Omega_1$

(ii) Die Abbildung $x \mapsto Q(x,C)$ ist $\mathcal{F}_1 - \mathcal{B}$-meßbar $\quad \forall C \in \mathcal{C}_2$

(iii) $\mathrm{P}[\,X \in A, Y \in C\,] = \int\limits_A \mathrm{P}X^{-1}\,(dx)\,Q(x,C) \quad \forall A \in \mathcal{F}_1\,,\; C \in \mathcal{C}_2$

Beweis: "$\Rightarrow$": klar.

"$\Leftarrow$": Die Bedingungen (i) und (ii) kennzeichnen nach Definition 1 und Bemerkung 1 Q als Übergangswahrscheinlichkeit.
Die Bedingung 2 aus Definition 2 a.) erhält man aus obiger Bedingung (iii) leicht per Dynkinargument. (man zeigt, daß für $A \in \mathcal{F}_1$ die Menge

$$\mathcal{D}_A := \{\, B \in \mathcal{F}_2 : \mathrm{P}[\,X \in A, Y \in B\,] = \int\limits_A \mathrm{P}X^{-1}\,(dx)\,Q(x,B)\,\} \text{ ein Dynkinsystem ist)}$$

$\square$

Satz 1.5: (Existenz der bedingten Verteilung)

Ist Ω_2 borel'sche Teilmenge eines polnischen Raumes, so gilt: $\quad \mathbb{P}Y^{-1} \mid X = \cdot \neq \emptyset$

Bemerkung: Ein topologischer Raum E heißt polnisch, wenn es eine seine Topologie definierende, vollständige Metrik gibt und wenn E eine abzählbare Basis besitzt. Eine Metrik heißt dabei vollständig, wenn der zugehörige metrische Raum vollständig ist. Eine abzählbare Basis ist ein System S von abzählbar vielen offenen Mengen von E derart, daß sich jede offene Menge aus E als Vereinigung von Mengen aus S darstellen lässt.
Die Bedingung "polnisch" deckt alle Fälle der Praxis ab.

Beweis: Hier wird der Fall $\Omega_2 = \mathbb{R}$ behandelt. Der allgemeine Fall wird dadurch gelöst, daß falls Ω_2 abzählbar ist, die elementare bedingte Verteilung existiert. Ist Ω_2 überabzählbar, so ist Ω_2 borel-isomorph zu $\mathbb{R}$ (vgl. Übung). Ein Borel-Isomorphismus ist dabei eine bijektive, in beiden Richtungen borel-meßbare Abbildung.
Das Ziel des nun folgenden Beweises ist die Konstruktion einer Funktion $G : \mathbb{R}\times\Omega_1 \to [\,0,1\,]$ mit der Eigenschaft $r \mapsto G(r \mid x)$ ist Verteilungsfunktion zu $\mathrm{P}Y^{-1} \mid X = x$, d.h. für $x \in \Omega_1$ gilt: $\quad G(r \mid x) = \mathrm{P}[\,Y \leq r \mid X = x\,] \quad \forall r \in \mathbb{R}$.

Deswegen wird der Fall $\Omega_2 = \mathbb{R}$ gewählt, da ein Wahrscheinlichkeitsmaß auf $\mathcal{B}_1$ eindeutig durch seine Verteilungsfunktion bestimmt ist (W-Theorie 7.13) und eine Verteilungsfunktion einfacher zu konstruieren ist als ein Wahrscheinlichkeitsmaß.

<u>1. Schritt:</u> Konstruktion einer PX^{-1}-Dichte $F(r \mid \cdot)$ von $P_r X^{-1}$, $r \in \mathbb{Q}$,
wo $P_r X^{-1}[A] := P[Y \le r, X \in A]$ $(\le P[X \in A] \equiv PX^{-1}[A])$ für $A \in \mathcal{F}_1$.

<u>Beh.:</u> $\exists F(r \mid \cdot) : \Omega_1 \to [0,\infty)$, welches $\mathcal{F}_1$-meßbar ist und welches erfüllt:

$$P[Y \le r, X \in A] = \int_A PX^{-1}(dx)\, F(r \mid x)\,,\ r \in \mathbb{Q}$$

<u>1. Beweis:</u> $1_{\{Y \le r\}} \in \mathcal{L}^2(P) \equiv \{ f\colon \Omega \to \mathbb{R} : f\ \mathcal{F}\text{-meßbar, bzgl. P quadratisch integrierbar} \}$
Wähle $F(r \mid \cdot) \in \mathbb{E}[1_{\{Y \le r\}} \mid X = \cdot\,]$ = Menge der bedingten Erwartungswerte von $1_{\{Y \le r\}}$ gegeben X. ($\mathcal{L}^2$-Definition)
F ist nach dem Projektionssatz für Hilberträume existent (vgl. Statistik 4.3).
Nach Statistik 4.4 gilt dann: $E[1_{\{Y \le r\}}(1_A \circ X)] = E[F(r \mid X) \cdot (1_A \circ X)]$

$$\left(\iff P[Y \le r, X \in A] = \int_A PX^{-1}(dx)\, F(r \mid x) \right)$$

<u>2. Beweis:</u> Wie oben schon angedeutet, gilt $P_r X^{-1}[A] \le PX^{-1}[A] \quad \forall A \in \mathcal{F}_1$
$\implies P_r X^{-1} << PX^{-1}$, d.h. $P_r X^{-1}$ ist PX^{-1}-stetig. Daher besitzt $P_r X^{-1}$ nach dem Satz von Radon-Nikodym eine PX^{-1}-Dichte $F(r \mid \cdot)$.

<u>2. Schritt:</u> Konstruktion einer Menge $\mathcal{M} \subset \Omega_1$, so daß $F(\cdot \mid x)$ isoton auf $\mathbb{Q}$ $\forall x \in \mathcal{M}$.
Es gilt für alle $A \in \mathcal{F}_1$ und alle $r \le r'$ mit $r, r' \in \mathbb{Q}$

$$\int_A PX^{-1}(dx)\, F(r \mid x) \le \int_A PX^{-1}(dx)\, F(r' \mid x) \le \int_A 1\ PX^{-1}(dx)$$

wegen $P[Y \le r, X \in A] \le P[Y \le r', X \in A] \le P[X \in A]$.
Da die Ungleichungskette für alle $A \in \mathcal{F}_1$ gilt, sind die Integranden vergleichbar.
(W-Theorie 9.20 a.))
Mit $\mathcal{M}(r,r') := \{ x \in \Omega_1 : F(r \mid x) \le F(r' \mid x) \le 1 \}$ definiert für $r, r' \in \mathbb{Q}$ mit $r \le r'$ erhält man somit: $PX^{-1}[\mathcal{M}(r,r')] = 1$. Dies gilt ebenso für $\mathcal{M} := \bigcap_{r \le r';\, r,r' \in \mathbb{Q}} \mathcal{M}(r,r') \in \mathcal{F}_1$, da $\mathcal{M}$ abzählbarer Schnitt von Mengen mit dem Maß 1.

<u>3. Schritt:</u> Konstruktion einer Menge $\mathcal{M}' \subset \mathcal{M} \subset \Omega_1$, so daß $\forall x \in \mathcal{M}'$ $F(\cdot \mid x)$ rechtsstetig in allen $r \in \mathbb{Q}$ ist. Es gilt für $r \in \mathbb{Q}$ und alle $A \in \mathcal{F}_1$:

$$(+) \quad P[Y \le r, X \in A] = \lim_{r' \downarrow r,\ r' \in \mathbb{Q}} P[Y \le r', X \in A]$$

Ferner ist
$$\begin{aligned}
\int_A PX^{-1}(dx)\, F(r \mid x) \cdot 1_{\mathcal{M}}(x) &\overset{(*)}{=} \int_A PX^{-1}(dx)\, F(r \mid x) \\
&= P[Y \le r, X \in A] \\
&\overset{(+)}{=} \lim_{r' \downarrow r,\ r' \in \mathbb{Q}} P[Y \le r', X \in A] \\
&\equiv \lim_{r' \downarrow r,\ r' \in \mathbb{Q}} \int_A PX^{-1}(dx)\, F(r' \mid x) \\
&\overset{(*)}{=} \lim_{r' \downarrow r,\ r' \in \mathbb{Q}} \int_A PX^{-1}(dx)\, F(r' \mid x) \cdot 1_{\mathcal{M}}(x) \\
&\overset{(**)}{=} \int_A PX^{-1}(dx) \left(\lim_{r' \downarrow r,\ r' \in \mathbb{Q}} F(r' \mid x) \right) \cdot 1_{\mathcal{M}}(x)
\end{aligned}$$

Dabei gelten $(*)$ wegen $PX^{-1}[\mathcal{M}] = 1$ und $(**)$ aufgrund majorisierter Konvergenz (der Integrand ist fallend in r', beschränkt durch 1 und 0).

Da diese Gleichungskette für alle $A \in \mathcal{F}_1$ stattfindet, gilt nach W-Theorie 9.20 b.) für die Menge $\mathcal{M}' := \bigcap_{r \in \mathbb{Q}} \{ x \in \mathcal{M} : \lim_{r' \downarrow r\,,\, r' \in \mathbb{Q}} F(r' \mid x) = F(r \mid x) \}$: $\quad \mathrm{P}X^{-1}[\mathcal{M}'] = 1$.

<u>4. Schritt:</u> Völlig analog zu Schritt 3 konstruiert man eine Menge $\mathcal{M}'' \subset \mathcal{M}' \subset \mathcal{M} \subset \Omega_1$ mit den Eigenschaften $\mathrm{P}X^{-1}[\mathcal{M}''] = 1$, $\lim_{r \to -\infty\,,\, r \in \mathbb{Q}} F(r \mid x) = 0 \quad \forall x \in \mathcal{M}''$ und

$\lim_{r \to \infty\,,\, r \in \mathbb{Q}} F(r \mid x) = 1 \quad \forall x \in \mathcal{M}''$ mittels den Beziehungen

$\mathrm{P}[X \in A] = \lim_{r' \to \infty\,,\, r' \in \mathbb{Q}} \mathrm{P}[Y \le r', X \in A]$ und $0 = \mathrm{P}[\emptyset] = \lim_{r' \to -\infty\,,\, r' \in \mathbb{Q}} \mathrm{P}[Y \le r', X \in A]$

Somit hat man die gesuchte Funktion $G(r \mid x)$ für $r \in \mathbb{Q}$ auf einer Menge $\mathcal{M}'' \subset \Omega_1$ mit $\mathrm{P}X^{-1}[\mathcal{M}''] = 1$ in Form von F schon konstruiert.

<u>5. Schritt:</u> Ausdehnung der Konstruktion auf $\mathbb{R}$ und ganz Ω_1:

Definiere für $x \in \Omega_1\,,\ r \in \mathbb{R} \quad G(r \mid x) := \begin{cases} \inf_{r' \in \mathbb{Q};\, r' > r} F(r' \mid x) & \text{für } x \in \mathcal{M}'' \\ 1_{[0,\infty)}(r) & \text{für } x \notin \mathcal{M}'' \end{cases}$

Mit dieser Definition wird $G(\,\cdot \mid x)$ Verteilungsfunktion $\forall x \in \Omega_1$.

(Der Fall $x \notin \mathcal{M}''$ ist trivial. Für $x \in \mathcal{M}''$ sind die Bedingungen $G(-\infty \mid x) = 0$ und $G(\infty \mid x) = 1$ trivial. Die rechtsseitige Stetigkeit von $G(\,\cdot \mid x)$ in $r \in \mathbb{R}$ folgt aus der Tatsache, daß es wegen der Isotonie von $G(\,\cdot \mid x)$ genügt, die Bedingung $G(r' \mid x) \downarrow G(r \mid x)$ für $r' \downarrow r$ nur im Falle $r' \downarrow r\,,\ r' \in \mathbb{Q}$ zu betrachten.

Dann ist nämlich wegen Schritt 3 $G(r' \mid x) = F(r' \mid x)$ und $F(r' \mid x) \downarrow G(r \mid x)$ aufgrund der Infimumseigenschaft von $G(r \mid x)$.)

Ferner ist $G(r \mid \cdot\,)$ $\mathcal{F}_1$-meßbar $\forall r \in \mathbb{Q}$, da wegen Schritt 3 $G(r \mid x) = F(r \mid x)$ für $r \in \mathbb{Q}$ und $x \in \mathcal{M}''$, d.h. $G(r \mid x) = F(r \mid x) \cdot 1_{\mathcal{M}''}(x) + 1_{[0,\infty)}(r) \cdot 1_{\mathcal{M}''^C}(x)$.

<u>6. Schritt:</u> Überprüfung, ob G die bedingte Verteilung liefert:

Sei $Q(x, \cdot\,)$ das zur Verteilungsfunktion $G(\,\cdot \mid x)$ gehörende Wahrscheinlichkeitsmaß (vgl. W-Theorie 7.13) und $\tilde{J}_1' := \{ (-\infty, r] : r \in \mathbb{Q} \}$. $Q(x, \cdot\,)$ ist ein Wahrscheinlichkeitsmaß auf $\mathcal{B}_1 \quad \forall x \in \Omega_1$. Nach Schritt 5 ist die Abbildung $x \mapsto Q(x, C)$ für $C \equiv (-\infty, r] \in \tilde{J}_1'$ $\mathcal{F}_1 - \mathcal{B}_1$-meßbar, da $Q\big(x, (-\infty, r]\big) \equiv G(r \mid x)$.

Aufgrund dieser Identität gilt für C auch $\mathrm{P}[Y \in C, X \in A] = \int_A \mathrm{P}X^{-1}(dx)\, Q(x, C)$, da $Q(x, C) = Q\big(x, (-\infty, r]\big) \equiv G(r \mid x) = F(r \mid x)$ auf $\mathcal{M}''$.

$\tilde{J}_1'$ ist durchschnittstabiles Mengensystem und erzeugt $\mathcal{B}_1$ (vgl. W-Theorie 6.5), es folgt mit Lemma 1.5′: $\quad Q(\,\cdot\,, \cdot\,)$ ist die gesuchte bedingte Verteilung von Y bzgl. X.

q.e.d.

Somit ist die Existenz einer bedingten Verteilung geklärt, es stellt sich nun die Frage, inwiefern sie eindeutig ist:

<u>Satz 1.6:</u> (Eindeutigkeit der bedingten Verteilung)

Sei $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$ und Q' eine Übergangswahrscheinlichkeit von Ω_1 nach Ω_2. Dann sind äquivalent:

(1) $\quad Q' \in \mathbb{P}Y^{-1} \mid X = \cdot$

(2) $\quad Q(\,\cdot\,, B) = Q'(\,\cdot\,, B) \quad \mathrm{P}X^{-1}$-fast sicher $\quad \forall B \in \mathcal{F}_2$.

<u>Beweis:</u> Der Beweis ist leicht, vgl. Übung.

□

Bemerkung: Verschiedene Elemente aus $\mathbb{P}Y^{-1} \mid X = \cdot$ bezeichnet man auch als <u>Versionen</u> von $\mathrm{P}Y^{-1} \mid X = \cdot$

Es folgen nun einige Eigenschaften der bedingten Verteilung wie das Verhalten bei stochastischer Unabhängigkeit sowie die Einführung eines "bedingten Bildmaßes":

Satz 1.7: a.) Es sind äquivalent:

(1) X, Y sind unabhängig

(2) $\exists\, Q \in \mathbb{P}Y^{-1} \mid X = \cdot$, so daß $Q(x, \cdot\,)$ von x unabhängig ist

(3) Die Abbildung $(x, B) \mapsto \mathrm{P}[\,Y \in B\,]$ ist eine Version von $\mathrm{P}Y^{-1} \mid X = \cdot$

b.) Sind $X, (Y, Z)$ unabhängig, so gilt:
Die Abbildung $(x, z, B) \mapsto \mathrm{P}[\,Y \in B \mid Z = z\,]$ ist eine Version von $\mathrm{P}Y^{-1} \mid (X, Z) = \cdot$

c.) "bedingtes Bildmaß": Sei $T : \Omega_1 \times \Omega_2 \to \Omega'$ meßbar und $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$
Dann ist $\big(\mathrm{P}\big(T \circ (X,Y)\big)^{-1} \mid X = x\big) = (\mathrm{P}Y^{-1} \mid X = x)\,\big(T(x, \cdot\,)\big)^{-1}$, d.h. mit

$Q_T := Q\big(x, \{\, y : T(x,y) \in B\,\}\big)$ gilt: $\quad Q_T \in \mathbb{P}\big(T \circ (X,Y)\big)^{-1} \mid X = \cdot$

Beweis: a.) ist leicht zu beweisen, die Teile b.) und c.) beweist man mit Hilfe von (1.3), den elementaren Eigenschaften der stochastischen Unabhängigkeit sowie der Integration mittels Bildmaß (vgl. Übung).

Beweis von a.):

"(1)$\Rightarrow$(3)": Für $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$ gilt:

$$\mathrm{P}[\,X \in A, Y \in B\,] = \mathrm{P}[\,X \in A\,] \cdot \mathrm{P}[\,Y \in B\,] \quad \text{nach Definition der Unabhängigkeit}$$
$$\equiv \int_A \mathrm{P}X^{-1}\,(dx)\, \mathrm{P}[\,Y \in B\,]$$

und somit $\mathrm{P}Y^{-1} \in \mathbb{P}Y^{-1} \mid X = \cdot$ nach Definition 2.

"(3)$\Rightarrow$(2)": $Q := \mathrm{P}Y^{-1}$

"(2)$\Rightarrow$(1)":
$$\mathrm{P}[\,X \in A, Y \in B\,] = \int_A \mathrm{P}X^{-1}\,(dx)\, Q(B) \quad \text{nach Voraussetzung}$$
$$\equiv Q(B) \cdot \mathrm{P}[\,X \in A\,]$$

Setze $A := \Omega_1 \implies \mathrm{P}[\,Y \in B\,] = Q(B)$.

q.e.d.

Soviel zur Definition der bedingten Verteilung und ihren Eigenschaften. Analog zum nicht bedingten Fall definiert man den bedingten Erwartungswert mittels der bedingten Verteilung:

Definition 3: Sei Y eine Zufallsvariable mit Werten in $(\overline{\mathbb{R}}, \overline{\mathcal{B}})$ (also $\Omega_2 = \overline{\mathbb{R}}$), so daß $\mathrm{E}[\,Y\,]$ existiert. Dann heißt $f : \Omega_1 \to \overline{\mathbb{R}}$ ein <u>*bedingter Erwartungswert von Y bzgl. X*</u>, falls ein $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$ existiert mit der Eigenschaft $f(x) = \int_{\overline{\mathbb{R}}} Q(x, dy)\, y$ für $x \in \Omega_1$.

<u>Schreibweisen:</u> $f = \mathrm{E}[\,Y \mid X = \cdot\,]$, $f(x) = \mathrm{E}[\,Y \mid X = x\,]$ und $f \in \mathbb{E}[\,Y \mid X = \cdot\,]$

Bemerkung: Man vergleiche: $\mathrm{E}[\,Y\,] = \int_{\overline{\mathbb{R}}} \mathrm{P}Y^{-1}\,(dy)\, y$

Mit Satz 1.7 ergibt sich ein

Spezialfall: Sind X, Y unabhängig, so gilt: $\mathrm{E}[Y] \in \mathbb{E}[Y \mid X = \cdot]$.

Beweis: X, Y unabhängig $\Longrightarrow \mathrm{P}Y^{-1} \in \mathbb{P}Y^{-1} \mid X = \cdot$

$$\Longrightarrow f(x) := \int_{\overline{\mathbb{R}}} \mathrm{P}Y^{-1}(dy)\, y \equiv \mathrm{E}[Y] \in \mathbb{E}[Y \mid X = \cdot]$$

q.e.d.

Der nachstehende Satz gibt einige Eigenschaften des bedingten Erwartungswertes wieder, speziell die Existenz- und Eindeutigkeitsfrage sowie eine Charakterisierung, die manche Autoren als Definition des bedingten Erwartungswertes verwenden:

Satz 1.8: Sei $\Omega_2 = \overline{\mathbb{R}}$ und es existiere $\mathrm{E}[Y]$. Dann gelten die folgenden Aussagen:

a.) Existenz: $\mathbb{E}[Y \mid X = \cdot] \neq \emptyset$

b.) Eindeutigkeit: Sei $g \in \mathbb{E}[Y \mid X = \cdot]$ und $f : \Omega_1 \to \overline{\mathbb{R}}$ eine meßbare Funktion.
Dann gilt: $f \in \mathbb{E}[Y \mid X = \cdot] \iff f = g \quad \mathrm{P}X^{-1}$-fast sicher

c.) Charakterisierung:

$$f \in \mathbb{E}[Y \mid X = \cdot] \iff \begin{cases} (1) \quad f : \Omega_1 \to \overline{\mathbb{R}} \text{ ist } \mathcal{F}_1 - \overline{\mathcal{B}}\text{-meßbar} \\ (2) \quad \int_{\Omega_1} f \, d\mathrm{P}X^{-1} \quad \text{existiert} \\ (3) \quad \int_{\{X \in A\}} Y \, d\mathrm{P} = \int_A f \, d\mathrm{P}X^{-1} \quad \forall A \in \mathcal{F}_1 \end{cases}$$

Beweis: a.) Nach Satz 1.5 existiert ein $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$. Mit $\pi_2 : \Omega_1 \times \overline{\mathbb{R}} \to \overline{\mathbb{R}}$ als Projektion auf die zweite Komponente gilt dann: $\mathrm{E}[Y]$ existiert $\iff \mathrm{E}[\pi_2 \circ (X, Y)]$ existiert

Es ist $\mathrm{E}[Y] = \mathrm{E}[\pi_2 \circ (X,Y)] \equiv \int_\Omega \pi_2 \circ (X,Y)\, d\mathrm{P} = \int_{\Omega_1 \times \overline{\mathbb{R}}} \pi_2 \, d\mathrm{P}(X,Y)^{-1}$

$$\equiv \int_{\Omega_1 \times \overline{\mathbb{R}}} \pi_2 \, d(\mathrm{P}X^{-1} \otimes Q) \quad \text{nach Definition der bedingten Verteilung.}$$

Da $\mathrm{E}[\pi_2 \circ (X,Y)]$ existiert, ist $\pi_2 \in \mathcal{L}'_{\mathrm{P}X^{-1} \otimes Q}$ und (1.4) b.) ist anwendbar:

$$\int_{\Omega_1 \times \overline{\mathbb{R}}} \pi_2 \, d(\mathrm{P}X^{-1} \otimes Q) = \int_{\Omega_1} \mathrm{P}X^{-1}(dx) \int_{\overline{\mathbb{R}}} Q(x, dy)\, \pi_2(x,y)$$

$$\equiv \int_{\Omega_1} \mathrm{P}X^{-1}(dx) \int_{\overline{\mathbb{R}}} Q(x, dy)\, y$$

Mit (1.4) b.) folgt ferner: Es existiert eine Menge $\mathcal{M} \in \mathcal{F}_1$ mit der Eigenschaft $\mathrm{P}X^{-1}[\mathcal{M}] = 1$, so daß $\int_{\overline{\mathbb{R}}} Q(x, dy)\, y$ existiert $\forall\, x \in \mathcal{M}$.

Wähle nun $Q^*(x, \cdot) := \begin{cases} Q(x, \cdot) & \text{für } x \in \mathcal{M} \\ \delta_0(\cdot) & \text{für } x \notin \mathcal{M} \end{cases}$ mit δ_0 wie in Bemerkung 3.

Dann ist Q^* Übergangswahrscheinlichkeit und es folgt mit (1.6): $Q^* \in \mathbb{P}Y^{-1} \mid X = \cdot$

Außerdem existiert dann $\int_{\overline{\mathbb{R}}} Q^*(x, dy)\, y \quad \forall x \in \Omega_1$, eine Version aus $\mathbb{E}[Y \mid X = \cdot]$ ist somit konstruiert.

b.) "$\Leftarrow$": Sei $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$ mit $g(x) = \int\limits_{\mathbb{R}} Q(x, dy)\, y$. Es gilt für die Menge $\mathcal{M} := \{ f = g \}$ nach Voraussetzung: $\mathrm{P}X^{-1}[\mathcal{M}] = 1$.

Wähle $Q^*(x, \cdot) := \begin{cases} Q(x, \cdot) & \text{für } x \in \mathcal{M} \\ \delta_{f(x)}(\cdot) & \text{für } x \notin \mathcal{M} \end{cases}$

Dann ist $Q^* \in \mathbb{P}Y^{-1} \mid X = \cdot$ $\big($ analog zu a.) $\big)$ und es gilt:

$$f(x) = \int\limits_{\mathbb{R}} Q^*(x, dy)\, y \quad \forall x \in \Omega_1, \text{ also ist } f \text{ aus } \mathbb{E}[\,Y \mid X = \cdot\,].$$

c.) "$\Rightarrow$": Sei $f \in \mathbb{E}[\,Y \mid X = \cdot\,] \Longrightarrow f(x) = \int\limits_{\mathbb{R}} Q(x, dy)\, y$ für ein $Q \in \mathbb{P}Y^{-1} \mid X = \cdot$

Aus (1.2) folgt sofort: f ist $\mathcal{F}_1 - \overline{\mathcal{B}}$-meßbar, d.h. (1).

Da π_2 wie in a.) gezeigt aus $\mathcal{L}'_{\mathrm{P}X^{-1} \otimes Q}$ ist, so ist nach (1.4)

$$f(x) = \int\limits_{\mathbb{R}} Q(x, dy)\, \pi_2(x, y) \equiv \int\limits_{\mathbb{R}} Q(x, dy)\, y \quad \text{aus } \mathcal{L}'_{\mathrm{P}X^{-1}}, \text{ d.h. (2).}$$

Für (3) hat man die Gleichungskette ($A \in \mathcal{F}_1$):

$$\int\limits_{\{X \in A\}} Y\, d\mathrm{P} \equiv \int\limits_{\Omega} (1_A \circ X) \cdot Y\, d\mathrm{P}$$

$$= \int\limits_{\Omega_1 \times \mathbb{R}} \mathrm{P}(X, Y)^{-1} \big(d(x, y)\big)\; 1_A(x) \cdot y \quad (\text{ Bildmaßformel })$$

$$\equiv \int\limits_{\Omega_1 \times \mathbb{R}} (\mathrm{P}X^{-1} \otimes Q) \big(d(x, y)\big)\; 1_A(x) \cdot y \quad (\text{ Definition der bedingten Verteilung })$$

$$= \int\limits_{\Omega_1} \mathrm{P}X^{-1}(dx)\, 1_A(x) \int\limits_{\mathbb{R}} Q(x, dy)\, y \quad (\text{ Fubini, da } 1_A(x) \cdot y \in \mathcal{L}'_{\mathrm{P}X^{-1} \otimes Q} \,)$$

$$\equiv \int\limits_{A} \mathrm{P}X^{-1}\,(dx)\, f(x)$$

"$\Leftarrow$": Nach a.) existiert $g \in \mathbb{E}[\,Y \mid X = \cdot\,]$. Mit c.) "$\Rightarrow$" für g statt f gilt: ($A \in \mathcal{F}_1$)

$$\int\limits_{\{X \in A\}} Y\, d\mathrm{P} = \int\limits_{A} g\, d\mathrm{P}X^{-1} \quad \text{und somit nach Voraussetzung:}$$

$$\int\limits_{A} f\, d\mathrm{P}X^{-1} = \int\limits_{\{X \in A\}} Y\, d\mathrm{P} = \int\limits_{A} g\, d\mathrm{P}X^{-1} \quad \forall A \in \mathcal{F}_1.$$

$\Longrightarrow f = g$ $\mathrm{P}X^{-1}$ fast sicher $\big($ W-Theorie 18.6 d.) $\big)$

Mithin gilt nach b.) "$\Leftarrow$": $f \in \mathbb{E}[\,Y \mid X = \cdot\,]$.

b.) "$\Rightarrow$": Bis auf den letzten Schluß genau wie c.) "$\Leftarrow$".

q.e.d.

Am Schluß dieses Abschnitts betrachtet man noch einige Spezialfälle der bedingten Verteilung und des bedingten Erwartungswertes:

Definition 4: Für $C \in \mathcal{F}$ definiert man $\mathrm{P}[\,C \mid X = \cdot\,] := \mathrm{E}[\,1_C \mid X = \cdot\,]$ als die *bedingte Wahrscheinlichkeit von C bzgl. X*, analog $\mathbb{P}[C \mid X = \cdot\,] := \mathbb{E}[\,1_C \mid X = \cdot\,]$.

Beachte: Die bedingte Wahrscheinlichkeit ist wohldefiniert, da $\mathrm{E}[\,1_C\,] \equiv \mathrm{P}[\,C\,]$ existiert.

Definition 5: Sei $\mathcal{G}$ Unter-σ-Algebra von $\mathcal{F}$ und $X : \Omega \to \Omega$ die Identität, welche als $\mathcal{F}-\mathcal{G}$-meßbare Abbildung aufgefaßt wird (d.h. $\Omega_1 \equiv \Omega$, $\mathcal{F}_1 \equiv \mathcal{G}$). Dann definiere für $C \in \mathcal{F}$
$\mathrm{P}[\,C \mid \mathcal{G}\,](\,\cdot\,) := \mathrm{P}[\,C \mid X = \cdot\,]$ als <u>*bedingte Wahrscheinlichkeit von C bzgl. $\mathcal{G}$*</u>.
Analog bezeichnet man $\mathrm{P}Y^{-1} \mid \mathcal{G} := \mathrm{P}Y^{-1} \mid X = \cdot$ als <u>*bedingte Verteilung von Y bzgl. $\mathcal{G}$*</u>. Ist Y eine erweitert reelle Zufallsvariable, d.h. $(\Omega_2, \mathcal{F}_2) \equiv (\overline{\mathbb{R}}, \overline{\mathcal{B}})$, und existiert $\mathrm{E}[\,Y\,]$, so heißt $\mathrm{E}[\,Y \mid \mathcal{G}\,](\,\cdot\,) := \mathrm{E}[\,Y \mid X = \cdot\,]$ <u>*bedingter Erwartungswert von Y bzgl. $\mathcal{G}$*</u>.
<u>Schreibweisen:</u> $\mathbb{E}[\,Y \mid \mathcal{G}\,] := \mathbb{E}[\,Y \mid X = \cdot\,]$, $\mathbb{P}Y^{-1} \mid \mathcal{G} := \mathbb{P}Y^{-1} \mid X = \cdot$

Bemerkung: Es gilt: $Q \in \mathbb{P}Y^{-1} \mid \mathcal{G}$

$\overset{\text{I.)}}{\iff}$ 1.) Q ist Übergangswahrscheinlichkeit von $(\Omega, \mathcal{G})$ nach $(\Omega_2, \mathcal{F}_2)$

2.) $\mathrm{P}[\,G \cap \{\,Y \in B\,\}\,] = \int\limits_G Q(\omega, B)\,\mathrm{P}(d\omega) \quad \forall G \in \mathcal{G},\ B \in \mathcal{F}_2$

$\overset{\text{II.)}}{\iff}$ 1.) und 2.)* $Q(\,\cdot\,, B)$ ist bedingte Wahrscheinlichkeit $\mathrm{P}[\,Y \in B \mid \mathcal{G}\,] \quad \forall B \in \mathcal{F}_2$

<u>Beweis:</u> I.) ist genau Definition 2 a.), es ist hier $\{\,X \in G\,\} = G \quad \forall G \in \mathcal{G}$
sowie $\mathrm{P}X^{-1}\,|_{\mathcal{G}} = \mathrm{P}\,|_{\mathcal{G}}$ ($|_{\mathcal{G}}$ bedeutet: eingeschränkt auf $\mathcal{G}$)
II.) gilt, da 2.)* $\iff$ 2.) mittels den Identitäten
$Q(\,\cdot\,, B) = \mathrm{P}[\,Y \in B \mid \mathcal{G}\,] \equiv \mathrm{P}[\,Y \in B \mid X = \cdot\,] \equiv \mathrm{E}[\,1_{\{\,Y \in B\,\}} \mid X = \cdot\,]$
Somit gilt nach (1.8) c.) :

$$\mathrm{P}[\,G \cap \{\,Y \in B\,\}\,] \equiv \int\limits_G 1_{\{\,Y \in B\,\}}\,d\mathrm{P} = \int\limits_G Q(\omega, B)\,\mathrm{P}(d\omega) \quad \forall G \in \mathcal{G},\ B \in \mathcal{F}_2.$$

Mit geringer Transferleistung läßt sich analog zu (1.8) formulieren:

Satz 1.9: a.) Existenz: $\mathbb{E}[\,Y \mid \mathcal{G}\,] \neq \emptyset$

b.) Eindeutigkeit: Sei $g \in \mathbb{E}[\,Y \mid \mathcal{G}\,]$, $f : \Omega \to \overline{\mathbb{R}}$ $\mathcal{G}-\overline{\mathcal{B}}$-meßbar. Dann gilt:

$$f \in \mathbb{E}[\,Y \mid \mathcal{G}\,] \iff f = g \quad \mathrm{P}(|_{\mathcal{G}}) - \text{fast sicher}$$

c.) Charakterisierung:

$$f \in \mathbb{E}[\,Y \mid \mathcal{G}\,] \iff \begin{cases} (1) & f : \Omega \to \overline{\mathbb{R}} \ \text{ ist } \mathcal{G}-\overline{\mathcal{B}}\text{-meßbar} \\ (2) & \int\limits_\Omega f\,d\mathrm{P} \quad \text{existiert} \\ (3) & \int\limits_G Y\,d\mathrm{P} = \int\limits_G f\,d\mathrm{P} \quad \forall\, G \in \mathcal{G} \end{cases}$$

Bemerkung: 1.) Diesen Satz beweist man mit der Vorgehensweise in obiger Bemerkung.

2.) Da in b.) f und g $\mathcal{G}-\overline{\mathcal{B}}$-meßbar sind, so sind die Aussagen [$f = g$ P-fast sicher] und [$f = g$ $\mathrm{P}\,|_{\mathcal{G}}$-fast sicher] hier äquivalent. (leicht)

Definition 6: Mit $\sigma(X) := X^{-1}(\mathcal{F}_1) \equiv \{\,\{\,X \in A\,\} : A \in \mathcal{F}_1\,\}$ definiert man
$\mathbb{E}[\,Y \mid X\,] := \mathbb{E}[\,Y \mid \sigma(X)\,]$ als die <u>*Menge der bedingten Erwartungswerte von Y bzgl. $\sigma(X)$*</u>.

Bemerkung: Beachte, daß $\sigma(X)$ eine σ-Algebra auf $\mathcal{F}$ ist, daher macht diese Definition Sinn.
Es gilt: $\mathbb{E}[\,Y \mid X\,] = \{\,f \circ X : f \in \mathbb{E}[\,Y \mid X = \cdot\,]\,\}$ (vgl. Übung)
Dies zeigt man mit Hilfe der Sätze (1.8) c.) bzw. (1.9) c.) unter Zuhilfenahme der Aussage, daß eine Funktion g genau dann $\sigma(X)-\overline{\mathcal{B}}$-meßbar ist, wenn eine meßbare Funktion f existiert, welche die Gleichung $g = f \circ X$ erfüllt.

Die letzten Sätze dieses Abschnitts erfassen einige wichtige Eigenschaften der bedingten Erwartung. Speziell werden die schon aus dem nicht bedingten Fall her bekannten Aussagen über monotone Konvergenz bzw. das Lemma von Fatou verallgemeinert.

Satz 1.10: Sei $\mathcal{G} \subset \mathcal{F}$ eine Unter-σ-Algebra sowie Y, Y' quasiintegrierbare erweitert reelle Zufallsvariablen. $\mathrm{E}[\,\cdot \mid \mathcal{G}\,]$ stehe für eine beliebige Version. Dann gelten:

a.) Ist $\mathrm{E}[\,|\,Y\,|\,] < \infty$, so existiert eine reellwertige Version $f \in \mathbb{E}[\,Y \mid \mathcal{G}\,]$.

b.) Ist $Y \leq Y'$ fast sicher $\Longrightarrow$ $\mathrm{E}[\,Y \mid \mathcal{G}\,] \leq \mathrm{E}[\,Y' \mid \mathcal{G}\,]$ fast sicher. (Isotonie)
Ist speziell $Y \geq r \in \mathbb{R}$, so existiert ein $f \in \mathbb{E}[\,Y \mid \mathcal{G}\,]$ mit $f \geq r$.

c.) Gilt i.) $\mathrm{E}[\,|\,Y\,|\,] < \infty$, $\mathrm{E}[\,|\,Y'\,|\,] < \infty$, $\alpha, \beta \in \mathbb{R}$
oder ii) $Y \geq 0$, $Y' \geq 0$, $\alpha, \beta \geq 0$, dann gilt:
$\mathrm{E}[\,\alpha Y + \beta Y' \mid \mathcal{G}\,] = \alpha \mathrm{E}[\,Y \mid \mathcal{G}\,] + \beta \mathrm{E}[\,Y' \mid \mathcal{G}\,]$ fast sicher. (Linearität)

d.) Ist Y $\mathcal{G}$-meßbar, so ist $Y \in \mathbb{E}[\,Y \mid \mathcal{G}\,]$.

e.) Sei $\mathcal{G}'$ Unter-σ-Algebra von $\mathcal{G}$. Dann gilt:
$\mathrm{E}\big[\,\mathrm{E}[\,Y \mid \mathcal{G}\,] \mid \mathcal{G}'\,\big] = \mathrm{E}[\,Y \mid \mathcal{G}'\,]$ fast sicher. (Sukzessive Konditionierung)

f.) Ist $\mathrm{E}[\,|\,Y\,|\,] < \infty$, $g : \mathbb{R} \to \mathbb{R}$ konvex und $\mathrm{E}[\,|\,g \circ Y\,|\,] < \infty$, so gilt:
$g(\mathrm{E}[\,Y \mid \mathcal{G}\,]) \leq \mathrm{E}[\,g \circ Y \mid \mathcal{G}\,]$ fast sicher. (Jensen'sche Ungleichung)

g.) Seien Y, $Y \cdot Y'$ integrierbar oder $Y, Y' \geq 0$. Ist ferner Y' $\mathcal{G}$-meßbar, so gilt:
$\mathrm{E}[\,Y \cdot Y' \mid \mathcal{G}\,] = Y' \cdot \mathrm{E}[\,Y \mid \mathcal{G}\,]$. (Faktorisierung)

h.) Sei Ω_2 beliebig, d.h. Y nicht notwendigerweise erweitert reell. Des weiteren seien $T : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ $\mathcal{F}_1 \otimes \mathcal{F}_2 - \overline{\mathcal{B}}$-meßbar, X $\mathcal{G} - \mathcal{F}_1$-meßbar, $\mathrm{E}[\,T(X,Y)\,]$ existent und $\mathbb{P}Y^{-1} \mid \mathcal{G} \neq \emptyset$. Existiert ferner für $Q \in \mathbb{P}Y^{-1} \mid \mathcal{G}$ und alle $(x, \omega) \in \Omega_1 \times \Omega_2$
$h(x,\omega) := \int\limits_{\Omega_2} Q(\omega, dy)\, T\big(X(\omega), y\big)$, so ist h $\mathcal{F}_1 \otimes \mathcal{F}_2$-meßbar und es gilt
$\mathrm{E}[\,T(X,Y) \mid \mathcal{G}\,](\omega) = h\big(X(\omega), \omega\big)$ P-fast sicher.

Beweise/Bemerkungen: (vgl. auch Übung)

Die Aussagen a.) - f.) gelten auch für Erwartungswerte $\mathrm{E}[\,\cdot \mid X = \cdot\,]$ anstelle von $\mathrm{E}[\,\cdot \mid \mathcal{G}\,]$.
Im Beweis der Aussage a.) gilt für jede Version, daß sie fast sicher reell ist. Man ändere sie auf der Nullmenge, wo sie die Werte $+\infty$ bzw. $-\infty$ annimmt, ab und gebe ihr dort reelle Werte.
In b.) erhält man analog Versionen, die fast sicher $\geq r$ sind. Man gebe ihnen auf der Menge $\{\, f < r \,\}$ den Wert r.
Die Aussagen c.) - e.) beweist man leicht.

Beweis von f.):
$$g(\mathrm{E}[\,Y \mid \mathcal{G}\,]) \overset{\text{h.)}}{=} g\Big(\int\limits_{\mathbb{R}} Q(\,\cdot\,, dy)\, y\Big) \quad \text{für } Q \in \mathbb{P}Y^{-1} \mid \mathcal{G}, \text{ h.) angewandt auf } \pi_2$$
$$\leq \int\limits_{\overline{\mathbb{R}}} Q(\,\cdot\,, dy)\, g(y) \quad \text{nach Jensen (vgl. W-Theorie 14.13)}$$
$$\overset{\text{h.)}}{=} \mathrm{E}[\,g \circ Y \mid \mathcal{G}\,] \quad \text{wie oben}$$

Beweis von g.): Beweisprinzip für Integrale (maßtheoretische Induktion):

Für $Y' := 1_A,\ A \in \mathcal{G}$ und $Y \geq 0$ gilt für alle $G \in \mathcal{G}$:

$$\int\limits_G 1_A \cdot \mathrm{E}[\,Y \mid \mathcal{G}\,]\, d\mathrm{P} = \int\limits_{G\cap A} \mathrm{E}[\,Y \mid \mathcal{G}\,]\, d\mathrm{P}$$

$$= \int\limits_{G\cap A} Y\, d\mathrm{P} \quad \text{nach Satz 1.9 c.)}$$

$$= \int\limits_G 1_A \cdot Y\, d\mathrm{P} \equiv \int\limits_G \mathrm{E}[\,1_A \cdot Y \mid \mathcal{G}\,]\, d\mathrm{P} \quad \text{nach d.)}$$

Die restlichen Schritte sind leicht.

Beweis von h.): Mit $X^0 : (\Omega, \mathcal{F}) \to (\Omega, \mathcal{G})$ als Identität und $G \in \mathcal{G}$ gilt:

$$\int\limits_G T(X,Y)\, d\mathrm{P} \equiv \int\limits_\Omega T(X \circ X^0, Y) \cdot 1_G \circ X^0\, d\mathrm{P}$$

$$= \int\limits_{\Omega\times\Omega_2} T\big(X(\omega), y\big) \cdot 1_G(\omega)\, \mathrm{P}(X^0, Y)^{-1}\big(d(\omega, y)\big) \quad \text{(Bildmaßformel)}$$

$$= \int\limits_{\Omega\times\Omega_2} T\big(X(\omega), y\big) \cdot 1_G(\omega) \big(\mathrm{P}(X^0)^{-1} \otimes Q\big)\, \big(d(\omega, y)\big)$$

(für $Q \in \mathbb{P}Y^{-1} \mid \mathcal{G}$)

$$= \int\limits_G \mathrm{P}(X^0)^{-1}(d\omega) \int\limits_{\Omega_2} Q(\omega, dy)\, T\big(X(\omega), y\big) \quad \text{nach Fubini}$$

$$\equiv \int\limits_G \mathrm{P}(d\omega)\, h\big(X(\omega), \omega\big)$$

□

Satz 1.11: Seien $(Y_n)_{n\in\mathbb{N}_0}$ erweitert reelle Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit $\mathrm{E}[\,Y_0^-\,] < \infty$.
Ist dann $\mathcal{G}$ eine Unter-σ-Algebra von $\mathcal{F}$, so gelten:

a.) Monotone Konvergenz: Ist $Y_n \leq Y_{n+1}$ fast sicher $\forall\, n \in \mathbb{N}_0$, so ist
$\mathrm{E}[\lim\limits_{n\to\infty} Y_n \mid \mathcal{G}\,] = \lim\limits_{n\to\infty} \mathrm{E}[\,Y_n \mid \mathcal{G}\,]$ fast sicher.

b.) Lemma von Fatou: Gilt $Y_n \geq Y_0$ fast sicher $\forall\, n \in \mathbb{N}$, so hat man
$\mathrm{E}[\,\underline{\lim}_{n\to\infty}\, Y_n \mid \mathcal{G}\,] \leq \underline{\lim}_{n\to\infty}\, \mathrm{E}[\,Y_n \mid \mathcal{G}\,]$ fast sicher.

Beweis: a.) Man betrachte dazu W-Theorie 10.1, (1.8) c.), 1.10 b.) sowie W-Theorie 8.14 .

b.) Der Beweis erfolgt analog zu dem im nicht bedingten Fall (siehe hierfür W-Theorie 10.6).

□

Soweit die Ausführungen zu bedingten Verteilungen und bedingten Erwartungswerten.
Es folgt nun eine Erweiterung/Verallgemeinerung der Maßtheorie durch die Einführung von Maßen auf abzählbar unendlichen Produkträumen.

Abschnitt 2: Maße auf unendlichen Produkträumen

Durch die bisherige Theorie der Produkte endlich vieler Wahrscheinlichkeitsräume war es möglich, für die Durchführung von $n \in \mathbb{N}$ unabhängigen Experimenten mit zufälligem Ausgang ein stochastisches Modell zu erstellen (vgl. W-Theorie 13.5):

Es seien $n \in \mathbb{N}$ unabhängige Experimente E_i durch Wahrscheinlichkeitsräume $(\Omega_i, \mathcal{F}_i, \mathrm{P}_i)$, $1 \le i \le n$, beschrieben (d.h. es treten bei einem Versuch E_i Werte aus Ω_i auf). Man sucht einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ und Zufallsvariablen X_i auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in $(\Omega_i, \mathcal{F}_i)$, $1 \le i \le n$, so daß die X_i jeweils das Experiment E_i beschreiben. Das heißt für die X_i:

1.) $X_1, \ldots, X_n$ sind voneinander unabhängig

2.) $X_i \sim \mathrm{P}_i$, d.h. $\mathrm{P}X_i^{-1} = \mathrm{P}_i$

Als Lösung des Problems erhält man: $\Omega = \bigtimes_{i=1}^{n} \Omega_i$, $\mathcal{F} = \bigotimes_{i=1}^{n} \mathcal{F}_i$, $\mathrm{P} = \bigtimes_{i=1}^{n} \mathrm{P}_i$ und $X_i : \Omega \to \Omega_i$ als Projektion auf die i-te Komponente, $1 \le i \le n$.

Mit dieser Theorie erfaßt man jedoch nicht den Fall, daß man eine Folge $(E_i)_{i \in \mathbb{N}}$ von Experimenten hat. Handelt es sich in obigem endlichen Fall noch um ein n-stufiges Zufallsexperiment, wo man für jedes n ein anderes stochastisches Modell benötigt, so möchte man für den jetzigen Fall ein einheitliches Modell mit einem einzigen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ konstruieren.
Die Konstruktion eines solchen Wahrscheinlichkeitsraumes wird ein wesentliches Ziel dieses Abschitts sein. Zunächst jedoch noch einige

<u>Vorüberlegungen:</u> Seien Meßräume $(\Omega_i, \mathcal{F}_i)$ und Projektionen $\pi_i : \bigtimes_{j=1}^{\infty} \Omega_j \to \Omega_i$ gegeben, $i \in \mathbb{N}$.

Man definiert sich folgende nützliche Mengensysteme:

$$\widehat{\mathcal{F}_n} := \{ \Omega_1 \times \ldots \times \Omega_{n-1} \times A_n \times \Omega_{n+1} \times \ldots \mid A_n \in \mathcal{F}_n \} = \pi_n^{-1}(\mathcal{F}_n) \equiv \sigma(\pi_n)$$

$$\widehat{\bigtimes_{i=1}^{n} \mathcal{F}_i} := \{ A_1 \times \ldots \times A_n \times \Omega_{n+1} \times \ldots \mid A_i \in \mathcal{F}_i,\ 1 \le i \le n \} = (\pi_1, \ldots, \pi_n)^{-1}\big(\bigtimes_{i=1}^{n} \mathcal{F}_i\big)$$

sowie
$$\widehat{\bigotimes_{i=1}^{n} \mathcal{F}_i} := \{ B_n \times \Omega_{n+1} \times \ldots \mid B_n \in \bigotimes_{i=1}^{n} \mathcal{F}_i \} = (\pi_1, \ldots, \pi_n)^{-1}\big(\bigotimes_{i=1}^{n} \mathcal{F}_i\big) \equiv \sigma\big((\pi_1, \ldots, \pi_n)\big)$$

<u>Schreibweisen:</u> $\sigma(Y_t, t \in T) := \sigma\big(\bigcup_{t \in T} \sigma(Y_t)\big)$, wo $\sigma(Y_t)$ schon in Abschnitt 1 definiert.

Es gilt: $\sigma(Y_t, t \in T) = \sigma\big((Y_t, t \in T)\big)$ (vgl. Übung)

Nun ein Lemma, welches die Struktur von $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$ näher beschreibt, was für die nächsten Beweise hilfreich sein wird:

<u>Lemma 2.1:</u> Es gelten die Identitäten

$$\underbrace{\bigotimes_{i=1}^{\infty} \mathcal{F}_i}_{\mathrm{I}} = \sigma\big(\underbrace{\bigcup_{n=1}^{\infty} \widehat{\mathcal{F}_n}}_{\mathrm{II}}\big) = \sigma\big(\underbrace{\bigcup_{n=1}^{\infty} \widehat{\bigtimes_{i=1}^{n} \mathcal{F}_i}}_{\mathrm{III}}\big) = \sigma\big(\underbrace{\bigcup_{n=1}^{\infty} \widehat{\bigotimes_{i=1}^{n} \mathcal{F}_i}}_{\mathrm{IV}}\big)$$

Beweis: Da $\widehat{\mathcal{F}_n} \subset \widehat{\bigtimes_{i=1}^{n} \mathcal{F}_i} \subset \widehat{\bigotimes_{i=1}^{n} \mathcal{F}_i}$ folgt sofort II$\subset$III$\subset$IV.

$\pi_n : \bigtimes_{i=1}^{\infty} \Omega_i \to \Omega_n$ ist nach obigem II $-$ $\mathcal{F}_n$-meßbar. Mit W-Theorie 8.10 folgt daher:

$(\pi_1, \pi_2, \ldots)$ ist II$-$I-meßbar. Da aber $(\pi_1, \pi_2, \ldots) \equiv \mathrm{id}_{\bigtimes_{i=1}^{\infty} \Omega_i}$ folgt: I$\subset$II

Des weiteren ist π_n I$-\mathcal{F}_n$-meßbar, da $\widehat{\mathcal{F}_n} \subset$ I.

$\Longrightarrow$ $(\pi_1, \ldots, \pi_n)$ ist I$-\bigotimes_{i=1}^{n} \mathcal{F}_i$-meßbar nach W-Theorie 8.10

$\Longrightarrow$ $\widehat{\bigotimes_{i=1}^{n} \mathcal{F}_i} \equiv (\pi_1, \ldots, \pi_n)^{-1}(\bigotimes_{i=1}^{n} \mathcal{F}_i) \subset$ I $\Longrightarrow$ IV$\subset$I.

q.e.d.

Eine Verallgemeinerung dieses Lemmas stellt dar:

Lemma 2.2: Seien Meßräume $(\Omega, \mathcal{F})$, $(\Omega_i, \mathcal{F}_i)_{i \in \mathbb{N}}$ und Abbildungen $X_i : \Omega \to \Omega_i$, $i \in \mathbb{N}$, gegeben. Dann ist $\mathcal{A} := \bigcup_{n=1}^{\infty} \sigma(X_1, \ldots, X_n)$ eine $\sigma(X_1, X_2, \ldots)$ erzeugende Algebra.

Beweis: 1.) $\mathcal{A}$ ist eine Algebra, da $\sigma(X_1, \ldots, X_n)$ Algebra $\quad \forall n \in \mathbb{N}$ (leicht).

2.) Da $\sigma(X_1, \ldots, X_n) \subset \sigma(X_1, X_2, \ldots) \quad \forall n \in \mathbb{N}$, folgt sofort $\sigma(\mathcal{A}) \subset \sigma(X_1, X_2, \ldots)$.
Ferner ist X_i $\sigma(\mathcal{A}) - \mathcal{F}_i$-meßbar, d.h. $\sigma(X_i) \subset \sigma(\mathcal{A}) \quad \forall i \in \mathbb{N}$.
Daher wieder mit W-Theorie 8.10 : $\sigma\big((X_1, X_2, \ldots)\big) \equiv \sigma(X_1, X_2, \ldots) \subset \sigma(\mathcal{A})$

q.e.d.

Beispiel: $\bigcup_{n=1}^{\infty} \widehat{\bigotimes_{i=1}^{n} \mathcal{F}_i}$ ist eine Algebra, welche $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$ erzeugt.

Beweis: Setze $X_i := \pi_i$. Der Rest ist nach obigem klar.

Nun der angekündigte wichtige Satz, der ein Wahrscheinlichkeitsmaß P auf $(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{F}_i)$ liefert.

Diese Version des Produktmaßsatzes für den abzählbar unendlichen Fall wird nur für Wahrscheinlichkeitsmaße gezeigt, die nachstehende Konstruktion ist nicht auf beliebige Maße ausdehnbar. Der Fall überabzählbarer Produkträume wird hier nicht behandelt.

Zunächst jedoch noch eine weitere

Vorüberlegung: Seien Meßräume $(\Omega_i, \mathcal{F}_i)_{i \in \mathbb{N}}$ gegeben. Q_1 sei ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_1$ und Q_{n+1} eine Übergangswahrscheinlichkeit von $\bigtimes_{i=1}^{n} \Omega_i$ nach Ω_{n+1}, $n \in \mathbb{N}$. Dann ist die Setzung

$\bigotimes_{i=1}^{n+1} Q_i := (\bigotimes_{i=1}^{n} Q_i) \otimes Q_{n+1}$ wohldefiniert, da man induktiv zeigt:

$\bigotimes_{i=1}^{n} Q_i$ ist ein Wahrscheinlichkeitsmaß auf $\bigotimes_{i=1}^{n} \mathcal{F}_i$. Die Verknüpfung

$(\bigotimes_{i=1}^{n} Q_i) \otimes Q_{n+1}$ ist daher nach (1.3) definiert, es gilt für $A_i \in \mathcal{F}_i$, $1 \leq i \leq n$:

$$\bigotimes_{i=1}^{n} Q_i \,[\, A_1 \times \ldots \times A_n \,] = \int_{A_1} Q_1\,(d\omega_1) \int_{A_2} Q(\omega_1; d\omega_2) \ldots$$
$$\ldots \int_{A_{n-1}} Q_{n-1}(\omega_1, \ldots, \omega_{n-2}; d\omega_{n-1})\, Q_n(\omega_1, \ldots, \omega_{n-1}; A_n)$$

Satz 2.3: (Ionescu-Tulcea)

Seien Meßräume $(\Omega_i, \mathcal{F}_i)_{i\in\mathbb{N}}$ gegeben, Q_1 sei ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_1$ und Q_{n+1} eine Übergangswahrscheinlichkeit von $\bigtimes_{i=1}^{n} \Omega_i$ nach Ω_{n+1}, $n \in \mathbb{N}$. Ferner seien $\pi_j : \bigtimes_{i=1}^{\infty} \Omega_i \to \Omega_j$ die Projektionen, $j \in \mathbb{N}$. Dann gelten:

a.) Es existiert genau ein Wahrscheinlichkeitsmaß P auf $(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{F}_i)$ mit der Eigenschaft

$$\begin{aligned}
\mathrm{P}[A_1 \times \ldots \times A_n \times \Omega_{n+1} \times \ldots] &:= \bigotimes_{i=1}^{n} Q_i\,[A_1 \times \ldots \times A_n] \\
&\equiv \int_{A_1} Q_1(d\omega_1) \int_{A_2} Q(\omega_1; d\omega_2) \ldots \\
&\ldots \int_{A_{n-1}} Q_{n-1}(\omega_1, \ldots, \omega_{n-2}; d\omega_{n-1})\, Q_n(\omega_1, \ldots, \omega_{n-1}; A_n)
\end{aligned}$$

$\forall A_i \in \mathcal{F}_i \quad \forall n \in \mathbb{N}$. Schreibweise: $\mathrm{P} =: \bigotimes_{i=1}^{\infty} Q_i$

b.) P ist das einzige Wahrscheinlichkeitsmaß auf $(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{F}_i)$ mit $\mathrm{P}(\pi_1, \ldots, \pi_n)^{-1} = \bigotimes_{i=1}^{n} Q_i$,

d.h. $\bigotimes_{i=1}^{n} Q_i$ ist die gemeinsame Verteilung der π_i, $1 \le i \le n$, bzgl. P.

c.) P ist das einzige Wahrscheinlichkeitsmaß auf $(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{F}_i)$ mit $Q_1 = \mathrm{P}\pi_1^{-1}$ und

$Q_{n+1} = \mathrm{P}\pi_{n+1}^{-1} \mid (\pi_1, \ldots, \pi_n) = \cdot$, d.h. Q_{n+1} ist bedingte Verteilung von π_{n+1} bzgl. $(\pi_1, \ldots, \pi_n)$ unter P.

d.) Berechnung von Integralen:

Für jede Abbildung $f : \bigtimes_{i=1}^{n} \Omega_i \to \overline{\mathbb{R}}$ mit $f \circ (\pi_1, \ldots, \pi_n) \in \mathcal{L}'_{\mathrm{P}}$ gilt:

$$\begin{aligned}
\int_{\bigtimes_{i=1}^{\infty} \Omega_i} f \circ (\pi_1, \ldots, \pi_n)\, d\mathrm{P} &= \int_{\bigtimes_{i=1}^{n} \Omega_i} f \; d\bigotimes_{i=1}^{n} Q_i \\
&= \int_{\Omega_1} Q_1(d\omega_1) \int_{\Omega_2} \ldots \int_{\Omega_n} Q_n(\omega_1, \ldots, \omega_{n-1}; d\omega_n)\, f(\omega_1, \ldots, \omega_n)
\end{aligned}$$

Beweis: a.) Dieser Teil wird nun mit der Vorgehensweise aus Abschnitt 0 bewiesen.

I.) Die Setzung $\bigotimes_{i=1}^{n+1} Q_i := (\bigotimes_{i=1}^{n} Q_i) \otimes Q_{n+1}$ ist wohldefiniert (siehe oben).

II.) Definiere $\mathcal{A}_n := \sigma(\pi_1, \ldots, \pi_n)$. Dann ist $\mathcal{A}_n$ σ-Algebra auf $\bigtimes_{i=1}^{\infty} \Omega_i$ und es gilt

$\mathcal{A}_n \subset \mathcal{A}_{n+1}$ (d.h. $(\mathcal{A}_n)_{n\in\mathbb{N}}$ ist eine Filterung in $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$, siehe später)

Nach obigen Betrachtungen war $\mathcal{A}_n \equiv \sigma(\pi_1, \ldots, \pi_n) = \widehat{\bigotimes_{i=1}^{n}} \mathcal{F}_i$. Daher wird durch die Setzung $\mathrm{P}_n[\,B' \times \Omega_{n+1} \times \ldots] := \bigotimes_{i=1}^{n} Q_i\,[\,B'\,]$, wo $B' \in \bigotimes_{i=1}^{n} \mathcal{F}_i$, ein Wahrscheinlichkeitsmaß auf $\mathcal{A}_n$ definiert.

III.) $\mathcal{A} := \bigcup_{n=1}^{\infty} \mathcal{A}_n$ ist nach (2.2) eine $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$ erzeugende Algebra.

Auf dieser Algebra (= Ring mit $\bigtimes_{i=1}^{\infty} \Omega_i$) definiert man nun einen Inhalt P_∞:

IV.) Definiere $\mathrm{P}_\infty : \mathcal{A} \to [\,0,1\,]$ durch $\mathrm{P}_\infty[\,B\,] := \mathrm{P}_n[\,B\,]$, falls $B \in \mathcal{A}_n$.

Beh.: P_∞ ist wohldefiniert

Beweis: Da $(\mathcal{A}_n)_{n\in\mathbb{N}}$ Filterung gilt: $B \in \mathcal{A}_m \implies B \in \mathcal{A}_n \quad \forall\, n \geq m$

Sei daher $B = B' \times \Omega_{n+1} \times \ldots = B'' \times \Omega_{m+1} \times \ldots$, o.E. $n > m$

$\implies B' = B'' \times \Omega_{m+1} \times \ldots \times \Omega_n$

$$\implies \mathrm{P}_n[\,B\,] \equiv \bigotimes_{i=1}^{n} Q_i\,[\,B'\,] \equiv \bigotimes_{i=1}^{n} Q_i\,[\,B'' \times \Omega_{m+1} \times \ldots \times \Omega_n\,]$$

$$\stackrel{(+)}{=} \bigotimes_{i=1}^{m} Q_i\,[\,B''\,] \equiv \mathrm{P}_m[\,B\,]$$

(+) berechnet man leicht mit (1.3) oder (1.4). Dabei wird benutzt, daß $Q_i\,(\,\cdot\,, \Omega_i) = 1$ ist. Hier versagt die Konstruktion für allgemeine Maße.

V.) P_∞ ist ein Inhalt auf $\mathcal{A}$, da gilt:

$\mathrm{P}_\infty[\,\emptyset\,] \equiv \mathrm{P}_1[\,\emptyset \times \Omega_2 \times \ldots] \equiv Q_1\,[\,\emptyset\,] = 0$

Sind $A, B \in \mathcal{A}$ disjunkt, $A \in \mathcal{A}_n$, $B \in \mathcal{A}_m$ (o.E. $m \leq n$)

$\implies A \cup B \in \mathcal{A}_n$, da $(\mathcal{A}_n)_{n\in\mathbb{N}}$ Filterung

$\implies \mathrm{P}_\infty[\,A \cup B\,] \equiv \mathrm{P}_n[\,A \cup B\,] = \mathrm{P}_n[\,A\,] + \mathrm{P}_n[\,B\,]$ da P_n ein Maß und $B \in \mathcal{A}_n$

$\equiv \mathrm{P}_\infty[\,A\,] + \mathrm{P}_\infty[\,B\,]$

Als nächsten Schritt benötigt man (vgl. Abschnitt 0) das Vorliegen eines Prämaßes auf $\mathcal{A}$, dies zeigt

VI.) P_∞ ist Prämaß auf $\mathcal{A}$. Zum Beweis dieser Aussage benutzt man das Kriterium aus (0.4), d.h. es ist zu zeigen: P_∞ ist $\emptyset$-stetig.

Annahme: $(B_n)_{n\in\mathbb{N}} \subset \mathcal{A}$ ist antiton und $\lim_{n\to\infty} \mathrm{P}_\infty[\,B_n\,] > 0$

Dann ist zu zeigen: $\bigcap_{n=1}^{\infty} B_n \neq \emptyset$.

Sei dazu o.E. $B_n \in \mathcal{A}_n$, d.h. $B_n = B_n' \times \Omega_{n+1} \times \ldots$ mit $B_n' \in \bigotimes_{i=1}^{n} \mathcal{F}_i$

[Gilt nämlich $B_n \in \mathcal{A}_{k_n}$, so setze $m_n := \max\{\,k_1, \ldots, k_n\,\} + n$.

$\implies B_n \in \mathcal{A}_{m_n} \subset \mathcal{A}_{m_{n+1}}$, da $k_n \leq m_n < m_{n+1}$

$$\text{Setze nun } \tilde{B}_m := \begin{cases} B_n & \text{falls } m_n \leq m < m_{n+1} \\ \bigtimes_{i=1}^{\infty} \Omega_i & \text{falls } m < m_1 \end{cases}$$

$\implies \tilde{B}_m \in \mathcal{A}_m$, $\bigcap_{m=1}^{\infty} \tilde{B}_m = \bigcap_{n=1}^{\infty} B_n$ und $\lim_{m\to\infty} \mathrm{P}_\infty[\,\tilde{B}_m\,] = \lim_{n\to\infty} \mathrm{P}_\infty[\,B_n\,]$, da die Auffüllung der Folge $(B_n)_{n\in\mathbb{N}}$ zu $(\tilde{B}_n)_{n\in\mathbb{N}}$ deren Grenzwertverhalten nicht ändert.]

(i) Es ist $\bigotimes_{i=1}^{n} Q_i\,[\,B'_n\,] = \int\limits_{\Omega_1} Q_1\,(d\omega_1) \bigotimes_{i=2}^{n} Q_i\,(\omega_1; B'_n)$

$$= \int\limits_{\Omega_1} Q_1\,(d\omega_1) \int\limits_{\Omega_2} Q_2\,(\omega_1; d\omega_2) \bigotimes_{i=3}^{n} Q_i\,(\omega_1, \omega_2; B'_n)$$

wo $\bigotimes_{i=j}^{n} Q_i\,(\omega_1, \ldots, \omega_{j-1}; B'_n) := \int\limits_{\Omega_j} Q_j\,(\omega_1, \ldots, \omega_{j-1}; d\omega_j) \int\limits_{\Omega_{j+1}} \ldots$

$$\ldots \int\limits_{\Omega_n} Q_n\,(\omega_1, \ldots, \omega_{n-1}; d\omega_n)\, 1_{B'_n}(\omega_1, \ldots \omega_n).$$

Mit $B_{n+1} = B'_{n+1} \times \Omega_{n+2} \times \ldots$, wo $B'_{n+1} \in \bigotimes_{i=1}^{n+1} \mathcal{F}_i$, gilt:

$B_n \supset B_{n+1} \iff B'_n \times \Omega_{n+1} \supset B'_{n+1}$

Es folgt: $\quad (++) \quad \int\limits_{\Omega_{n+1}} Q_{n+1}\,(\,\cdot\,; d\omega_{n+1})\, 1_{B'_{n+1}}(\,\cdot\,, \omega_{n+1}) \leq 1_{B'_n}(\,\cdot\,)$

Daraus ergibt sich:

(ii) $1 \geq 1_{B'_j}(\,\cdot\,) \geq \bigotimes_{i=j+1}^{n} Q_i\,(\,\cdot\,; B'_n) \geq \bigotimes_{i=j+1}^{n+1} Q_i\,(\,\cdot\,; B'_{n+1})$ für $j \in \mathbb{N}$, da

$$\bigotimes_{i=j+1}^{n+1} Q_i\,(\,\cdot\,; B'_{n+1}) \equiv \int\limits_{\Omega_{j+1}} Q_{j+1}\,(\,\cdot\,; d\omega_{j+1}) \int\limits_{\Omega_{j+2}} \ldots$$

$$\ldots \int\limits_{\Omega_{n+1}} Q_{n+1}\,(\,\cdot\,, \omega_{j+1}, \ldots, \omega_n; d\omega_{n+1})\, 1_{B'_{n+1}}(\,\cdot\,, \omega_{j+1}, \ldots, \omega_{n+1})$$

$$\stackrel{(++)}{\leq} \int\limits_{\Omega_{j+1}} Q_{j+1}\,(\,\cdot\,; d\omega_{j+1}) \int\limits_{\Omega_{j+2}} \ldots$$

$$\ldots \int\limits_{\Omega_n} Q_n\,(\,\cdot\,, \omega_{j+1}, \ldots, \omega_{n-1}; d\omega_n)\, 1_{B'_n}(\,\cdot\,, \omega_{j+1}, \ldots, \omega_n)$$

$$\equiv \bigotimes_{i=j+1}^{n} Q_i\,(\,\cdot\,; B'_n) \leq \ldots \leq \int\limits_{\Omega_{j+1}} Q_{j+1}\,(\,\cdot\,, d\omega_{j+1})\, 1_{B'_{j+1}}(\,\cdot\,, \omega_{j+1})$$

$$\leq 1_{B'_j}(\,\cdot\,)$$

Mithin folgt aus (ii) mit monotoner Konvergenz:

$$0 < \lim_{n\to\infty} \bigotimes_{i=1}^{n} Q_i\,[\,B'_n\,] \equiv \lim_{n\to\infty} \int\limits_{\Omega_1} Q_1(d\omega_1) \bigotimes_{i=2}^{n} Q_i\,(\omega_1; B'_n) = \int\limits_{\Omega_1} Q_1(d\omega_1) \lim_{n\to\infty} \bigotimes_{i=2}^{n} Q_i\,(\omega_1; B'_n)$$

$\Longrightarrow \exists\ \overline{\omega}_1 \in \Omega_1$ mit $\lim_{n\to\infty} \bigotimes_{i=2}^{n} Q_i\,(\overline{\omega}_1; B'_n) > 0$

$$0 < \lim_{n\to\infty} \bigotimes_{i=2}^{n} Q_i\,(\overline{\omega}_1; B'_n) \equiv \lim_{n\to\infty} \int\limits_{\Omega_2} Q_2\,(\overline{\omega}_1; d\omega_2) \bigotimes_{i=3}^{n} Q_i\,(\overline{\omega}_1, \omega_2; B'_n)$$

$$= \int\limits_{\Omega_2} Q_2\,(\overline{\omega}_1; d\omega_2) \lim_{n\to\infty} \bigotimes_{i=3}^{n} Q_i\,(\overline{\omega}_1, \omega_2; B'_n)$$

$\Longrightarrow \exists\ \overline{\omega}_2 \in \Omega_2$ mit $\lim_{n\to\infty} \bigotimes_{i=3}^{n} Q_i\,(\overline{\omega}_1, \overline{\omega}_2; B'_n) > 0$

Induktiv: $\exists\ (\overline{\omega}_n)_{n\in\mathbb{N}} \in \mathop{\times}\limits_{i=1}^{\infty} \Omega_i$ mit

$$0 < \lim_{n\to\infty} \bigotimes_{i=j+1}^{n} Q_i\,(\overline{\omega}_1,\ldots,\overline{\omega}_j\,;B_n') \le 1_{B_j'}(\overline{\omega}_1,\ldots,\overline{\omega}_j) \quad \forall j\in\mathbb{N}$$

$$\Longrightarrow (\overline{\omega}_1,\overline{\omega}_2,\ldots)\in B_j \quad \forall j\in\mathbb{N} \;\Longrightarrow\; \bigcap_{j=1}^{\infty} B_j \ne \emptyset$$

Man hat also ein Prämaß auf $\mathcal{A}$, der Maßfortsetzungssatz ist anwendbar:

VII.) P_∞ kann nach (0.7) zu einem Maß P auf $\sigma(\mathcal{A}) = \bigotimes\limits_{i=1}^{\infty} \mathcal{F}_i$ fortgesetzt werden.

VIII.) Da $P|_{\mathcal{A}_n} \equiv P_\infty\,|_{\mathcal{A}_n} \equiv P_n$ ist P normiert, d.h. P ist ein Wahrscheinlichkeitsmaß.

Wegen $A_1\times\ldots\times A_n\times\Omega_{n+1}\times\ldots\in\mathcal{A}_n$ falls $A_i\in\mathcal{F}_i$ gilt:

$$\begin{aligned} P[\,A_1\times\ldots\times A_n\times\Omega_{n+1}\times\ldots] &\equiv P_n[\,A_1\times\ldots\times A_n\times\Omega_{n+1}\times\ldots] \\ &\equiv \bigotimes_{i=1}^{n} Q_i\,[\,A_1\times\ldots\times A_n\,] \end{aligned}$$

d.h. P besitzt die in a.) geforderte Eigenschaft.

Zur Eindeutigkeit:

IX.) Es war nach (2.1) $\sigma(\bigcup\limits_{n=1}^{\infty} \widehat{\mathop{\times}\limits_{i=1}^{n} \mathcal{F}_i}) = \bigotimes\limits_{i=1}^{\infty} \mathcal{F}_i$ und $\bigcup\limits_{n=1}^{\infty} \widehat{\mathop{\times}\limits_{i=1}^{n} \mathcal{F}_i}$ ist durchschnittstabil,

daher folgt mit dem Eindeutigkeitssatz für Maße (W-Theorie 16.2):

P ist als Wahrscheinlichkeitsmaß durch seine Werte auf dem Erzeugendensystem

$\bigcup\limits_{n=1}^{\infty} \widehat{\mathop{\times}\limits_{i=1}^{n} \mathcal{F}_i} \subset \bigcup\limits_{n=1}^{\infty} \widehat{\bigotimes\limits_{i=1}^{n} \mathcal{F}_i} \equiv \mathcal{A}$ festgelegt.

b.) Für $A_i\in\mathcal{F}_i$, $1\le i\le n$, war nach obiger Konstruktion:

$$\begin{aligned} \bigotimes_{i=1}^{n} Q_i\,[\,A_1\times\ldots\times A_n\,] &\equiv P[\,A_1\times\ldots\times A_n\times\Omega_{n+1}\times\ldots] \\ &= P(\pi_1,\ldots,\pi_n)^{-1}[\,A_1\times\ldots\times A_n\,] \end{aligned}$$

Die Eindeutigkeit folgt wie in a.) IX.)

c.) Es gelten die folgenden Äquivalenzen:

$P\pi_1^{-1} = Q_1;\ Q_{n+1} = P\pi_{n+1}^{-1}\,|\,(\pi_1,\ldots,\pi_n) = \cdot \quad \forall n\in\mathbb{N}$

$\Longleftrightarrow\ P\pi_1^{-1} = Q_1;\ P(\pi_1,\ldots,\pi_n,\pi_{n+1})^{-1} = P(\pi_1,\ldots,\pi_n)^{-1}\otimes Q_{n+1} \quad \forall n\in\mathbb{N}$

nach Definition 2.a) in Abschnitt 1

$\Longleftrightarrow\ P(\pi_1,\ldots,\pi_n)^{-1} = \bigotimes\limits_{i=1}^{n} Q_i \quad \forall n\in\mathbb{N}$, was nach b.) gilt.

d.) Die Aussage folgt aus b.) und Integration mittels Bildmaß.

q.e.d.

Als Korollar dieses Satzes hat man den

Spezialfall: Sei $Q_1 = P_1$ und $Q_{n+1}(\omega_1,\ldots,\omega_n;\,\cdot\,) = P_{n+1}$ unabhängig von $(\omega_1,\ldots,\omega_n)$, $n\in\mathbb{N}$, d.h. P_n sei ein Wahrscheinlichkeitsmaß auf $\mathcal{F}_n$. Dann erhält man aus (1.3) per Induktion

$\bigotimes\limits_{i=1}^{n} Q_i = \mathop{\times}\limits_{i=1}^{n} P_i$ und die Aussage b.) aus (2.3) modifiziert sich zu

$$\bigotimes_{i=1}^{\infty} Q_i\,[\,A_1\times\ldots\times A_n\times\Omega_{n+1}\times\ldots] = (\mathop{\times}_{i=1}^{n} P_i)\,[\,A_1\times\ldots\times A_n\,] \equiv \prod_{i=1}^{n} P_i[\,A_i\,]$$

Bemerkung: P_{n+1} kann gemäß Bemerkung 2 in Abschnitt 1 als Übergangswahrscheinlichkeit aufgefaßt werden.

Dies erfaßt exakt

Korollar 2.4: Seien $(\Omega_n, \mathcal{F}_n, \mathrm{P}_n)$ Wahrscheinlichkeitsräume, $n \in \mathbb{N}$. Dann existiert genau ein Wahrscheinlichkeitsmaß P auf $(\bigtimes_{i=1}^{\infty} \Omega_i, \bigotimes_{i=1}^{\infty} \mathcal{F}_i)$ mit der Eigenschaft

$$\mathrm{P}[A_1 \times \ldots \times A_n \times \Omega_{n+1} \times \ldots] = \prod_{i=1}^{n} \mathrm{P}_i[A_i] \quad \text{für } A_i \in \mathcal{F}_i,\ 1 \leq i \leq n,\ n \in \mathbb{N}.$$

$\mathrm{P} =: \bigtimes_{i=1}^{\infty} \mathrm{P}_i$ heißt (*unendliches*) *Produktmaß* der Wahrscheinlichkeitsmaße P_n, $n \in \mathbb{N}$.

Bemerkung: Mit Hilfe dieses Korollars ist eine Berechnung der Wahrscheinlichkeit einer Menge $\bigtimes_{i=1}^{\infty} A_i$, $A_i \in \mathcal{F}_i$, möglich, da $A_1 \times \ldots \times A_n \times \Omega_{n+1} \times \ldots \downarrow \bigtimes_{i=1}^{\infty} A_i$, $n \to \infty$.

Damit ist es möglich, eine Lösung für das Einführungsproblem zu finden:

Konstruktion eines Wahrscheinlichkeitsmodells unabhängiger Zufallsvariablen

Gegeben: Wahrscheinlichkeitsräume $(\Omega_n, \mathcal{F}_n, \mathrm{P}_n)_{n \in \mathbb{N}}$

Gesucht: Ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ sowie Zufallsvariablen X_n auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in Meßräumen$(\Omega_n, \mathcal{F})_{n \in \mathbb{N}}$ derart, daß gelten: 1.) $X_1, X_2, \ldots$ sind unabhängig
2.) $X_n \sim \mathrm{P}_n$, d.h. $\mathrm{P}{X_n}^{-1} = \mathrm{P}_n \quad \forall\, n \in \mathbb{N}$

Lösung: (analog zum endlichen Fall)

Setze $\Omega := \bigtimes_{i=1}^{\infty} \Omega_i$, $\mathcal{F} := \bigotimes_{i=1}^{\infty} \mathcal{F}_i$, $X_n := \pi_n : \Omega \to \Omega_n$ Projektion, $n \in \mathbb{N}$, sowie $\mathrm{P} := \bigtimes_{i=1}^{\infty} \mathrm{P}_i$.

Dann gilt für $A_i \in \mathcal{F}_i$, $1 \leq i \leq n$:

$$\begin{aligned} \mathrm{P}[X_1 \in A_1, \ldots, X_n \in A_n] &= \mathrm{P}[A_1 \times \ldots \times A_n \times \Omega_{n+1} \times \ldots] \\ &= \prod_{i=1}^{n} \mathrm{P}_i[A_i] \quad \text{nach (2.4)} \end{aligned}$$

Wählt man speziell $A_j := \Omega_j$ für $j \neq i$, $1 \leq j, i \leq n$, so erhält man $\mathrm{P}X_i^{-1}[A_i] = \mathrm{P}_i[A_i]$.
Daraus folgen die Punkte 1 und 2.

Man besitzt nun das nötige Handwerkszeug, um als wichtige Beispiele von stochastischen Prozessen Markoff-Prozesse einzuführen (diese werden in den späteren Abschnitten weiter vertieft).

Definition 1: Ein stochastischer Prozeß $X = (X_n)_{n \in \mathbb{N}_0}$ mit Zustandsraum $(S, \mathcal{S})$ heißt ein (*zeit-*) *homogener Markoff-Prozeß* (*mit Übergangswahrscheinlichkeit Q*), wenn eine Übergangswahrscheinlichkeit Q von S nach S existiert derart, daß die Abbildung $\widehat{Q}_n : S^{n+1} \times \mathcal{S} \to \mathbb{R}$ mit $\widehat{Q}_n\big((s_0, \ldots, s_n); B\big) := Q(s_n; B)$ erfüllt:

(ME) $\widehat{Q}_n$ ist unter P eine bedingte Verteilung von X_{n+1} bzgl. $(X_0, \ldots, X_n)$, $n \in \mathbb{N}_0$.

Die Verteilung von X_0 wird als *Startverteilung* von $(X_n)_{n \in \mathbb{N}_0}$ bezeichnet. Ist S abzählbar, so heißt X *Markoff-Kette* mit *Übergangsmatrix* $\mathbb{P} = (p_{ij})_{i,j \in S}$, wo $p_{ij} := Q(i; \{j\})$.

Bemerkungen: 1.) (ME) heißt *Markoff-Eigenschaft*: Bei Kenntnis des gegenwärtigen Zustands eines stochastischen Systems $\big($ beschrieben durch die $(X_n)_{n \in \mathbb{N}_0}$ $\big)$ hängt dessen zukünftige Entwicklung nur von der Gegenwart ab, nicht von der Vergangenheit.

2.) $\widehat{Q}_n$ in obiger Definition ist eine Übergangswahrscheinlichkeit von S^n nach S, da
$\widehat{Q}_n\big((s_0,\dots,s_n);\,\cdot\,\big) = Q(s_n;\,\cdot\,) \quad \forall s_i \in S,\ 0 \le i \le n$, sowie
$\widehat{Q}_n(\,\cdot\,;B) = Q\big(\pi_n(\,\cdot\,);B\big) \quad \forall\, B \in \mathcal{S}$, wo π_n die Projektion auf die n-te Komponente.

3.) Aus (ME) folgt: Q ist eine bedingte Verteilung von X_{n+1} bzgl. X_n, $n \in \mathbb{N}_0$
Denn für $A, B \in \mathcal{S}$ hat man die Gleichungskette

$$\begin{aligned}
\mathrm{P}[\,X_n \in A, X_{n+1} \in B\,] &= \mathrm{P}[\,(X_0,\dots,X_n) \in S^n \times A, X_{n+1} \in B\,] \\
&\overset{(\mathrm{ME})}{=} \int_{S^n \times A} \mathrm{P}(X_0,\dots,X_n)^{-1}\big(d(x_0,\dots,x_n)\big)\, \widehat{Q}_n\big((x_0,\dots,x_n);B\big) \\
&= \int_{\{(X_0,\dots,X_n) \in S^n \times A\}} \mathrm{P}(d\omega)\, \widehat{Q}_n\big((X_0,\dots,X_n)(\omega);B\big) \\
&= \int_{\{X_n \in A\}} \mathrm{P}(d\omega)\, Q\big(X_n(\omega);B\big) = \int_A \mathrm{P}X_n^{-1}(dx_n)\, Q(x_n;B)
\end{aligned}$$

q.e.d.

4.) Zeitliche Homogenität bedeutet, daß das System $(X_n)_{n\in\mathbb{N}_0}$ zeitlich gleichbleibenden Umweltbedingungen unterliegt. Dies wird hier durch die Unabhängigkeit von Q von n ausgedrückt.

Als direkte Folgerung aus (2.3) ergibt sich:

Korollar 2.5: Sei $(S,\mathcal{S})$ ein meßbarer Raum, ν ein Wahrscheinlichkeitsmaß auf $\mathcal{S}$ und Q eine Übergangswahrscheinlichkeit von S nach S. Seien ferner $(\Omega,\mathcal{F}) := (\mathop{\times}\limits_{i=0}^{\infty} S, \bigotimes\limits_{i=0}^{\infty} \mathcal{S})$ und $\pi_n : \Omega \to S$ die Projektionen, $n \in \mathbb{N}_0$. Dann existiert genau ein Wahrscheinlichkeitsmaß P_ν auf $(\Omega,\mathcal{F})$ mit der Eigenschaft, daß $(\pi_n)_{n\in\mathbb{N}_0}$ unter P_ν ein homogener Markoff-Prozeß mit Übergangswahrscheinlichkeit Q und Startverteilung ν ist.

Beweis: Setze in (2.3) $Q_1 := \nu$ und $Q_{n+1}(x_1,\dots,x_n;B) := Q(x_n;B)$. Dann liefert c.) in (2.3) zusammen mit einer Indexverschiebung die Behauptung.

q.e.d.

Damit hat man bereits die Existenz eines kanonischen Markoff-Prozesses auf einem speziellen Raum gesichert, welcher in späteren Abschnitten noch eine Rolle spielen wird.
Von besonderer Gestalt ist der folgende Existenzsatz:

Satz 2.6: a.) Gegeben seien Meßräume $(S,\mathcal{S})$ und $(E,\mathcal{E})$, eine $\mathcal{S}\otimes\mathcal{E}-\mathcal{S}$-meßbare Funktion $f : S \times E \to S$ sowie eine Folge unabhängiger, identisch verteilter Zufallsvariablen $(Z_n)_{n\in\mathbb{N}}$ auf einem Wahrscheinlichkeitsraum $(\Omega,\mathcal{F},\mathrm{P})$ mit Werten in E. Durch die Definition $X_0 := x_0 \in S$, $X_{n+1} := f(X_n, Z_{n+1})$, $n \in \mathbb{N}_0$, wird $(X_n)_{n\in\mathbb{N}_0}$ zu einem homogenen Markoff-Prozeß mit Übergangswahrscheinlichkeit $Q(x,B) := \mathrm{P}[\,f(x,Z_1) \in B\,]$, $B \in \mathcal{S}$, und Zustandsraum S.

b.) Sei S borelsche Teilmenge eines polnischen Raumes und $\mathcal{S}$ die Borel-σ-Algebra auf S. Ist weiter $x_0 \in S$ und Q eine Übergangswahrscheinlichkeit von S nach S, so existieren $(\Omega,\mathcal{F},\mathrm{P})$, $(E,\mathcal{E})$, $(Z_n)_{n\in\mathbb{N}}$ und f gemäß a.) derart, daß $(X_n)_{n\in\mathbb{N}_0}$ definiert wie in a.) auch ein homogener Markoff-Prozeß mit Übergangswahrscheinlichkeit Q und Zustandsraum S ist.

Beweis: a.) Man zeigt: Q ist eine Übergangswahrscheinlichkeit von S nach S und $\widehat{Q}_n\big((x_0,\ldots,x_n);B\big) := Q(x_n,B)$ ein Element aus $\mathbb{P}X_{n+1}^{-1} \mid (X_0,\ldots,X_n) = \cdot$

b.) Man setzt $(E,\mathcal{E}) := \big((0,1),(0,1)\cap\mathcal{B}_1\big)$, $\Omega := \bigtimes_{i=1}^{\infty} E$, $\mathcal{F} := \bigotimes_{i=1}^{\infty}\mathcal{E}$ und $\mathrm{P} := \bigtimes_{i=1}^{\infty} \mathrm{U}(0,1)$ sowie $Z_n : \Omega \to E$ als die Projektion auf die n-te Komponente. Dann nutzt man aus, daß für Q ein f mit den gewünschten Eigenschaften mittels Pseudoinverser der Verteilungsfunktion von $Q(x,\cdot\,)$ im Fall $(S,\mathcal{S}) = (\mathbb{R},\mathcal{B})$ gefunden werden kann.

□

Als Beispiele für Markoff-Prozesse kann man angeben:

Beispiele: 1.) Lagerhaltung unter einer (s^*,S^*)-Auffüllstrategie, $s^* \le S^*$.
Man hat ein Lager mit der Kapazität S^* und füllt dieses, sobald der Bestand unter eine kritische Grenze s^* sinkt, wieder ganz auf. Dabei werde die Nachfrage nach Gütern durch eine Folge von unabhängigen, identisch verteilten Zufallsvariablen $(Z_n)_{n\in\mathbb{N}}$ beschrieben, d.h. Z_n ist die Nachfrage nach Gütern in der n-ten Periode. Somit ergibt sich für den Lagerbestand X_n nach der n-ten Bestellung die Formel

$$X_0 := S^* \quad (\text{ z.B.})$$

$$X_{n+1} := \begin{cases} X_n - Z_{n+1} & \text{falls } X_n - Z_{n+1} \ge s^* \\ S^* & \text{falls } X_n - Z_{n+1} < s^* \end{cases}$$

$$=: f(X_n, Z_{n+1})$$

Nach (2.6) a.) ist damit $(X_n)_{n\in\mathbb{N}_0}$ ein homogener Markoff-Prozeß.

2.) d-dimensionale Irrfahrt
Sei $(Z_n)_{n\in\mathbb{N}}$ eine Folge von d-dimensionalen, unabhängigen, identisch verteilten Zufallsvektoren und $X_n := x_0 + Z_1 + \ldots + Z_n$, d.h. $X_{n+1} = X_n + Z_{n+1} =: f(X_n, Z_{n+1})$.
Für $(X_n)_{n\in\mathbb{N}_0}$ gilt gleiches wie in 1.). Dieses Beispiel wird später weiter vertieft.

Soviel zunächst zu Markoff-Prozessen. Eine wichtige Verallgemeinerung von (2.4) ist

Satz 2.7: (Kolmogoroff, zeitdiskrete Version)

Gegeben seien Meßräume $(\Omega_n,\mathcal{F}_n)_{n\in\mathbb{N}}$, wo Ω_n borelsche Teilmenge eines polnischen Raumes und $\mathcal{F}_n$ die σ-Algebra seiner borelschen Mengen ist. Ferner sei $(\mathrm{P}_n)_{n\in\mathbb{N}}$ eine Familie von Wahrscheinlichkeitsmaßen, wo P_n auf $\bigotimes_{i=1}^{n}\mathcal{F}_i$ definiert. $(\mathrm{P}_n)_{n\in\mathbb{N}}$ sei *konsistent* in dem Sinne, daß für $B_n \in \bigotimes_{i=1}^{n}\mathcal{F}_i$, $n\in\mathbb{N}$, die Gleichung $\mathrm{P}_{n+1}[\,B_n\times\Omega_{n+1}\,] = \mathrm{P}_n[\,B_n\,]$ stattfindet.

Dann existiert genau ein Wahrscheinlichkeitsmaß P auf $\bigotimes_{i=1}^{\infty}\mathcal{F}_i$ mit der Eigenschaft

$\mathrm{P}[\,B_n\times\Omega_{n+1}\times\ldots] = \mathrm{P}_n[\,B_n\,]$ für $B_n \in \bigotimes_{i=1}^{n}\mathcal{F}_i$, $n\in\mathbb{N}$.

Bemerkung: Setzt man P_n als das Produktmaß der P_i aus (2.4), $1 \le i \le n$, so erhält man die Aussage von (2.4). Die Konsistenz ist auch in diesem Fall gegeben.

Beweis: Setze $\pi_n : \bigtimes_{i=1}^{m} \Omega_i \to \Omega_n$ als Projektion, $n \le m \le \infty$, so ist $\mathrm{P}_{n+1}(\pi_1, \ldots, \pi_n)^{-1} = \mathrm{P}_n$ nach Voraussetzung. Gemäß (1.5) besitzt π_{n+1} unter P_{n+1} eine bedingte Verteilung Q_{n+1} bzgl. $(\pi_1, \ldots, \pi_n)$, deshalb $\mathrm{P}_{n+1} \equiv \mathrm{P}_{n+1}(\pi_1, \ldots, \pi_n, \pi_{n+1})^{-1} = \mathrm{P}_{n+1}(\pi_1, \ldots, \pi_n)^{-1} \otimes Q_{n+1}$
$\equiv \mathrm{P}_n \otimes Q_{n+1} = \mathrm{P}_1 \otimes Q_2 \otimes \ldots \otimes Q_{n+1}$ per Induktion.

Setze gemäß (2.3) $\mathrm{P} := \mathrm{P}_1 \otimes Q_2 \otimes Q_3 \otimes \ldots$, dann gilt wieder mit (2.3):

$\mathrm{P}(\pi_1, \ldots, \pi_n)^{-1} = \mathrm{P}_1 \otimes Q_2 \otimes \ldots \otimes Q_n \equiv \mathrm{P}_n$, d.h. P besitzt die gewünschten Eigenschaften.

Die Eindeutigkeit von P folgt wieder mit dem Eindeutigkeitssatz, da $\bigcup_{n=1}^{\infty} \sigma\big((\pi_1, \ldots, \pi_n)\big)$ ein durchschnittstabiles Erzeugendensystem von $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$ ist.

q.e.d.

Definition 2: $\mathrm{P}_n \equiv \mathrm{P}(\pi_1, \ldots, \pi_n)^{-1}$ heißt *n-dimensionale Randverteilung von P*.

In Abschnitt 3 werden nun einige wichtige und interessante Resultate über unabhängige Zufallsvariablen bzw. unabhängige Ereignisse zusammengestellt:

Abschnitt 3: Null-Eins-Gesetze

Hat man reellwertige unabhänigige Zufallsvariablen $(X_i)_{i \in \mathbb{N}}$, so interessiert z.B. die Frage, mit welcher Wahrscheinlichkeit $\frac{1}{n} \sum_{i=1}^{n} X_i$ für $n \to \infty$ konvergiert. A priori könnten hier alle reellen Zahlen zwischen 0 und 1 als Wahrscheinlichkeiten auftreten. Es wird sich jedoch herausstellen, daß als Konvergenzwahrscheinlichkeiten nur die Zahlen 0 und 1 in Frage kommen. Dies ist ein Beispiel eines sogenannten 0-1-Gesetzes, genauso wie:

Ist A ein Ereignis, welches von allen Ereignissen/sich selbst unabhängig ist, so gilt: $\mathrm{P}[A] \in \{0, 1\}$.
Denn es ist $\mathrm{P}[A] = \mathrm{P}[A \cap A] = \mathrm{P}[A] \cdot \mathrm{P}[A]$.

Weitere Kandidaten für solche Ereignisse und Zufallsvariablen werden in der nächsten Definition vorgestellt:

Definition 1: a.) Sei $X = (X_n)_{n \in \mathbb{N}}$ eine Folge von Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in Meßräumen $(\Omega_n, \mathcal{F}_n)_{n \in \mathbb{N}}$. Dann wird die *σ-Algebra der X-terminalen Ereignisse* definiert durch $\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$

b.) Eine Zufallsvariable $Y : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ heißt *X-terminal*, falls sie sogar $\mathcal{T} - \mathcal{F}'$-meßbar ist.

Bemerkung: $\mathcal{T}$ ist als Schnitt von σ-Algebren wieder eine σ-Algebra. Es ist $\mathcal{T} \neq \emptyset$, da z.B. $\emptyset, \Omega \in \mathcal{T}$. $\mathcal{T}$ beinhaltet die Ereignisse, deren Auftreten bzw. Nichtauftreten nicht vom Verhalten endlich vieler X_i abhängt. Mit anderen Worten: $\mathcal{T}$ beschreibt das asymptotische Verhalten der X_i. Eine X-terminale Zufallsvariable tritt bei Grenzwertbetrachtungen auf.

Nun speziell im Hinblick auf obige erste Motivation das

Beispiel: Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge reeller Zufallsvariablen. Dann ist $Y := \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i$ eine X-terminale Zufallsvariable. Aufgrund von $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{k-1} X_i = 0 \quad \forall\, k \in \mathbb{N}$ ist nämlich $Y = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=k}^{n} X_i \quad \forall k \in \mathbb{N}$. Da aber $\frac{1}{n}\sum_{i=k}^{n} X_i \quad \sigma(X_k, X_{k+1}, \ldots) - \mathcal{B}$-meßbar ist für alle $k, n \in \mathbb{N}$, so auch Y. Mithin ist Y X-terminal.

q.e.d.

Nun wird gezeigt, daß die Ereignisse aus Definition 1 den 0-1-Gesetzen genügen sowie X-terminale Zufallsvariablen charakterisiert:

Satz 3.1: (0-1-Gesetz von Kolmogoroff)
Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge von unabhängigen Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$. Dann gelten die folgenden Aussagen:

a.) Für jedes X-terminale Ereignis A gilt $\mathrm{P}[\,A\,] \in \{\,0, 1\,\}$.

b.) Jede erweitert reelle X-terminale Zufallsvariable Y ist P-fast sicher konstant.

Beweis: a.) Für $A \in \mathcal{T}$ muß man nur zeigen, daß A von allen andnere Ereignissen unabhängig ist.

Sei also $A \in \mathcal{T} \subset \sigma(X) \equiv \sigma(X_1, X_2, \ldots) = \sigma(\Re)$, wo $\Re := \bigcup_{n=1}^{\infty} \sigma(X_1, \ldots, X_n)$ (vgl. (2.2))

Setze $\mathcal{D} := \{\, D \in \sigma(X) : \mathrm{P}[\,A \cap D\,] = \mathrm{P}[\,A\,] \cdot \mathrm{P}[\,D\,]\,\}$. $\mathcal{D}$ ist ein Dynkinsystem (leicht).
Für $R_n \in \sigma(X_1, \ldots, X_n)$ sind A und R_n unabhängig, da $A \in \sigma(X_{n+1}, X_{n+2}, \ldots)$ nach Definition von $\mathcal{T}$ und $(X_1, \ldots, X_n)$, $(X_{n+1}, X_{n+2}, \ldots)$ unabhängig nach W-Theorie 24.2.
$\Longrightarrow R_n \in \mathcal{D}$, d.h. $\Re \subset \mathcal{D} \implies \sigma(X) \equiv \sigma(\Re) \subset \mathcal{D}$, da $\Re$ durchschnittsstabil.

b.) Betrachte die Verteilungsfunktion $F(t) := \mathrm{P}[\,Y \leq t\,]$, $t \in \mathbb{R}$.
Da Y $\mathcal{T} - \overline{\mathcal{B}}$-meßbar und somit die Menge $\{\,Y \leq t\,\} \in \mathcal{T}$ ist, gilt nach a.): $F(t) \in \{\,0, 1\,\}$
F ist isoton in t, daher lassen sich folgende drei Fälle unterscheiden:

Fall 1: $\exists\, a \in \mathbb{R}$ mit $F(t) = 0$ für $t < a$ und $F(t) = 1$ für $t \geq a \implies \mathrm{P}[\,Y = a\,] = 1$
Fall 2: $F = \underline{1} \implies \mathrm{P}[\,Y = -\infty\,] = 1$
Fall 3: $F = \underline{0} \implies \mathrm{P}[\,Y = +\infty\,] = 1$

q.e.d.

Dieses wichtige Resultat findet seine Anwendung u.a. beim starken Gesetz der großen Zahlen, d.h. bei der Beantwortung der eingangs dieses Abschnitts gestellten Frage:

Anwendung 3.2: Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger reeller Zufallsvariablen. Analog zu obigem Beispiel kann man zeigen: $\underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i$ ist X-terminal.

Damit folgt: $\{\,\frac{1}{n}\sum_{i=1}^{n} X_i$ konvergiert $\} = \{\,\underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i\,\} \in \mathcal{T}$.

Nach (3.1) konvergiert $\frac{1}{n}\sum_{i=1}^{n} X_i$ fast sicher oder divergiert fast sicher.

Ferner existieren $\overline{c}, \underline{c} \in \overline{\mathbb{R}}$ mit $\mathrm{P}[\,\overline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{c}\,] = \mathrm{P}[\,\underline{\lim_{n\to\infty}} \frac{1}{n}\sum_{i=1}^{n} X_i = \underline{c}\,] = 1$.

Die Frage, wann $\underline{c} = \overline{c}$ gilt, wird vom starken Gesetz der großen Zahlen beantwortet.

Anwendung 3.3: Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger Zufallsvariablen mit Werten in Meßräumen $(\Omega_n, \mathcal{F}_n)_{n\in\mathbb{N}}$. Dann gilt für $B_n \in \mathcal{F}_n$:

$$\overline{\lim_{n\to\infty}}\{ X_n \in B_n \} = \{ X_n \in B_n \text{ für unendlich viele } n \}$$

$$\equiv \bigcap_{n=1}^{\infty} \bigcup_{i=n}^{\infty} \{ X_i \in B_i \} \overset{(*)}{=} \bigcap_{n=k}^{\infty} \bigcup_{i=n}^{\infty} \{ X_i \in B_i \} \quad \forall\, k \in \mathbb{N}$$

Dabei besteht $(*)$ wegen der Antitonie von $\bigcup_{i=n}^{\infty} \{ X_i \in B_i \}$.

Da aber $\bigcup_{i=n}^{\infty} \{ X_i \in B_i \} \in \sigma(X_k, X_{k+1}, \ldots) \quad \forall\, k \in \mathbb{N}$ mit $n \geq k$, folgt:

$\overline{\lim_{n\to\infty}}\{ X_n \in B_n \} \in \mathcal{T}$, d.h. $\mathrm{P}[\,\overline{\lim_{n\to\infty}}\{ X_n \in B_n \}\,] \in \{ 0, 1 \}$.

Die Anwendung (3.3) wird im nächsten Satz näher charakterisiert:

Satz 3.4: (Lemma von Borel-Cantelli)

Sei $X = (X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in $(\Omega_n, \mathcal{F}_n)_{n\in\mathbb{N}}$. Dann gilt für $B_n \in \mathcal{F}_n$:

a.) $\sum_{n=1}^{\infty} \mathrm{P}[\,X_n \in B_n\,] < \infty \Longrightarrow \mathrm{P}[\,\overline{\lim_{n\to\infty}}\{ X_n \in B_n \}\,] = 0$

b.) Sind die X_n ferner unabhängig, so gilt sogar

$$\sum_{n=1}^{\infty} \mathrm{P}[\,X_n \in B_n\,] = \infty \Longleftrightarrow \mathrm{P}[\,\overline{\lim_{n\to\infty}}\{ X_n \in B_n \}\,] = 1$$

Beweis: a.) Es gilt $\bigcup_{m=n}^{\infty} \{ X_m \in B_m \} \downarrow \overline{\lim_{l\to\infty}}\{ X_l \in B_l \}, \ (n \to \infty)$ (vgl. (3.3))

$$\Longrightarrow \mathrm{P}[\,\overline{\lim_{n\to\infty}}\{ X_n \in B_n \}\,] = \lim_{n\to\infty} \mathrm{P}[\,\bigcup_{m=n}^{\infty} \{ X_m \in B_m \}\,]$$

$$\leq \lim_{n\to\infty} \sum_{m=n}^{\infty} \mathrm{P}[\,X_m \in B_m\,] = 0$$

b.) "$\Leftarrow$" ist die Negation der Aussage a.)

"$\Rightarrow$": Für $n, N \in \mathbb{N}$ hat man die Kette

$$\mathrm{P}[\,\bigcup_{m=n}^{N} \{ X_m \in B_m \}\,] = 1 - \mathrm{P}[\,\bigcap_{m=n}^{N} \{ X_m \in B_m^C \}\,]$$

$$= 1 - \prod_{m=n}^{N} \mathrm{P}[\,X_m \in B_m^C\,] \quad \text{da die } X_i \text{ unabhängig}$$

$$= 1 - \prod_{m=n}^{N} (1 - \mathrm{P}[\,X_m \in B_m\,])$$

$$\geq 1 - \prod_{m=n}^{N} \mathrm{e}^{-\mathrm{P}[\,X_m \in B_m\,]} \quad \text{da } 1 - x \leq \mathrm{e}^{-x} \text{ auf } [0,1]$$

$$= 1 - \mathrm{e}^{-\sum_{m=n}^{N} \mathrm{P}[\,X_m \in B_m\,]} \longrightarrow 1, \ (N \to \infty)$$

Da aber $\mathrm{P}[\,\bigcup_{m=n}^{N} \{ X_m \in B_m \}\,] \uparrow \mathrm{P}[\,\bigcup_{m=n}^{\infty} \{ X_m \in B_m \}\,], \ (N \to \infty)$, folgt die Behauptung.

q.e.d.

Am Schluß dieses Abschnitts noch eine nützliche

Anwendung: Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von identisch verteilten, unabhängigen Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in $(\Omega', \mathcal{F}')$. Dann gilt für $B \in \mathcal{F}'$:

$$\mathrm{P}[\overline{\lim_{n\to\infty}} \{ X_n \in B \}] = \begin{cases} 0 & \text{falls } \mathrm{P}[X_1 \in B] = 0 \\ 1 & \text{falls } \mathrm{P}[X_1 \in B] > 0 \end{cases}$$

<u>Beweis:</u> $\mathrm{P}[X_1 \in B] = \mathrm{P}[X_n \in B] \quad \forall n \in \mathbb{N}$, da die X_n identisch verteilt. Somit gilt:

$$\begin{aligned} \mathrm{P}[X_1 \in B] = 0 &\Longrightarrow \sum_{n=1}^{\infty} \mathrm{P}[X_n \in B] = 0 \\ &\Longrightarrow \mathrm{P}[\overline{\lim_{n\to\infty}} \{ X_n \in B \}] = 0 \text{ nach (3.4) a.)} \\ \mathrm{P}[X_1 \in B] > 0 &\Longrightarrow \sum_{n=1}^{\infty} \mathrm{P}[X_n \in B] = \infty \\ &\Longrightarrow \mathrm{P}[\overline{\lim_{n\to\infty}} \{ X_n \in B \}] = 1 \text{ nach (3.4) b.)} \end{aligned}$$

q.e.d.

Nachdem in Abschnitt 2 schon der in der Theorie der stochastischen Prozesse wichtige Begriff der Markoff-Eigenschaft eingeführt wurde, wird nun mit Hilfe des bedingten Erwartungswertes der ebenfalls fundamentale Martingalbegriff definiert und charakterisiert.

Abschnitt 4: Zeitdiskrete Martingale

Die Martingaltheorie hat, wie viele andere Elemente der Stochastik, ihren Ursprung in der Behandlung/Auswertung von Glücksspielchancen. (Martingal = Strategie beim Roulette, wobei man den beim vorangegangenen Spiel verlorenen Einsatz nun verdoppelt einsetzt)

Ferner wird sie u.a. auch zur Kontrolle von stochastischen Prozessen und zur Herleitung von Aussagen über die fast sichere Konvergenz von Folgen von Zufallsvariablen herangezogen.
Man sucht z.B. in der Stochastik ein Analogon zu der aus der Analysis her bekannnten Aussage:

"Jede isotone, nach oben beschränkte Folge hat einen reellen Grenzwert."

Dieses Resultat wird später im Martingalkonvergenzsatz hergeleitet, ein wichtiges Ziel dieses Abschnitts.

Sei dafür im folgenden (falls nichts anderes vorausgestzt) $X = (X_n)_{n\in I}$ eine Familie reeller Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$. Dabei drückt die Indexmenge $I \subset \overline{\mathbb{Z}} = \mathbb{Z} \cup \{+\infty, -\infty\}$ die zeitliche Komponente von X aus ($n \in I$ Zeitparameter). Es wird daher für I die Bedingung $[n \in I,\ n < \sup I \Rightarrow n+1 \in I]$ vorausgesetzt, d.h. hat man in einem Zeitpunkt $n \in I$ eine Beobachtung gemacht, so möchte man diese mit dem Wert im Zeitpunkt $n+1$ vergleichen können.

Um in einem Zeitpunkt $n \in I$ Ereignisse der Vorgeschichte einschließlich der Gegenwart interpretieren/auswerten zu können, definiert man:

Definition 1: Sei $\mathbb{F} = (\mathcal{F}_n)_{n\in I}$ eine Familie von Unter-σ-Algebren von $\mathcal{F}$, I wie oben. Dann heißt $\mathbb{F}$ <u>*Filterung*</u> (<u>*in $\mathcal{F}$*</u>), falls $\mathcal{F}_m \subset \mathcal{F}_n \subset \mathcal{F} \quad \forall m \le n,\ m, n \in I$. Die Filterung $\mathbb{F}^X := (\mathcal{F}_n^X)_{n\in I}$ mit $\mathcal{F}_n^X := \sigma(X_m, m \le n, m \in I)$ heißt die zu $(X_n)_{n\in I}$ gehörige <u>*kanonische Filterung*</u>. $(X_n)_{n\in I}$ ist <u>*adaptiert an eine Filterung $\mathbb{F}$*</u>, wenn für alle $n \in I$ X_n $\mathcal{F}_n - \mathcal{B}$-meßbar ist.

Bemerkungen: 1.) Die Unter-σ-Algebra $\mathcal{F}_n$ ist als System von Ereignissen der Vorgeschichte bis zum Zeitpunkt n (einschließlich) konzipiert, d.h. die Zufallsvariable 1_A mit $A \in \mathcal{F}_n$ ist bis zum Zeitpunkt n beobachtbar.

2.) Der Begriff der Filterung wird der Tatsache gerecht, daß bei fortschreitender Zeit die Informationsmenge anwächst.

3.) Die Adaptiertheit eines Prozesses $(X_n)_{n\in I}$ bedeutet, daß im Zeitpunkt $n \in I$ die Beobachtungen von X_n Ereignisse aus $\mathcal{F}_n$ sind. Dies verdeutlicht die Aussage

$$(X_n)_{n\in I} \text{ ist an } \mathbb{F} \text{ adaptiert } \iff \mathbb{F}^{\mathrm{X}} \subset \mathbb{F}$$

<u>Beweis:</u> $(X_n)_{n\in I}$ ist adaptiert an $\mathbb{F}$

$$\iff X_n \text{ ist } \mathcal{F}_n\text{-meßbar} \quad \forall\, n \in I$$

$$\iff \sigma(X_m) \subset \mathcal{F}_m \subset \mathcal{F}_n \quad \forall\, m \leq n,\ m \in I \quad \forall\, n \in I$$

$$\iff \mathcal{F}_n^X \equiv \sigma\Big(\bigcup_{m\leq n,\ m\in I} \sigma(X_m)\Big) \subset \mathcal{F}_n \quad \forall\, n \in I.$$

q.e.d.

Zur Adaptiertheit noch ein

Beispiel: $X = (X_n)_{n\in I}$ mit $X_n := f_n(Z_m, m \leq n, m \in I)$ ist ein an $\mathbb{F}^Z$ adaptierter Prozeß, wobei $f_m : (\bigtimes_{l\leq m, l\in I} \Omega_l, \bigotimes_{l\leq m, l\in I} \mathcal{F}_l) \to (\mathbb{R}, \mathcal{B})$, $m \in I$, eine Familie von meßbaren Abbildungen sowie $Z_m : (\Omega, \mathcal{F}) \to (\Omega_m, \mathcal{F}_m)$, $m \in I$, ein überlagerter Prozeß ist.

<u>Beweis:</u> Es ist zu zeigen: $\mathbb{F}^X \subset \mathbb{F}^Z$

Dafür hat man $\mathcal{F}_n^X \equiv \sigma\Big(\bigcup_{m\leq n, m\in I} \sigma\big(f_m(Z_l, l \leq m, l \in I)\big)\Big)$ und ferner

$$\begin{aligned}
\sigma\big(f_m(Z_l, l \leq m, l \in I)\big) &\equiv (Z_l, l \leq m, l \in I)^{-1} f_m^{-1}(\mathcal{B}) \\
&\subset (Z_l, l \leq m, l \in I)^{-1}\Big(\bigotimes_{l\leq m, l\in I} \mathcal{F}_l\Big). \\
&\subset \mathcal{F}_m^Z \subset \mathcal{F}_n^Z \quad \forall m \leq n,\ m \in I
\end{aligned}$$

q.e.d.

Nun zu der schon angekündigten Definition des Martingalbegriffs:

Definition 2: Seien $(\Omega, \mathcal{F}, \mathrm{P})$ und I wie oben, $\mathbb{F} = (\mathcal{F}_{\mathrm{n}})_{\mathrm{n}\in\mathrm{I}}$ eine Filterung in $\mathcal{F}$ sowie $X = (X_n)_{n\in I}$ eine Familie von reellen Zufallsvariablen, welche an $\mathbb{F}$ adaptiert ist. Dann heißt X <u>*Submartingal*</u> (<u>*bzgl.* $\mathbb{F}$</u>), wenn für alle $n \in I$ gilt: X_n ist integrierbar und

$$(*) \quad X_m \leq \mathrm{E}[\,X_n \mid \mathcal{F}_m\,] \quad \text{fast sicher} \quad \forall\, m \leq n, m \in I$$

X heißt <u>*Supermartingal*</u>, falls $-X$ Submartingal ist $\big($ d.h. in $(*)$ steht "$\geq$" statt "$\leq$" $\big)$

X wird als <u>*Martingal*</u> bezeichnet, falls X sowohl Sub- als auch Supermartingal ist $\big($ d.h. in $(*)$ steht "$=$" statt "$\leq$" $\big)$

Bemerkungen: 0.) $(*)$ in Definition 2 ist, daX_m $\mathcal{F}_m - \mathcal{B}$-meßbar, äquivalent zu
$\mathrm{E}[\,X_n - X_m \mid \mathcal{F}_m\,] \geq 0$ fast sicher $\forall\, n, m \in I$ mit $n \geq m$.
Dies bedeutet: $\int\limits_A X_m \, d\mathrm{P} \leq \int\limits_A X_n \, d\mathrm{P} \quad \forall\, A \in \mathcal{F}_m$, wo m, n wie oben.

1.) Die Martingaltheorie eignet sich vorzüglich für Prognosen. Denn nach 0.) gilt insbesondere (da $\Omega \in \mathcal{F}_m$): $\mathrm{E}[\,X_m\,] \leq \mathrm{E}[\,X_n\,] \quad \forall\, m \leq n\,,\,,\ m, n \in I$
Beschreibt X ein Glücksspiel, so ist es im Submartingalfall günstig weiterzuspielen, da in der Zukunft n die zu erwartende Ausschüttung höher ist als die Ausschüttung in der Gegenwart m.

2.) Für $I \subset \mathbb{Z}$ reicht zum Nachweis von $(*)$ eine Untersuchung der Zeitpunkte n und $n+1$ aus (wo $n < \sup I$).
Gilt nämlich für alle $n < \sup I$ die Beziehung $\quad (+)\ \mathrm{E}[\,X_{n+1} \mid \mathcal{F}_n\,] \geq X_n$, so folgt:

$$\begin{aligned} X_{m-1} &\leq \mathrm{E}[\,X_m \mid \mathcal{F}_{m-1}\,] \quad \text{nach } (+) \\ &\leq \mathrm{E}\big[\,\mathrm{E}[\,X_{m+1} \mid \mathcal{F}_m\,] \mid \mathcal{F}_{m-1}\big] \quad \text{nach } (+) \text{ und (1.10) b.)} \\ &= \mathrm{E}[\,X_{m+1} \mid \mathcal{F}_{m-1}\,] \quad \text{nach (1.10) e.)} \\ &= \ldots\ldots \end{aligned}$$

d.h. man erhält nach sukzessiver Berechnung $(*)$ aus Definition 2.

Nun einige

Beispiele: a.) Eindimensionale Irrfahrt

Sei $(Z_i)_{i\in\mathbb{N}}$ eine Folge unabhängiger, integrierbarer Zufallsvariablen, $x_0 \in \mathbb{R}$ der Startpunkt und $X_n := x_0 + \sum\limits_{i=1}^{n} Z_i$ für $n \in \mathbb{N}_0$. Dann gilt mit der Filterung $\mathbb{F} = (\mathcal{F}_\mathrm{n})_{\mathrm{n}\in\mathbb{N}_\circ}$ wo $\mathcal{F}_n := \sigma(Z_1, \ldots, Z_n)\,,\ n \in \mathbb{N}\,,\ \mathcal{F}_\circ := \{\,\emptyset, \Omega\,\}$:

$$\mathrm{E}[\,Z_i\,] \left\{ \begin{array}{c} \geq 0 \\ = 0 \\ \leq 0 \end{array} \right\} \quad \forall\, i \in \mathbb{N} \Longrightarrow (X_n)_{n\in\mathbb{N}_\circ} \text{ ist ein } \left\{ \begin{array}{c} \text{Sub-} \\ \\ \text{Super-} \end{array} \right\} \text{Martingal.}$$

$$\begin{aligned} \text{Denn für } n \in \mathbb{N} \text{ hat man } \mathrm{E}[\,X_{n+1} - X_n \mid \mathcal{F}_n\,] &\equiv \mathrm{E}[\,Z_{n+1} \mid Z_1, \ldots, Z_n\,] \\ &= \mathrm{E}[\,Z_{n+1}\,] \quad \text{nach (1.7)} \end{aligned}$$

b.) Sei $\mathbb{F} = (\mathcal{F}_\mathrm{n})_{\mathrm{n}\in\mathbb{N}}$ eine Filterung in $\mathcal{F}$ und Z eine integrierbare Zufallsvariable. Dann ist $X = (X_n)_{n\in\overline{\mathbb{N}}}$ mit $X_n := \mathrm{E}[\,Z \mid \mathcal{F}_n\,]$ und $\mathcal{F}_\infty := \sigma(\bigcup\limits_{n=1}^{\infty} \mathcal{F}_n)$ ein Martingal bzgl. der "abgeschlossenen" Filterung $(\mathcal{F}_n)_{n\in\overline{\mathbb{N}}}$.

<u>Beweis:</u> X ist an $(\mathcal{F}_n)_{n\in\overline{\mathbb{N}}}$ adaptiert und integrierbar nach Definition der bedingten Erwartung. Ferner gilt für $m, n \in \overline{\mathbb{N}}$ mit $m \leq n$:

$$\begin{aligned} \mathrm{E}[\,X_n \mid \mathcal{F}_m\,] &\equiv \mathrm{E}\big[\mathrm{E}[\,Z \mid \mathcal{F}_n\,] \mid \mathcal{F}_m\big] = \mathrm{E}[\,Z \mid \mathcal{F}_m\,] \quad \text{nach (1.10) a.)} \\ &\equiv X_m \end{aligned}$$

q.e.d.

Eine für weitere Betrachtungen wichtige Eigenschaft von Submartingalen beschreibt das

Lemma 4.1: Es sei $X = (X_n)_{n\in I}$ ein [Sub-] Martingal (bzgl. einer beliebigen Filterung) und $\varphi : \mathbb{R} \to \mathbb{R}$ eine [isotone] konvexe Funktion derart, daß $\varphi(X_n)$ integrierbar ist $\quad \forall\, n \in I$.
Dann ist $\big(\varphi(X_n)\big)_{n\in I}$ ein Submartingal.

Beweis: $\left(\varphi(X_n)\right)_{n\in I}$ ist an $\mathbb{F}^X$ adaptiert, da φ jeweils $\mathcal{B}-\mathcal{B}$-meßbar. Für $m,n\in I$ mit $m\le n$ gilt:

Fall "konvex": $\varphi(X_m)\equiv\varphi(\mathrm{E}[\,X_n\mid\mathcal{F}_m^X\,])\overset{(+)}{\le}\mathrm{E}[\,\varphi(X_n)\mid\mathcal{F}_m^X\,]$

Fall "isoton konvex": $\varphi(X_m)\le\varphi(\mathrm{E}[\,X_n\mid\mathcal{F}_m^X\,])\overset{(+)}{\le}\mathrm{E}[\,\varphi(X_n)\mid\mathcal{F}_m^X\,]$

wo (+) jeweils wegen (1.10) f.) besteht.

q.e.d.

Beispiel: Ein wichtiger Vertreter ist die Funktionenklasse $\varphi_a(x):=a\vee x\equiv\max(a,x)$

Spezialfall: $\varphi_0(x)=x^+$

Nun eine für die Martingaltheorie (im Bezug auf Konvergenz wie am Anfang motiviert) und die Integrationstheorie an sich unerläßliche

Definition 2: Sei $X=(X_n)_{n\in I}$ eine Familie von erweitert-reellen Zufallsvariablen auf $(\Omega,\mathcal{F},\mathrm{P})$, I eine beliebige Indexmenge. Dann heißt X *gleichmäßig integrierbar*

$$:\Longleftrightarrow\quad\sup_{n\in I}\int\limits_{\{\,|X_n|>\alpha\,\}}|\,X_n\,|\;d\mathrm{P}\longrightarrow 0\,,\;(\alpha\to\infty).$$

Bemerkung 4.2: 1.) Es gilt für integrierbares X stets $\int\limits_{\{\,|X|>\alpha\,\}}|\,X\,|\;d\mathrm{P}\longrightarrow 0\,,\;(\alpha\to\infty)$. Dies folgt mit majorisierter Konvergenz, da $|\,X\,|\cdot 1_{\{\,|X|>\alpha\,\}}\downarrow\underline{0}$ P-fast sicher, sofern $\alpha\to\infty$.

2.) Man kann zeigen (Übung):

$(X_n)_{n\in I}$ gleichmäßig integrierbar $\Longleftrightarrow\sup\limits_{n\in I}\mathrm{E}[\,(|\,X_n\,|-\alpha)^+\,]\longrightarrow 0\,,\;(\alpha\to\infty)$

Analog zu 1.) gilt für jedes integrierbare X: $\mathrm{E}[\,(|\,X\,|-\alpha)^+\,]\to 0\,,\;(\alpha\to\infty)$

Es folgt nun eine Erweiterung des Satzes der majorisierten Konvergenz:

Korollar 4.3: Ist $X=(X_n)_{n\in\mathbb{N}}$ gleichmäßig integrierbar und $X_n\longrightarrow X_\infty$, $(n\to\infty)$, fast sicher, so ist X_∞ integrierbar und $\mathrm{E}[\,X_n\,]\longrightarrow\mathrm{E}[\,X_\infty\,]$, $(n\to\infty)$.

Beweis: Sei o.E. $X_n\ge 0$ (ansonsten Zerlegung $X_n=X_n^+-X_n^-$ und getrennte Betrachtungen)
Aus der gleichmäßigen Integrierbarkeit folgt sofort (Übung): $\sup\limits_{n\in\mathbb{N}}\mathrm{E}[\,X_n\,]<\infty$
Damit erhält man mit Hilfe von Fatou (Minorante 0):

$\infty>\sup\limits_{n\in\mathbb{N}}\mathrm{E}[\,X_n\,]\ge\varliminf\limits_{n\to\infty}\mathrm{E}[\,X_n\,]\ge\mathrm{E}[\,X_\infty\,]$, d.h. X_∞ ist integrierbar.

Für die zweite Behauptung wähle man zu $\epsilon>0$ ein $\alpha>0$ derart, daß $\int\limits_{\{\,X_n>\alpha\,\}}X_n\,d\mathrm{P}<\epsilon\quad\forall\,n\in\mathbb{N}$.
Dies ist aufgrund der gleichmäßigen Integrierbarkeit von X möglich.

Mit Hilfe von α "stutzt" man X_∞, um eine integrierbare Majorante zu erhalten:

$$
\begin{aligned}
\underline{\lim}_{n\to\infty} \mathrm{E}[\, X_n\,] &\geq \mathrm{E}[\, X_\infty\,] \geq \mathrm{E}[\min(X_\infty,\alpha)] \equiv \mathrm{E}[\,\alpha\wedge X_\infty\,] \\
&= \mathrm{E}[\,\overline{\lim_{n\to\infty}}\ \alpha\wedge X_n\,] \geq \overline{\lim_{n\to\infty}}\ \mathrm{E}[\,\alpha\wedge X_n\,] \quad \text{mit Fatou} \\
&\geq \overline{\lim_{n\to\infty}} \int\limits_{\{X_n\leq\alpha\}} (X_n\wedge\alpha)\,d\mathrm{P} \qquad \text{da } X_n\wedge\alpha\geq 0 \\
&\equiv \overline{\lim_{n\to\infty}} \int\limits_{\{X_n\leq\alpha\}} X_n\ d\mathrm{P} \equiv \overline{\lim_{n\to\infty}}\,(\mathrm{E}[\,X_n\,] - \int\limits_{\{X_n>\alpha\}} X_n\,d\mathrm{P}) \geq \overline{\lim_{n\to\infty}}\ \mathrm{E}[\,X_n\,]-\epsilon
\end{aligned}
$$

Mit $\epsilon\to 0$ ergibt sich die Konvergenzbehauptung.

q.e.d.

Im vorigen Beweis sah man schon die Vorteile beim "Stutzen" bzw. "Abschneiden" eines Prozesses. Eine weitere Eigenschaft zeigt

Lemma 4.4: Sei $X=(X_n)_{n\in I}$ ein Submartingal bzgl. $\mathbb{F}=(\mathcal{F}_\mathrm{n})_{\mathrm{n\in I}}$, I habe die Eigenschaft $\sup I\in I$. Dann ist für alle $a\in\mathbb{R}$ $(a\vee X_n)_{n\in I}$ ein gleichmäßig integrierbares Submartingal. (Beispiel: Für $a=0$: $a\vee X_n = X_n^+$)

Beweis: Sei $s := \sup I$. Nach (4.1) ist $(a\vee X_n)_{n\in I}$ ein Submartingal. Sei o.E. $X_n\geq 0$, $a=0$ $\big($ ansonsten betrachte man $(a\vee X_n)-a$, was ebenfalls nach (4.1) wieder ein Submartingal ist $\big)$. Damit gilt für $\alpha>0$:

$$
\begin{aligned}
\int\limits_{\{X_n>\alpha^2\}} X_n\,d\mathrm{P} &\leq \int\limits_{\{X_n>\alpha^2\}} X_s\ \ d\mathrm{P} \quad (\text{ Submartingal }) \\
&\leq \int\limits_{\{X_n>\alpha^2\}} (X_s-\alpha)^+ + \alpha\ \ d\mathrm{P} \quad \text{da } x\leq (x-\alpha)^+ + \alpha \\
&\leq \mathrm{E}[\,(X_s-\alpha)^+\,] + \alpha\cdot\mathrm{P}[\,X_n>\alpha^2\,] \\
&\leq \mathrm{E}[\,(X_s-\alpha)^+\,] + \frac{1}{\alpha}\cdot\mathrm{E}[\,X_n\,] \quad \text{nach Markoff} \\
&\leq \mathrm{E}[\,(X_s-\alpha)^+\,] + \frac{1}{\alpha}\cdot\mathrm{E}[\,X_s\,] \quad (\text{ Submartingal }) \\
&\longrightarrow 0,\ \ (\alpha\to\infty) \quad \text{gleichmäßig in } n \text{ nach (4.2) 2.)}
\end{aligned}
$$

q.e.d.

Wie eingangs des Abschnitts schon angedeutet, wird nun die Rolle der Martingaltheorie als Kotrollinstrument von stochastischen Prozessen (z.B. bei Glücksspielen) erläutert:

Definition 3: Ist $X=(X_n)_{n\in\mathbb{N}_0}$ ein Submartingal [Supermartingal] bzgl. $\mathbb{F}=(\mathcal{F}_\mathrm{n})_{\mathrm{n}\in\mathbb{N}_0}$, so wird ein stochastischer Prozeß $U=(U_n)_{n\in\mathbb{N}_0}$ *Kontrolle des Submartingals* [*Supermartingals*] X oder *Spielsystem* genannt, falls $U\geq 0$ und an $\mathbb{F}$ adaptiert ist.

Der Prozeß $X^U=(X_n^U)_{n\in\mathbb{N}_0}$ mit $X_n^U := X_0 + \sum_{i=1}^{n} U_{i-1}(X_i - X_{i-1})$

wird als das *durch U kontrollierte Submartingal* [*Supermartingal*] bezeichnet.

Bemerkung: Der Kontrollprozeß U erlaubt es, im Zeitpunkt $n-1$ zu entscheiden, inwiefern der Zuwachs $(X_n - X_{n-1})$ sich im Ergebnis niederschlagen soll. ($U = \underline{1}$ bedeutet die unveränderte Übernahme des Ausgangsprozesses, $X^{\underline{1}} \equiv X$)
Da in X_n^U $n-1$ als höchster Index von U auftritt, bedeutet die Adaptiertheit von U, daß die Entscheidung über die Größenordnung von U_n nur von den Erfahrungswerten der Vergangenheit im Zeitpunkt n abhängt. Eine andere Möglichkeit für die Verwendung von U ist (wie nachher noch näher erläutert) die Entwicklung von Spielstrategien.

Nun eine Charakterisierung des Kontrollprozesses

Proposition 4.5: (über Spielsysteme)
Jedes kontrollierte Submartingal X^U ist wieder ein Submartingal, falls X^U integrierbar ist. Ist U $[0,1]$-wertig, so ist X^U integrierbar und es gilt:

$$(*) \quad \mathrm{E}[\,X_n^U\,] \leq \mathrm{E}[\,X_n\,] \quad \forall\, n \in \mathbb{N}_0$$

Beweis: Für $n \in \mathbb{N}$ hat man mit (1.10) g.), da $U_n(X_{n+1} - X_n)$ integrierbar und U_n $\mathcal{F}_n$-meßbar ist:

$$\begin{aligned}\mathrm{E}[\,X_{n+1}^U - X_n^U \mid \mathcal{F}_n\,] &= \mathrm{E}[\,U_n(X_{n+1} - X_n) \mid \mathcal{F}_n\,] \\ &= \underbrace{U_n}_{\mathrm{I}}\ \underbrace{\mathrm{E}[\,X_{n+1} - X_n \mid \mathcal{F}_n\,]}_{\mathrm{II}} \geq 0\end{aligned}$$

Dabei ist $\mathrm{I} \geq 0$ nach Voraussetzung und $\mathrm{II} \geq 0$, da X ein Submartingal.
Der Beweis der Ungleichung $(*)$ erfolgt durch

$$\begin{aligned}\mathrm{E}[\,X_n^U\,] &\equiv \mathrm{E}[\,X_0\,] + \sum_{i=1}^{n} \mathrm{E}\big[\,\mathrm{E}[\,U_{i-1}(X_i - X_{i-1}) \mid \mathcal{F}_{i-1}\,]\big] \\ &= \mathrm{E}[\,X_0\,] + \sum_{i=1}^{n} \mathrm{E}\big[\,U_{i-1}\mathrm{E}[\,X_i - X_{i-1} \mid \mathcal{F}_{i-1}\,]\big] \quad \text{analog zu oben} \\ &\leq \mathrm{E}[\,X_0\,] + \sum_{i=1}^{n} \mathrm{E}\big[\,\underbrace{\mathrm{E}[\,X_i - X_{i-1} \mid \mathcal{F}_{i-1}\,]}_{\mathrm{III}}\,\big] \quad \text{da } 0 \leq U_{i-1} \leq 1 \text{ und } \mathrm{III} \geq 0 \\ &= \mathrm{E}[\,X_0\,] + \sum_{i=1}^{n} \big(\mathrm{E}\big[\,\mathrm{E}[\,X_i \mid \mathcal{F}_{i-1}\,]\big] - \mathrm{E}[\,X_{i-1}\,]\big) \\ &\equiv \mathrm{E}[\,X_n\,] \quad \text{wegen } \mathrm{E}\big[\,\mathrm{E}[\,X_i - X_{i-1} \mid \mathcal{F}_{i-1}\,]\big] = \mathrm{E}[\,X_i\,] - \mathrm{E}[\,X_{i-1}\,]\end{aligned}$$

q.e.d.

Eine Anwendung beim Glücksspiel zeigt das

Beispiel: Sei $Z = (Z_n)_{n\in\mathbb{N}}$ eine Folge von identisch verteilten, unabhängigen $\{\,-1, 1\,\}$-wertigen Zufallsvariablen. Mit $x_0 \in \mathbb{R}$ definiere man $X_n := x_0 + \sum_{i=1}^{n} Z_i$, $n \in \mathbb{N}_0$.
Das Ereignis $Z_n = 1$ bedeute Gewinn im n-ten Spiel, $Z_n = -1$ analog Verlust, $p := \mathrm{P}[\,Z_n = 1\,]$ bezeichne die Gewinnwahrscheinlichkeit.
Nach obigem Beispiel definiert die Irrfahrt $X = (X_n)_{n\in\mathbb{N}_0}$ ein Supermartingal genau dann, wenn $p \leq \frac{1}{2}$ ist ($\Longleftrightarrow \mathrm{E}[\,Z_n\,] \leq 0$).
Als Filterung habe man die kanonische Filterung $\mathbb{F}^Z$ mit $\mathcal{F}_0^Z := \{\,\emptyset, \Omega\,\}$. Ist $U = (U_n)_{n\in\mathbb{N}_0}$ Kontrolle, wo U_n den Einsatz beim n-ten Spiel wiedergibt, so ist $X_n^U = X_{n-1}^U + U_{n-1}Z_n$ der Kapitalstand zum Zeitpunkt n.

Gemäß der Bemerkung nach Definition 6 in Abschnitt 1 existiert eine Folge meßbarer Abbildungen $(\delta_n)_{n\in\mathbb{N}_0}$ mit $U_n = \delta_n(Z_1, \ldots, Z_n)$, $n \geq 1$. Dabei ist δ_o konstant, sinnvollerweise $\leq x_0$. Die Integierbarkeit von X_n^U folgt aus der Tatsache, daß X_n^U nur endlich viele reelle Werte annimmt. Nach (4.5) ist X^U ein Supermartingal, falls $p \leq \frac{1}{2}$, insbesondere besteht die Beziehung $\mathrm{E}[\,X_n^U\,] \leq \mathrm{E}[\,X_0^U\,] \equiv X_0$ (d.h. spielen lohnt nicht).
Beim Petersburger Spiel, einem Münzwurfspiel, bei dem man mit dem Einsatz 1 anfängt und seinen Einsatz solange verdoppelt, bis man gewinnt, hat man demnach
$\delta_0 = 1$, $\delta_n(\underbrace{-1, \ldots, -1}_{n\text{-mal}}) = 2^n$, $\delta_n = 0$ sonst.

Nachdem das allgemeine Kontrollkonzept vorgestellt wurde, nun zu einem speziellen Kontrollinstrument, der Stoppzeit. Die zugrundeliegende Idee ist, einen stochastischen Prozeß beim Vorliegen bestimmter Abbruchbedingungen (z.B. Erreichen eiener Gewinngrenze, Ruin oder allgemeiner: erster Eintritt des Prozesses in eine bestimmte Ereignismenge) abbrechen zu lassen (zu stoppen).
Dafür die

Definition 4: Sei eine Filterung $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ gegeben. Dann heißt $\tau : \Omega \to \overline{\mathbb{N}}_0$ *Stoppzeit* (*bzgl.* $\mathbb{F}$), falls $\{\,\tau = n\,\} \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}_0$.

Bemerkungen: 1.) Es gilt: $\{\,\tau = n\,\} \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}_0 \iff \{\,\tau \leq n\,\} \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}_0$

Beweis: "$\Rightarrow$": $\{\,\tau \leq n\,\} = \sum_{i=0}^{n} \{\,\tau = i\,\}$

"$\Leftarrow$": $\{\,\tau = n\,\} = \{\,\tau \leq n\,\} - \{\,\tau \leq n-1\,\}$.

q.e.d.

Die Bedingung $\{\,\tau \leq n\,\} \in \mathcal{F}_n$ besagt, daß die Entscheidung, in τ zu stoppen, keinen Vorgriff auf die Zukunft bedeutet.

2.) Mit $U_n := 1_{\{\,\tau > n\,\}} \geq 0$ ist U Kontrolle im Sinne von Definition 3 (U = Abbruch bei τ), da $\{\,\tau > n\,\} = \{\,\tau \leq n\,\}^C \in \mathcal{F}_n \quad \forall\, n \in \mathbb{N}_0$. Für X^U hat man dann:

$$X_n^U \equiv X_0 + \sum_{i=1}^{n} 1_{\{\,\tau \geq i\,\}}\,(X_i - X_{i-1}) = X_0 + \sum_{i=1}^{n\wedge\tau} (X_i - X_{i-1}) = X_{n\wedge\tau}$$

$X_{n\wedge\tau}$ ist der mittels τ gestoppte Prozeß.

Für einen gestoppten Prozeß kann man als Korollar zu (4.5) angeben:

Korollar 4.6: (Stoppsatz)
Sei $X = (X_n)_{n\in\mathbb{N}_0}$ ein Submartingal bzgl. einer Filterung $\mathbb{F}$ und τ eine Stoppzeit bzgl. gleicher Filterung. Dann gelten:

a.) $(X_{n\wedge\tau})_{n\in\mathbb{N}_0}$ ist ein Submartingal und $\mathrm{E}[\,X_0\,] \leq \mathrm{E}[\,X_{n\wedge\tau}\,] \leq \mathrm{E}[\,X_n\,]$.

b.) Ist ferner $X^+ = (X^+_{n\wedge\tau})_{n\in\mathbb{N}_0}$ gleichmäßig integrierbar mit $\tau < \infty$ fast sicher, so ist $\mathrm{E}[\,X_0\,] \leq \mathrm{E}[\,X_\tau\,] < \infty$.

Beweis: a.) Die zweite Ungleichung ist nach obiger Bemerkung 2.) die Aussage von (4.5) mit $U_n = 1_{\{\,\tau > n\,\}}$. Die erste Ungleichung folgt sofort aus den ersten drei Beziehungen des Beweises von (4.5).

b.) Es gilt: $\underline{\lim}_{n\to\infty} \mathrm{E}[\,X^-_{n\wedge\tau}\,] \geq \mathrm{E}[\,\underline{\lim}_{n\to\infty} X^-_{n\wedge\tau}\,]$ nach Fatou

$= \mathrm{E}[\,1_{\{\,\tau<\infty\,\}} X^-_\tau\,] \equiv \mathrm{E}[\,X^-_\tau\,]$ nach Voraussetzung

Jedoch hat man andererseits, da $X^+_{n\wedge\tau} \longrightarrow X^+_\tau$ fast sicher:

$$\begin{aligned}\lim_{n\to\infty} \mathrm{E}[\,X^+_{n\wedge\tau}\,] &= \lim_{n\to\infty} \mathrm{E}[\,1_{\{\tau<\infty\}} X^+_{n\wedge\tau}\,] = \mathrm{E}[\,1_{\{\tau<\infty\}} X^+_\tau\,] \text{ nach (4.3)}\\ &= \mathrm{E}[\,X^+_\tau\,]\end{aligned}$$

$$\begin{aligned}\Longrightarrow\ \mathrm{E}[\,X_0\,] &\le \varliminf_{n\to\infty} \mathrm{E}[\,X_{n\wedge\tau}\,] \text{ nach a.)}\\ &\le \varlimsup_{n\to\infty} \mathrm{E}[\,X_{n\wedge\tau}\,] = \varlimsup_{n\to\infty} (\mathrm{E}[\,X^+_{n\wedge\tau}\,] - \mathrm{E}[\,X^-_{n\wedge\tau}\,])\\ &\le \mathrm{E}[\,X^+_\tau\,] - \mathrm{E}[\,X^-_\tau\,] = \mathrm{E}[\,X_\tau\,] \le \mathrm{E}[\,X^+_\tau\,] < \infty\end{aligned}$$

q.e.d.

Bemerkung: Ist X ein Martingal, so gilt in a.) und b.) (falls $(X_{\tau\wedge n})_{n\in\mathbb{N}_0}$ gleichmäßig integrierbar) jeweils die Gleichheit (setze X und $-X$ ein). Dies gilt ebenso für (4.5).

Als Anwendung in der Spieltheorie, das sogenannte Ruinproblem:

Beispiel: Seien $(Z_n)_{n\in\mathbb{N}}$ unabhängige, identisch verteilte $\{\,-1,1\,\}$-wertige Zufallsvariablen und $\mathbb{F}^Z$ die kanonische Filterung mit $\mathcal{F}^Z_0 := \{\,\emptyset,\Omega\,\}$. Mit $a\in\mathbb{N}_0$ ist $X_n := a + \sum_{i=1}^n Z_i$ eine Irrfahrt auf $\mathbb{Z}$. Ist a das Startkapital eines Spielers und $Z_n = 1$ (bzw. -1) der Gewinn (bzw. Verlust) einer Einheit im n-ten Spiel, so gibt X_n den Kapitalstand nach dem n-ten Spiel wieder. Ist $0 < p := \mathrm{P}[\,Z_n = 1\,]$ die Gewinnwahrscheinlichkeit, so bezeichnet $\varrho := \dfrac{1-p}{p}$ das Verlust-Gewinnverhältnis.

Der Spieler gebe sich ein Ziel $c\in\mathbb{N}_0$ vor, bei dessen Erreichen er aufhört zu spielen, das Erreichen von 0 (gleichbedeutend mit dem Ruin) führt zum zwangsweisen Abbruch.
Man hat das Abbruchkriterium $\tau := \inf\{\,n\ge 0 : X_n \in \{\,0,c\,\}\,\}$, $\inf\emptyset := +\infty$.

τ ist eine Stoppzeit, da $\{\,\tau\le n\,\} = \bigcup_{i=0}^n \{\,X_i\in\{\,0,c\,\}\,\} \in \mathcal{F}^Z_n$, $n\in\mathbb{N}_0$. Wegen $\{\,\tau<\infty\,\} \supset \varlimsup_{n\to\infty}\{\,Z_{nc+i}=1\,,\ 0\le i<c\}$ und $\mathrm{P}[\,Z_{nc+i}=1\,,\ 0\le i<c\,] = p^c > 0$ erhält man mit Borel-Cantelli: $1 = \mathrm{P}[\,\varlimsup_{n\to\infty}\{\,Z_{nc+i}=1\,,\ 0\le i<c\,\}\,] \le \mathrm{P}[\,\tau<\infty\,]$, d.h. $\mathrm{P}[\,\tau<\infty\,] = 1$.

Betrachte nun die Funktion $h(x) := \begin{cases} \varrho^x & \text{für } p\ne\frac12\\ x & \text{für } p=\frac12\end{cases}$

Mit h gilt: $\big(h(X_n)\big)_{n\in\mathbb{N}_0}$ ist ein Martingal.

<u>Beweis:</u> Der Fall $p=\dfrac12$ wurde in einem der obigen Beispiele bereits behandelt. Für $p\ne\dfrac12$:
Es ist $\mathrm{E}[\,h(X_{n+1})\mid\mathcal{F}^Z_n\,](\,\cdot\,) \equiv \mathrm{E}[\,h(X_{n+1})\mid(Z_1,\dots,Z_n)\,](\,\cdot\,)$ nach Definition 6 in Abschnitt 1.

$$\begin{aligned}&\mathrm{E}[\,h(X_{n+1})\mid(Z_1,\dots,Z_n)=(z_1,\dots,z_n)\,]\\ &= \mathrm{E}[\,h(x+Z_{n+1})\mid(Z_1,\dots,Z_n)=(z_1,\dots,z_n)\,] \quad \text{falls } x = a+z_1+\ldots+z_n\\ &= \mathrm{E}[\,h(x+Z_{n+1})\,] \quad \text{nach dem Spezialfall von (1.7)}\\ &= p\cdot h(x+1) + (1-p)\cdot h(x-1) \equiv h(x)\end{aligned}$$

Mithin ist $\mathrm{E}[\,h(X_{n+1})\mid(Z_1,\dots,Z_n)\,] = h(X_n)$ gemäß der Bemerkung nach Definition 6 in Abschnitt 1 .

Die Integrabilität folgt, da X_n nur endlich viele reelle Werte annimmt.

Wegen $0\le X_{n\wedge\tau}\le c$ folgt: $0\le h(X_{n\wedge\tau}) \le \max_{0\le x\le c} h(x)$

Deshalb ergibt sich zusammen mit (4.6) b.):

$$h(a) \equiv \mathrm{E}[\,h(X_0)\,] = \mathrm{E}[\,h(X_\tau)\,] \equiv h(c) \cdot \mathrm{P}[\,X_\tau = c\,] + h(0) \cdot \mathrm{P}[\,X_\tau = 0\,]$$

$$\equiv h(c) \cdot (1 - \mathrm{P}[\,X_\tau = 0\,]) + h(0) \cdot \mathrm{P}[\,X_\tau = 0\,]$$

$\Longrightarrow \mathrm{P}[\,X_\tau = 0\,] = \dfrac{h(c) - h(a)}{h(c) - h(0)}$ ist die Ruin-Wahrscheinlichkeit, d.h. die Wahrscheinlichkeit, eher 0 als c zu erreichen.

Die Wahrscheinlichkeit, bei einem Spiel ohne Limit einmal ruiniert zu werden (d.h. $X_n = 0$ für ein $n \in \mathbb{N}_0$), errechnet sich damit wie folgt:

$$\begin{aligned}
\mathrm{P}[\bigcup_{n=0}^{\infty}\{\,X_n = 0\,\}\,] &= \mathrm{P}[\bigcup_{n=0}^{\infty}\;\bigcup_{c=0}^{\infty}\{\,X_n = 0\,,\; 0 < X_k < c \quad \forall\, k < n\,\}\,] \quad (\text{ leicht })\\
&= \mathrm{P}[\bigcup_{c=0}^{\infty}\;\underbrace{\bigcup_{n=0}^{\infty}\{\,X_n = 0\,,\; 0 < X_k < c \quad \forall\, k < n\,\}}_{M_c}\,]\\
&= \lim_{c\to\infty} \mathrm{P}[\,M_c\,] \quad \text{da } M_{c_1} \subset M_{c_2} \text{ für } c_1 < c_2\\
&\overset{(*)}{=} \lim_{c\to\infty} \frac{h(c) - h(a)}{h(c) - h(0)} = \begin{cases} 1 & \text{für } p \le \frac{1}{2} \\ \varrho^a & \text{für } p > \frac{1}{2} \end{cases}
\end{aligned}$$

$(*)$ gilt, da M_c das Ereignis ist, bei der Vorgabe eines Ziels c vor dessen Erreichen ruiniert zu werden.

Die Ermittlung der Ruin-Wahrscheinlichkeit beim Ruinproblem ist ein erstes Beispiel eines "Problems des ersten Eintritts in eine Ereignismenge", τ heißt die Eintrittszeit. Ein ähnliches Problem wird nun als Anwendung von (4.5) vorgestellt, das sogenannte "Überquerungsproblem":

Beweis der Doob'schen Ungleichung:

Die Doob'sche Ungleichung wird eine Abschätzung für das Oszillationsverhalten eines Submartingals liefern, was für die späteren Konvergenzsätze von Bedeutung sein wird.
Seien dazu ein Submartingal $X = (X_n)_{n\ge 0}$ bzgl. $\mathbb{F} = (\mathcal{F}_n)_{n\ge 0}$ und zwei Schranken r, s mit $r < s$ vorgegeben. Um das Überquerungsverhalten von X im Bezug auf das Intervall $[\,r, s\,]$ zu beschreiben, definiert man:

$$\begin{aligned}
\tau_0 &:= -1\\
\sigma_m &:= \inf\{\,j > \tau_{m-1} : X_j \le r\,\} \quad \text{für } m \in \mathbb{N}\\
\tau_m &:= \inf\{\,j > \sigma_m : X_j \ge s\,\} \quad \text{für } m \in \mathbb{N}
\end{aligned}$$

Dabei sei jeweils $\inf \emptyset := +\infty$. σ_m, $m \in \mathbb{N}$, gibt den nächsten Zeitpunkt nach τ_{m-1} an, in dem X ins Intervall $(-\infty, r]$ eintritt $\big($ also eine vollständige Überquerung von (r, s) nach unten hin $\big)$. Für τ_m gilt das entsprechende. Man zeigt leicht, daß τ_m, σ_m Stoppzeiten sind.
Als Kontrollprozeß U für die Oszillationsbewegung von X definiert man für $i \in \mathbb{N}_0$:

$$U_i := \begin{cases} 1 & \text{auf } \{\,\sigma_m \le i < \tau_m\,\} \\ 0 & \text{auf } \{\,\tau_m \le i < \sigma_{m+1}\,\} \end{cases}$$

U ist Kontrolle im obigen Sinne, da U an $\mathbb{F}$ adaptiert mittels $\{\,U_i = 1\,\} = \bigcup_{m=1}^{\infty}\{\,\sigma_m \le i\,\} \cap \{\,\tau_m \le i\,\}^C \in \mathcal{F}_i$
U_i hat also den Wert 1, falls sich X im Zeitpunkt i in einer Aufwärtsüberquerung von (r, s) befindet, 0 sonst.

Als weitere Kontrollgröße definiert man sich die Zufallsvariable $U_n[r,s] := \max\{m \in \mathbb{N}_0 : \tau_m \leq n\}$. $U_n[r,s]$ beschreibt die Anzahl der aufsteigenden Überquerungen des Intervalls (r,s) im Zeitraum $\{0,\ldots,n\}$. Deren mittlere Anzahl soll nun abgeschätzt werden. Man stutzt dazu den Prozeß X zu $\widetilde{X}$, wo

$$\widetilde{X}_n := r \vee X_n \equiv r + (X_n - r)^+ \geq r$$

$\widetilde{X}_n$ ist nach (4.1) ein Submartingal und besitzt durch die Stutzung die Eigenschaften

$$\text{i.)}\quad \widetilde{X}_n - \widetilde{X}_{\sigma_m} \geq 0 \qquad \text{ii.)}\quad \widetilde{X}_{\tau_m} - \widetilde{X}_{\sigma_m} \geq s - r \quad (\text{ leicht })$$

Daher gilt:

$$\begin{aligned}
\widetilde{X}_n^U &\equiv \widetilde{X}_0 + \sum_{i=1}^{n} U_{i-1}(\widetilde{X}_i - \widetilde{X}_{i-1}) = \widetilde{X}_0 + \sum_{i=1}^{n}\sum_{m=1}^{\infty} 1_{\{\sigma_m \leq i-1 < \tau_m\}}(\widetilde{X}_i - \widetilde{X}_{i-1}) \\
&= \widetilde{X}_0 + \sum_{m=1}^{\infty}\sum_{i=1}^{n} 1_{\{\sigma_m \leq i-1 < \tau_m\}}(\widetilde{X}_i - \widetilde{X}_{i-1}) = \widetilde{X}_0 + \sum_{m=1}^{\infty} \sum_{\big((\sigma_m+1)\wedge(n-1)\big) \leq i \leq (\tau_m \wedge n)} (\widetilde{X}_i - \widetilde{X}_{i-1}) \\
&= \widetilde{X}_0 + \sum_{m=1}^{\infty} (\widetilde{X}_{n\wedge\tau_m} - \widetilde{X}_{n\wedge\sigma_m}) \\
&= \widetilde{X}_0 + \sum_{m=1}^{U_n[r,s]} (\widetilde{X}_{\tau_m} - \widetilde{X}_{\sigma_m}) + (\widetilde{X}_n - \widetilde{X}_{\sigma_{(U_n[r,s]+1)}}) \cdot 1_{\{\sigma_{(U_n[r,s]+1)} < n\}} \quad \text{da } \tau_m > \sigma_m \geq m \\
&\geq r + \sum_{m=1}^{U_n[r,s]} (\widetilde{X}_{\tau_m} - \widetilde{X}_{\sigma_m}) \quad \text{wegen i.) und } \widetilde{X}_0 \geq r \\
&\geq r + U_n[r,s] \cdot (s-r) \quad \text{wegen ii.)}
\end{aligned}$$

Andererseits gilt nach (4.5): $\mathrm{E}[\widetilde{X}_n] \equiv r + \mathrm{E}[(X_n - r)^+] \geq \mathrm{E}[\widetilde{X}_n^U] \geq r + (s-r)\cdot \mathrm{E}\big[U_n[r,s]\big]$

Dies ergibt in der Summation aller Schritte den Beweis von

Proposition 4.7: (Doob'sche Ungleichung)

Sei $(X_n)_{n\geq 0}$ ein Submartingal und $U_n[r,s]$ wie oben definiert. Dann gilt:

$$\mathrm{E}\big[U_n[r,s]\big] \leq \frac{1}{s-r} \cdot \mathrm{E}[(X_n - r)^+]$$

Bemerkung 4.8: Man kann ein Submartingal "auffüllen":

Sei $I \subset \mathbb{Z}$ und $X = (X_n)_{n\in I}$ ein Submartingal bzgl. $(\mathcal{F}_n)_{n\in I}$. Dann existiert ein Submartingal $\widetilde{X} = (X_n)_{n\in\mathbb{Z}}$ bzgl. $(\mathcal{F}_n)_{n\in\mathbb{Z}}$, welches X "fortsetzt".
Definiere zur Fortsetzung nach rechts für $n > s \equiv \sup I$: $X_n := X_s$ und $\mathcal{F}_n := \mathcal{F}_s$.
Die Fortsetzung nach links geschieht analog.

Nun hat man alle technischen Voraussetzungen, um den Hauptsatz dieses Abschnitts zu beweisen. Aufgrund von (4.8) sei im weiteren o.E. $I = \mathbb{Z}$.

Satz 4.9: (Martingal-Konvergenzsatz)

Sei $X = (X_n)_{n\in\mathbb{Z}}$ ein Submartingal bzgl. der Filterung $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{Z}}$. Setzt man zusätzlich $\mathcal{F}_{-\infty} := \bigcap_{n\in\mathbb{Z}} \mathcal{F}_n$ und $\mathcal{F}_\infty := \sigma(\bigcup_{n\in\mathbb{Z}} \mathcal{F}_n)$, so gelten die Aussagen

a.) $X_{-\infty} := \lim_{n\to-\infty} X_n$ existiert fast sicher.

b.) Ist $\inf_{n\leq 0} \mathrm{E}[X_n] > -\infty$, so ist $X_{-\infty}$ integrierbar und $(X_n)_{n\in\mathbb{Z}\cup\{-\infty\}}$ ein Submartingal.
(Abschluß eines Submartingals nach links)

Unter der zusätzlichen Voraussetzung $\sup\limits_{n\geq 0} \mathrm{E}[X_n^+] < \infty$ ($\Longleftrightarrow \sup\limits_{n\geq 0} \mathrm{E}[\,|X_n|\,] < \infty$)

gelten ferner:

c.) $X_\infty := \lim\limits_{n\to\infty} X_n$ existiert fast sicher und ist integrierbar.

d.) $(X_n)_{n\in\mathbb{Z}\cup\{\infty\}}$ ist ein Submartingal $\Longleftrightarrow$ $(X_n^+)_{n\geq 0}$ ist gleichmäßig integrierbar.

Beweis: a.) Setze $\overline{X} := \overline{\lim\limits_{n\to-\infty}}\, X_n$ und $\underline{X} := \underline{\lim\limits_{n\to-\infty}}\, X_n$.

Dann ist z.z.: $\overline{X} = \underline{X}$ fast sicher, d.h. $\mathrm{P}[\overline{X} > \underline{X}] = 0$ (da stets $\overline{X} \geq \underline{X}$).

Man schreibt dazu die Menge $\{\overline{X} > \underline{X}\}$ als $\{\overline{X} > \underline{X}\} = \bigcup\limits_{\substack{r,s\in\mathbb{Q}\\ r<s}} B(r,s)$

mit $B(r,s) := \{\overline{X} > s > r > \underline{X}\}$. Für $r,s\in\mathbb{Q}$ mit $r<s$ bezeichne $U_n[r,s]$ wieder die Anzahl der Überquerungen des Intervalls $[r,s]$ von $(X_i)_{i\geq -n}$ im Zeitraum $\{-n,\ldots,0\}$. (Um auf die Form von (4.7) zu kommen, gehe man gedanklich zu dem zeitlich verschobenen Prozeß $\widetilde{X}_i := X_{i-n}$ über).

Nach (4.7) gilt daher: $\mathrm{E}\big[U_n[r,s]\big] \leq \dfrac{1}{s-r}\cdot \mathrm{E}[(X_0-r)^+] < \infty$, wobei die letzte Ungleichung aus (4.2) folgt. Da $U_n[r,s]$ isoton in n und $U_n[r,s]\geq 0 \quad \forall\, n\in\mathbb{N}$, hat man aufgrund monotoner Konvergenz und voriger Ungleichung:

$\mathrm{E}\big[\lim\limits_{n\to\infty} U_n[r,s]\big] = \lim\limits_{n\to\infty} \mathrm{E}\big[U_n[r,s]\big] < \infty \Longrightarrow \mathrm{P}\big[\lim\limits_{n\to\infty} U_n[r,s] = \infty\big] = 0$

Wegen $(*)$ $B(r,s) \subset \{\lim\limits_{n\to\infty} U_n[r,s] = \infty\}$ folgt $\mathrm{P}[B(r,s)] = 0 \quad \forall\, r,s\in\mathbb{Q}$ mit $r<s$

$\Longrightarrow \mathrm{P}[\overline{X} > \underline{X}] = 0$

(Beweis von $(*)$: $\omega\in B(r,s) \Longrightarrow |\{m : X_m(\omega) > s\}| = |\{m : X_m(\omega) < r\}| = \infty$)

b.) Da X ein Submartingal ist, gelten: (i) $n \mapsto \mathrm{E}[X_n]$ ist isoton in $n\in\mathbb{Z}$

(ii) $(X_n^+)_{n\in\mathbb{Z}}$ ist nach (4.1) ebenfalls ein Submartingal, daher ist die Abbildung $n\mapsto \mathrm{E}[X_n^+]$ ist ebenfalls isoton in $n\in\mathbb{Z}$.

$$\begin{aligned}
\Longrightarrow \quad 0 \leq \mathrm{E}[X_{-\infty}^+] &\leq \underline{\lim_{n\to-\infty}}\; \mathrm{E}[X_n^+] \quad \text{nach Fatou}\\
&\leq \mathrm{E}[X_0^+] \;\; (<\infty) \quad \text{nach ii.)}\\
\text{und } 0 \leq \mathrm{E}[X_{-\infty}^-] &\leq \underline{\lim_{n\to-\infty}}\; \mathrm{E}[X_n^-] \quad \text{nach Fatou}\\
&\equiv \sup_{n\leq 0}\; \inf_{m\leq n} (\mathrm{E}[X_m^+] - \mathrm{E}[X_m])\\
&\leq \sup_{n\leq 0} (\mathrm{E}[X_n^+] - \mathrm{E}[X_n])\\
&\leq \mathrm{E}[X_0^+] - \inf_{n\leq 0} \mathrm{E}[X_n] \;\; (<\infty) \quad \text{nach (i) und (ii)}
\end{aligned}$$

$\Longrightarrow X_{-\infty}$ ist integrierbar.

Da ferner $X_{-\infty} \stackrel{\text{Def.}}{=} \underline{\lim\limits_{n\to-\infty}}\; X_n = \sup\limits_{n\leq -k}\; \inf\limits_{m\leq n} X_m$, d.h. $X_{-\infty}$ ist $\mathcal{F}_{-k}$-meßbar $\quad\forall\, k\in\mathbb{N}$, ist $X_{-\infty}$ $\mathcal{F}_{-\infty}$-meßbar. Für $a\in\mathbb{R}$ ist nach (4.4) der gestutzte Prozeß $(a\vee X_n)_{-\infty<n\leq 0}$ gleichmäßig integrierbar.

Es gilt daher für $m, n \in \mathbb{Z}$ mit $-\infty < m < n < \infty$ und $A \in \mathcal{F}_{-\infty} \subset \mathcal{F}_m \subset \mathcal{F}_n$:

$\int\limits_A a \vee X_m \, d\mathrm{P} \leq \int\limits_A a \vee X_n \, d\mathrm{P}$ wegen der Submartingaleigenschaft

$\Longrightarrow \int\limits_A a \vee X_{-\infty} \, d\mathrm{P} \leq \int\limits_A a \vee X_n \, d\mathrm{P}$ nach (4.3), da mit $(a \vee X_n)_{-\infty < n \leq 0}$ auch $\big(1_A \cdot (a \vee X_N)\big)_{-\infty < n \leq 0}$ gleichmäßig integrierbar ist. Aus $a \vee X_{-\infty} \downarrow X_{-\infty}$ bzw. $a \vee X_n \downarrow X_n$ für $a \to -\infty$ folgt mit monotoner Konvergenz $\int\limits_A X_{-\infty} \, d\mathrm{P} \leq \int\limits_A X_n \, d\mathrm{P}$.

c.) Beh.: $\sup\limits_{n \geq 0} \mathrm{E}[\, X_n^+ \,] < \infty \iff \sup\limits_{n \geq 0} \mathrm{E}[\, |\, X_n \,|\,] < \infty$

Beweis: "$\Leftarrow$" ist klar, da $|\, X_n \,| = X_n^+ + X_n^- \geq X_n^+$.

"$\Rightarrow$": $|\, X_n \,| = X_n^+ + X_n^- = 2 \cdot X_n^+ - X_n$

$$\Longrightarrow \sup_{n \geq 0} \mathrm{E}[\, |\, X_n \,|\,] = \sup_{n \geq 0} (2 \cdot \mathrm{E}[\, X_n^+ \,] - \mathrm{E}[\, X_n \,]) \leq 2 \cdot \sup_{n \geq 0} \mathrm{E}[\, X_n^+ \,] - \mathrm{E}[\, X_0 \,] < \infty$$

Die Existenz von X_∞ zeigt man nun analog zu a.):

Ist $U_n[\, r, s \,]$ die Anzahl der aufsteigenden Überquerungen des Intervalls $[\, r, s \,]$ $\big(\, r, s$ wie in a.) $\big)$ von $(X_n)_{n \geq 0}$ im Zeitraum $\{0, \ldots, n\}$, so hat man:

$$\mathrm{E}\big[\, U_n[\, r, s \,]\big] \leq \frac{1}{s - r} \mathrm{E}[\, (X_n - r)^+ \,] \leq \frac{1}{s - r} (\sup_{n \geq 0} \mathrm{E}[\, X_n^+ \,] + |\, r \,|) < \infty \quad \text{da } (a + b)^+ \leq a^+ + b^+$$

Der Rest des Beweises verläuft wie in a.).

Zur Integrierbarkeit von X_∞: $\mathrm{E}[\, |\, X_\infty \,|\,] \leq \varliminf\limits_{n \to \infty} \mathrm{E}[\, |\, X_n \,|\,]$ nach Fatou

$\leq \sup\limits_{n \geq 0} \mathrm{E}[\, |\, X_n \,|\,] \; (< \infty)$ analog zu a.).

d.) "$\Leftarrow$": Für $a \in \mathbb{R}$ ist $(a \vee X_n)_{n \geq 0}$ gleichmäßig integrierbar, denn per Fallunterscheidung $|\, a \vee X_n| \leq |X_n^+| + |a|$. Der Rest gilt ananlog zu b.).

"$\Rightarrow$": Dies ist die Aussage (4.4) mit $a = 0$

q.e.d.

Korollar 4.9′: 1.) Ist X in (4.9) sogar ein Martingal, so existiert $X_{-\infty}$ und $(X_n)_{n \in \mathbb{Z} \cup \{-\infty\}}$ ist ein Martingal. Die Aussage d.) modifiziert sich zu

$(X_n)_{n \in \mathbb{Z} \cup \{+\infty\}}$ ist ein Martingal $\iff (X_n)_{n \geq 0}$ ist gleichmäßig integrierbar.

Dies erhält man wieder durch Einsetzen von X und $-X$ in die Behauptung von (4.9).

2.) In (4.9) c.) gilt nicht notwendigerweise $\lim\limits_{n \to \infty} \mathrm{E}[\, X_n \,] = \mathrm{E}[\, X_\infty \,]$. Dies ist jedoch stets dann erfüllt, wenn die stärkere Voraussetzung "$(X_n)_{n \geq 0}$ ist gleichmäßig integrierbar" gilt. $\big($ vgl. Aussage (4.9) d.) und (4.3) $\big)$

□

Nun einige Anwendungen im Fall $n \to -\infty$:

Dazu zunächst eine wichtige Charakterisierung für gleichmäßig integrierbare Martingale:

Korollar 4.10: Sei $X = (X_n)_{n \in \mathbb{Z}}$ ein stochastischer Prozeß und $\mathbb{F} = (\mathcal{F}_\mathrm{n})_{\mathrm{n} \in \mathbb{Z}}$ eine Filterung. Dann sind äquivalent:

(i) X ist ein gleichmäßig integrierbares Martingal bzgl. $\mathbb{F}$

(ii) Es existiert eine integrierbare Zufallsvariable Z mit $X_n = \mathrm{E}[\, Z \mid \mathcal{F}_n \,]$

Beweis: (ii) $\Rightarrow$ (i): Nach Beispiel b.) hinter Definition 2 existiert ein $\mathcal{F}_\infty$ und damit $X_\infty \equiv \mathrm{E}[\,Z \mid \mathcal{F}_\infty\,]$, so daß $\widetilde{X} := (X_n)_{n\leq\infty}$ ein Martingal ist. Daher ist nach (4.4) $\widetilde{X}$ gleichmäßig integrierbar, da $\widetilde{X}^+$ und $\widetilde{X}^-$ gleichmäßig integrierbar und somit

$$\begin{aligned}\mathrm{E}[\,(\mid X\mid -2\cdot\alpha)^+\,] &= \mathrm{E}[\,(X^+ + X^- - 2\cdot\alpha)^+\,]\\ &\leq \mathrm{E}[\,(X^+ - \alpha)^+\,] + \mathrm{E}[\,(X^- - \alpha)^+\,] \longrightarrow 0 \quad (\alpha\to\infty)\end{aligned}$$

(i) $\Rightarrow$ (ii): Nach (4.9) c.), d.) existiert $X_\infty := \lim\limits_{n\to\infty} X_n$ fast sicher, nach Korollar 4.9$'$ ist $(X_n)_{n\leq\infty}$ ein Martingal, d.h. $\mathrm{E}[\,X_\infty \mid \mathcal{F}_n\,] = X_n$ fast sicher, also $X_n \in \mathbb{E}[\,X_\infty \mid \mathcal{F}_n\,]$. Mit anderen Worten: $Z := X_\infty$.

q.e.d.

Im Beweisteil (ii) $\Rightarrow$ (i) war von X_∞ die Rede. Es wird nun im nächsten Korollar gezeigt, daß die Schreibweise X_∞ ihre Berechtigung hat, d.h. daß gilt $X_\infty = \lim\limits_{n\to\infty} X_n$:

Korollar 4.11: (Lévy'scher Martingalsatz)
Sei Z eine integrierbare Zufallsvariable, $\mathbb{F} = (\mathcal{F}_\mathrm{n})_{\mathrm{n}\in\mathbb{N}_0}$ eine Filterung.
Mit der Setzung $\mathcal{F}_\infty := \sigma(\bigcup\limits_{n\in\mathbb{N}_0} \mathcal{F}_n)$ gilt dann:

$$\mathrm{E}[\,Z \mid \mathcal{F}_n\,] \longrightarrow \mathrm{E}[\,Z \mid \mathcal{F}_\infty\,],\ (n\to\infty).$$

Beweis: O.E. $Z \geq 0$, zerlege sonst Z in $Z^+ - Z^-$ und führe getrennte Betrachtungen durch.
In diesem Fall ist $(\mathrm{E}[\,Z \mid \mathcal{F}_n\,])_{n\in\mathbb{N}_0}$ nicht negativ und gemäß (4.10) ein gleichmäßig integrierbares Martingal. Nach (4.9) c.) und (4.9$'$) existiert $X_\infty = \lim\limits_{n\to\infty} \mathrm{E}[\,Z \mid \mathcal{F}_n\,]$ fast sicher, so daß $(\mathrm{E}[\,Z \mid \mathcal{F}_n\,], X_\infty)_{n\in\mathbb{N}_0}$ ein Martingal ist. Daher hat man für $A \in \mathcal{F}_n \subset \mathcal{F}_\infty$:

$$\begin{aligned}\int_A X_\infty\, d\mathrm{P} &= \int_A \mathrm{E}[\,Z \mid \mathcal{F}_n\,]\, d\mathrm{P} \quad \text{nach Martingaldefinition}\\ &\equiv \int_A Z\, d\mathrm{P} \quad \text{nach Definition der bedingten Erwartung.}\end{aligned}$$

X_∞ , Z sind nichtnegativ und integrierbar, mithin sind die Abbildungen $A \mapsto \int_A Z\, d\mathrm{P}$ und (aufgrund der Gleichhheit) $A \mapsto \int_A X_\infty\, d\mathrm{P}$ endliche Maße auf $\mathcal{F}_n$, die dort übereinstimmen $\forall\, n$.
Aus dem Eindeutigkeitssatz für Maße (W-Theorie 7.5) folgt, daß die Gleichung

$$\int_A X_\infty\, d\mathrm{P} = \int_A Z\, d\mathrm{P} \quad \forall\, A \in \mathcal{F}_\infty \equiv \sigma(\bigcup_{n\in\mathbb{N}_0} \mathcal{F}_n) \ \text{ stattfindet.}$$

Dies heißt aber nach Charakterisierung der bedingten Erwartung, daß $X_\infty = \mathrm{E}[\,Z \mid \mathcal{F}_\infty\,]$.

q.e.d.

Eine fundamentale Anwendung von (4.9) im Fall $n \to -\infty$ ist der Beweis des starken Gesetzes der großen Zahlen, welcher nun getätigt wird. Die zugrundeliegende Idee ist, das zu untersuchende arithmetische Mittel als Martingal aufzufassen:

Seien $(Z_i)_{i\in\mathbb{N}}$ unabhängige, identisch verteilte, integrierbare Zufallsvariablen, $S_n := \sum\limits_{i=1}^{n} Z_i$. Dann gilt:

$$\frac{1}{n} S_n \stackrel{(1)}{=} \mathrm{E}[\,Z_1 \mid S_n\,] \stackrel{(2)}{=} \mathrm{E}[\,Z_1 \mid S_n, S_{n+1}, \ldots].$$

Beweis: (2) folgt mit Hilfe von $\sigma(S_n, S_{n+1}, \ldots) = \sigma(S_n, Z_m, m > n)$

"$\subset$" folgt dabei aus $S_m = S_n + Z_{n+1} + \ldots + Z_m$; "$\supset$" aus $Z_m = S_m - S_{m-1}$, $m > n$.

Daher ist (1.7) b.) anwendbar, setze dort $X := (Z_m, m > n)$, $Y := Z_1$ sowie $Z := S_n$.

(1) besteht, da $\mathrm{E}[\, Z_1 \mid S_n \,] = \mathrm{E}[\, Z_j \mid S_n \,]$, $j = 1, \ldots, n$.

Denn: $A \in \sigma(S_n) \Longrightarrow A = S_n^{-1}(B)$ mit $B \in \mathcal{B}$

$$\Longrightarrow \int_A Z_j \, d\mathrm{P} = \int_\Omega (1_B \circ S_n) \cdot Z_j \, d\mathrm{P}$$

$$\equiv \int_{\mathbb{R}^n} (1_B \circ s) \cdot \Pi_j \, d(\mathrm{P}Z_1^{-1} \otimes \ldots \otimes \mathrm{P}Z_n^{-1}) \quad (\text{ Unabhängigkeit })$$

$$\equiv \int_{\mathbb{R}^n} (1_B \circ s) \cdot \Pi_j \, d(\bigtimes_{i=1}^{n} \mu)$$

wobei s Summationsabbildung, Π_j die Projektion auf die j-te Komponente und $\mu := \mathrm{P}Z_1^{-1}$. Da s eine symmetrische Funktion ist, also invariant unter Permutationen, zeigt man mit dem Satz von Fubini, daß obiges Integral von j unabhängig ist.

Dies ergibt: $\dfrac{1}{n} S_n \equiv \dfrac{1}{n} \mathrm{E}[\, S_n \mid S_n \,] = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \mathrm{E}[\, Z_i \mid S_n \,] = \mathrm{E}[\, Z_1 \mid S_n \,]$

□

Um die Resultate von (4.9) im Fall $n \to -\infty$ nutzen zu können, erfolgt nun eine Zeitumkehrung:
Definiere $\widetilde{Z}_{-n} := Z_n$, $\widetilde{S}_{-n} := \widetilde{Z}_{-n} + \ldots + \widetilde{Z}_{-1}$, d.h. nach obigem:
$\dfrac{1}{n} \widetilde{S}_{-n} = \mathrm{E}[\, \widetilde{Z}_{-1} \mid \widetilde{S}_{-n} \,] = \mathrm{E}[\, \widetilde{Z}_{-1} \mid \ldots, \widetilde{S}_{-n-1}, \widetilde{S}_{-n} \,]$

Damit nun das

Lemma 4.12: Seien $(Z_n)_{n \in \mathbb{N}}$ und $(S_n)_{n \in \mathbb{N}}$ wie oben definiert. Durch die Setzungen $X_{-n} := \dfrac{1}{n} S_n$ und $\mathcal{F}_{-n} := \sigma(S_m, m \geq n)$, $n \in \mathbb{N}$, wird $(X_n)_{n<0}$ zu einem Martingal bzgl. $(\mathcal{F}_n)_{n<0}$.

Beweis: 1.) $(\mathcal{F}_n)_{n<0}$ ist eine Filterung, da $\mathcal{F}_{-n} \subset \mathcal{F}_{-n+1} \equiv \mathcal{F}_{-(n-1)}$. $(X_n)_{n<0}$ ist adaptiert an $(\mathcal{F}_n)_{n<0}$, da $\sigma(X_{-n}) \equiv \sigma(S_n) \subset \sigma(S_m, m \geq n) \equiv \mathcal{F}_{-n}$.

2.) Nach obigem gilt $\mathrm{E}[\, Z_1 \mid S_n \,] = \dfrac{1}{n} S_n \quad (= \mathrm{E}[\, Z_i \mid S_n \,], 1 \leq i \leq n \,)$ und
$\mathrm{E}[\, Z_1 \mid S_n \,] = \mathrm{E}[\, Z_1 \mid S_n, S_{n+1}, \ldots \,] \equiv \mathrm{E}[\, Z_1 \mid \mathcal{F}_{-n} \,]$.
Mit anderen Worten: $X_{-n} = \mathrm{E}[\, Z_1 \mid \mathcal{F}_{-n} \,]$. Nach einer Auffüllung nach rechts $\big($ vgl. die Bemerkung vor (4.9) $\big)$ liefert (4.10) die Behauptung.

q.e.d.

Damit ist nun die Idee realisiert, $\dfrac{1}{n} S_n$ als Martingal darzustellen. Den letzten Schritt des Beweises der Konvergenz zeigt nun:

Satz 4.13: (Starkes Gesetz der großen Zahlen)
Sei $(Z_n)_{n \in \mathbb{N}}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen, deren Erwartungswert $\mathrm{E}[\, Z_n \,] = a$ existiert. Dann gilt:

$$\frac{1}{n} \sum_{i=1}^{n} Z_i \equiv \frac{1}{n} S_n \longrightarrow a \quad \text{fast sicher, } (n \to \infty).$$

Beweis: a.) Sei a reell. Dann sind die Z_i integrierbar und nach (4.12) ist $X = (X_n)_{n<0}$ mit $X_n := \frac{1}{n} S_n$ ein Martingal mit der Eigenschaft $\mathrm{E}[X_{-n}] = \mathrm{E}[X_{-1}] \equiv \mathrm{E}[Z_1] = a$.

Gemäß (4.9) a.) existiert $X_{-\infty} = \lim\limits_{n\to\infty} \frac{1}{n} S_n$ fast sicher. Das 0-1-Gesetz (3.1) und das dort vorangehende Beispiel liefern: $X_{-\infty}$ ist fast sicher konstant.

Da nach (4.9′) $(X_n)_{-\infty\le n<0}$ ein Martingal ist, hat man $\mathrm{E}[X_{-\infty}] = \mathrm{E}[X_{-1}] = a$

$\Longrightarrow X_{-\infty} = a$ fast sicher.

b.) Sei $a = +\infty$ (bei $a = -\infty$ betrachte man $-Z_n$). Um a.) anwenden zu können, stutzt man den Prozeß Z zu $Z^\alpha = (\alpha \wedge Z_n)_{n\in\mathbb{N}}$, $\alpha \in \mathbb{N}$. Es gilt: $-Z_1^- \equiv 0 \wedge Z_1 \le Z_1^\alpha \uparrow Z_1 \quad (\alpha \to \infty)$.
Da $\mathrm{E}[Z_1^\alpha] < \infty$ ($\mathrm{E}[Z_1]$ ist existent), ergibt sich mit monotoner Konvergenz (mit Minorante $-Z_1^-$): $a^\alpha := \mathrm{E}[Z_1^\alpha] \uparrow \mathrm{E}[Z_1] = \infty, \quad (\alpha \to \infty)$.

Z^α ist eine meßbare Funktion von Z, daher sind die $(Z_n^\alpha)_{n\in\mathbb{N}}$ weiterhin identisch verteilt und unhabhängig.

Mithin gilt nach a.) : $\frac{1}{n}\sum\limits_{i=1}^{n} Z_i^\alpha(\omega) \longrightarrow a^\alpha \quad \forall \omega \in A^\alpha$, wo $\mathrm{P}[A^\alpha] = 1$.

Mit $A := \bigcap\limits_{\alpha\in\mathbb{N}} A^\alpha$ ist $\mathrm{P}[A] = 1$ und $\underline{\lim\limits_{n\to\infty}} \frac{1}{n} S_n \ge \underline{\lim\limits_{n\to\infty}} \frac{1}{n} S_n^\alpha = a^\alpha \quad \forall \alpha \in \mathbb{N}$ auf A

$\Longrightarrow \underline{\lim\limits_{n\to\infty}} \frac{1}{n} S_n = \infty$ auf A, d.h. $\lim\limits_{n\to\infty} \frac{1}{n} S_n = \infty$ fast sicher.

q.e.d.

Als letzter Punkt der zeitdiskreten Martingaltheorie noch die für späteres wichtige

Proposition 4.14: Maximal-Ungleichung von Doob

Ist $X = (X_n)_{n\ge 0}$ ein nichtnegatives Submartingal bzgl. $\mathbb{F} = (\mathcal{F}_\mathrm{n})_{\mathrm{n}\ge 0}$ (z.B. $X_n = |M_n|$, M Martingal), so gilt für $\lambda > 0$ und $p \ge 1$:

$$\mathrm{P}[\max_{0\le m\le n} X_m \ge \lambda] \le \frac{1}{\lambda^p}\,\mathrm{E}[X_n^p]$$

Beweis: Es ist $\max\limits_{0\le m\le n} X_m \ge \lambda \Longleftrightarrow \max\limits_{0\le m\le n} X_m^p \ge \lambda^p$.

Mit X ist auch X^p ein Submartingal, denn $X^p = g(X)$, wobei die Abbildung $g(x) := (x^+)^p$ isoton und konvex ist $\big($ vgl. (4.1) $\big)$. Deshalb sei o.E. $p = 1$.

Setze $\tau := \inf\{ m \ge 0 : X_m \ge \lambda \}$, $\inf \emptyset := \infty$.

Dann ist $\{ \tau = m \} = \{ X_l < \lambda, l < m, X_m \ge \lambda \}$ und

$$\lambda\mathrm{P}[\max_{0\le m\le n} X_m \ge \lambda] = \lambda\mathrm{P}[\tau \le n] = \lambda \sum_{m=0}^{n} \mathrm{P}[\tau = m] \le \sum_{m=0}^{n} \int\limits_{\{\tau=m\}} X_m \, d\mathrm{P}$$

$$\le \sum_{m=0}^{n} \int\limits_{\{\tau=m\}} X_n \, d\mathrm{P} \le \mathrm{E}[X_n] \quad \text{da } X_n \ge 0$$

q.e.d.

Im weiteren Verlauf dieser Ausarbeitung wird nun der aus Abschnitt 2 her bekannte Begriff des Markoff-Prozesses weiter vertieft und anhand einiger Beispiele aus der Praxis verdeutlicht.

Abschnitt 5: Zeitdiskrete Markoff-Ketten

Markoff-Ketten beschreiben den zeitlichen Verlauf stochastischer Systeme mit abzählbarem Zustandsraum S. Dabei finden Sprünge von einem Zustand $i \in S$ in einen Zustand $j \in S$ mit einer festen Wahrscheinlichkeit p_{ij} (der sogenannten Übergangswahrscheinlichkeit) statt, welche von der Vorgeschichte/Vergangenheit des Systems unabhängig ist.
Die Abzählbarkeit des Zustandsraums S wird sich später in Meßbarkeitsfragen als großer Vorteil erweisen, da jede auf S (mit der Potenzmenge als σ-Algebra) definierte Funktion meßbar ist.
Zur Handhabung der Übergangswahrscheinlichkeiten ist es zweckmäßig, zu definieren:

Definition 1: Sei S ein abzählbarer Raum. Dann heißt $\mathbb{P} := (p_{ij})_{i,j \in S}$ *Übergangsmatrix* und die p_{ij} *Übergangswahrscheinlichkeiten*, falls $p_{ij} \geq 0\,,\ i,j \in S$ und $\sum\limits_{j \in S} p_{ij} = 1 \quad \forall\, i \in S$.

Die Übergangsmatrizen haben die nützliche Eigenschaft, eine Halbgruppe bezüglich der Matrixmultiplikation zu bilden:

$$\mathbb{P} = \mathbb{P}' \cdot \mathbb{P}'' \iff p_{ik} = \sum_{j \in S} p'_{ij} \cdot p''_{jk}$$

(d.h. das Produkt zweier Übergangsmatrizen ist wieder eine Übergangsmatrix)

Man führt für die Übergangsmatrizen folgende Bezeichnungen ein:

(5.1) $(p_{ij}^{(0)})_{i,j \in S} := \mathbb{P}^0$ sei die Einheitsmatrix.

(5.2) $(p_{ij}^{(n)})_{i,j \in S} := \mathbb{P}^n$ sei das n-fache Produkt von $\mathbb{P}$ mit sich selbst.

Daraus kann man unmittelbar folgern:

(5.3) $\mathbb{P}^{m+n} = \mathbb{P}^m \cdot \mathbb{P}^n$, d.h. $p_{ik}^{(m+n)} = \sum\limits_{j \in S} p_{ij}^{(m)} \cdot p_{jk}^{(n)} \quad \forall\, m,n \in \mathbb{N}_0$.

Den Begriff der Übergangsmatrizen $\mathbb{P}$ (zu S) kann man eineindeutig auf den schon bekannten Begriff der Übergangswahrscheinlichkeiten von S nach S zurückführen mittels der nachfolgenden Beziehung:
Ist $\mathbb{P} = (p_{ij})_{i,j \in S}$ eine Übergangsmatrix, so errechnet sich die zugehörige Übergangswahrscheinlichkeit Q aus $\mathbb{P}$ durch $Q(i,B) := \sum\limits_{j \in B} p_{ij}$ wo $B \in \mathrm{Pot}\,(S)$. Umgekehrt gilt damit $p_{ij} = Q(i,\{\,j\,\})$.

Nach diesen Vorüberlegungen/Vorbereitungen nun die exakte Definition einer Markoff-Kette:

Definition 2: Sei S ein abzählbarer Raum, $X = (X_n)_{n \in \mathbb{N}_0}$ ein stochastischer Prozeß auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Zustandsraum S sowie $\mathbb{F} = (\mathcal{F}_{\mathrm{n}})_{\mathrm{n} \in \mathbb{N}_0}$ eine Filterung in $\mathcal{F}$. Ist X an $\mathbb{F}$ adaptiert, so heißt $(X, \mathbb{F})$ eine (*zeithomogene*) *Markoff-Kette*, falls eine Übergangsmatrix $\mathbb{P} = (p_{ij})_{i,j \in S}$ existiert mit

(5.4) $\mathrm{P}[\,X_{n+1} = j \mid \mathcal{F}_n\,] = p_{ij}$ auf $\{\,X_n = i\,\}$ für $i,j \in S\,,\ n \in \mathbb{N}_0$

(d.h. $p_{X_n(\,\cdot\,),j}$ ist eine Version von $\mathrm{P}[\,X_{n+1} = j \mid \mathcal{F}_n\,](\,\cdot\,)$)
Dabei wird $\mathrm{P}X_0^{-1}$ als *Startverteilung* von X bezeichnet.

Bemerkungen: 1.) Die Bedingung (5.4) sagt aus, daß der Zustand des Systems in der Zukunft $n+1$ (hier: $X_{n+1}=j$) von der Vergangenheit inclusive den Ereignissen der Gegenwart (hier: $\mathcal{F}_n$) nur über den gegenwärtigen Zustand (hier: $X_n = i$) abhängt.
Dabei bedeutet (5.4) gerade, daß für $A \in \mathcal{F}_n$, $i,j \in S$ sowie $n \in \mathbb{N}_0$

(5.5) $$\mathrm{P}[A \cap \{X_{n+1}=j\}] = \sum_{i \in S} \mathrm{P}[A \cap \{X_n = i\}] \cdot p_{ij}$$

Beweis: Nach (5.4) ist $\mathrm{P}[X_{n+1}=j \mid \mathcal{F}_n] = p_{X_n,j}$

$$\Longrightarrow \mathrm{P}[A \cap \{X_{n+1}=j\}] \equiv \int_A 1_{\{X_{n+1}=j\}}\, d\mathrm{P} \equiv \int_A \mathrm{P}[X_{n+1}=j \mid \mathcal{F}_n]\, d\mathrm{P}$$
$$= \int_A p_{X_n,j}\, d\mathrm{P} = \sum_{i \in S} p_{ij} \int_A 1_{\{X_n=i\}} d\mathrm{P}$$
$$\equiv \sum_{i \in S} \mathrm{P}[A \cap \{X_n = i\}] \cdot p_{ij}$$

q.e.d.

Aus (5.5) hat man für $A \in \mathcal{F}_n$, $i,j \in S$, $n \in \mathbb{N}_0$ noch die Gleichung $\mathrm{P}[X_{n+1}=j \mid \{X_n = i\} \cap A] = p_{ij}$ falls $\mathrm{P}[\{X_n=i\} \cap A] > 0$, da

$$\mathrm{P}[X_{n+1}=j \mid \{X_n = i\} \cap A] \equiv \frac{\mathrm{P}[\{X_{n+1}=j\} \cap \{X_n=i\} \cap A]}{\mathrm{P}[\{X_n = i\} \cap A]}$$
$$\overset{(5.5)}{=} \sum_{l \in S} \frac{\mathrm{P}[\{X_n=i\} \cap A \cap \{X_n = i\}]}{\mathrm{P}[\{X_n=i\} \cap A]} \cdot p_{ij}$$
$$= p_{ij}$$

Wichtige Fälle: $A = \Omega$, $A = \{X_0 = i_0, \ldots, X_{n-1} = i_{n-1}\}$

2.) Zur zeitlichen Homogenität betrachte man die Bemerkung in Abschnitt 2.

Mit Hilfe von (5.5) hat man auch die Möglichkeit, die Wahrscheinlichkeit für das Eintreten einer gewissen Kette von zukünftigen Ereignissen zu berechnen. Dies zeigt

Lemma 5.6: Sei $(X, \mathbb{F})$ eine Markoff-Kette. Dann gilt für $n \in \mathbb{N}_0$, $t \in \mathbb{N}$ und $j_0, \ldots, j_t \in S$:
$\mathrm{P}[X_{n+1}=j_1, \ldots, X_{n+t}=j_t \mid \mathcal{F}_n] = p_{j_0j_1} \cdot \ldots \cdot p_{j_{t-1}j_t}$ auf $\{X_n = j_0\}$.
Insbesondere: $\mathrm{P}[X_0=j_0, \ldots, X_t = j_t] = \mathrm{P}[X_0 = j_0] \cdot p_{j_0j_1} \cdot \ldots \cdot p_{j_{t-1}j_t}$

Beweis: Es ist für $A \in \mathcal{F}_n$ zu zeigen:

$$\mathrm{P}[\{X_{n+1}=j_1\} \cap \ldots \cap \{X_{n+t}=j_t\} \cap A] = \sum_{j_0 \in S} \mathrm{P}[\{X_n = j_0\} \cap A] \cdot p_{j_0j_1} \cdot \ldots \cdot p_{j_{t-1}j_t}$$

$$\mathrm{P}[\bigcap_{k=1}^{t} \{X_{n+k}=j_k\} \cap A] = \mathrm{P}[\bigcap_{k=1}^{t-1} \{X_{n+k}=j_k\} \cap A] \cdot p_{j_{t-1}j_t}$$
$$= \underset{\cdots\cdots\cdots}{\text{induktiv}} = \sum_{j_0 \in S} \mathrm{P}[\{X_n = j_0\} \cap A] \cdot p_{j_0j_1} \cdot \ldots \cdot p_{j_{t-1}j_t}$$

q.e.d.

Aus Vereinfachungsgründen betrachtet man nun im weiteren den aus Abschnitt 2 mit Hilfe des Satzes von Ionescu-Tulcea gewonnenen, aus Projektionen bestehenden Markoff-Prozeß (vgl. (2.5)). Dabei wird die oben angesprochene Tatsache benutzt, daß eine vorgegebene Übergangsmatrix $\mathbb{P}$ eine eindeutig bestimmte Übergangswahrscheinlichkeit von S nach S definiert.

Definition 3: Seien $\nu \in \mathbb{P}(S)$ und eine Übergangsmatrix $\mathbb{P}$ gegeben. Mit P_ν bezeichne man das nach (2.5) existierende Wahrscheinlichkeitsmaß auf $(\Omega^0, \mathcal{F}^0) := \left(\mathop{\times}\limits_{j=0}^{\infty} S, \bigotimes\limits_{j=0}^{\infty} \mathrm{Pot}(S)\right)$, so daß für den Projektionsprozeß $X^0 = (X_n^0)_{n \in \mathbb{N}_0}$ (d.h. X_n^0 ist die Projektion von $\mathop{\times}\limits_{j=0}^{\infty} S$ auf die n-te Komponente) mit $\mathcal{F}_n^0 := \sigma(X_m^0, 0 \le m \le n)$ gilt:
$(X^0, \mathbb{F}^0)$ ist unter P_ν eine Markoff-Kette mit Übergangsmatrix $\mathbb{P}$, Startverteilung ν und Zustandsraum S. Ist die Startverteilung ν auf $i \in S$ konzentriert $\big($ d.h. $\nu(\{ i \}) = 1 \big)$, so schreibt man $\mathrm{P}_i := \mathrm{P}_\nu$. Analog: E_ν bzw. E_i als Erwartungswert unter P_ν bzw. P_i.
$(X^0, \{ \mathrm{P}_i : i \in S \})$ wird auch als *interne Markoff-Familie* zu $(X, \mathbb{F}, \mathrm{P})$ bezeichnet, wo $(X, \mathbb{F})$ unter P eine Markoff-Kette auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Startverteilung $\nu = \mathrm{P}X_0^{-1}$ und Übergangsmatrix $\mathbb{P}$ ist.

Bemerkung: $(X^0, \{ \mathrm{P}_i : i \in S \})$ ist auch interne Markoff-Familie zu $(X^0, \mathbb{F}^0, \mathrm{P}_\nu)$, da $(X^0)^0 = X^0$ etc.

Der entscheidende Vorteil des Übergangs zum kanonischen Prozeß $(X^0, \mathbb{F}^0)$ unter P_ν ist die Tatsache, daß alle Informationen im Wahrscheinlichkeitsmaß P_ν gespeichert sind. Mit anderen Worten: Hat man eine Markoff-Kette $(X, \mathbb{F})$ auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Startverteilung ν und Zustandsraum S, so reicht zu deren Beschreibung der kanonische Prozeß $(X^0, \mathbb{F}^0)$ mit P_ν aus, da die Verteilung von X unter P mit der von X^0 unter P_ν übereinstimmt (an dieser Verteilung ist man ja interessiert).

Dies zeigen der nächste Satz und das nachfolgende Korollar:

Satz 5.7: Sei $(X, \mathbb{F})$ eine Markoff-Kette, $(X^0, \mathbb{F}^0)$ der kanonische Prozeß. Dann gilt für alle $n \in \mathbb{N}_0$ und $B \in \bigotimes\limits_{j=0}^{\infty} \mathrm{Pot}(S)$: $\mathrm{P}_{X_n(\cdot)}[\,(X_0^0, X_1^0, \ldots) \in B\,] \equiv \mathrm{P}_{X_n(\cdot)}[\,B\,]$ ist eine Version von $\mathrm{P}[\,(X_n, X_{n+1}, \ldots) \in B \mid \mathcal{F}_n\,](\cdot)$, d.h. $\mathrm{P}_{X_n(\cdot)}[\,\cdot\,]$ ist bedingte Verteilung von $(X_n, X_{n+1}, \ldots)$ bzgl. $\mathcal{F}_n$.

Beweis: 1.) $\mathrm{P}_{X_n(\omega)}[\,\cdot\,]$ ist ein Wahrscheinlichkeitsmaß auf $\bigotimes\limits_{j=0}^{\infty} \mathrm{Pot}(S) \quad \forall \omega \in \Omega$

2.) Die Abbildung $\omega \mapsto \mathrm{P}_{X_n(\omega)}[\,B\,]$ ist $\mathcal{F}_n - \mathcal{B}$-meßbar für alle $B \in \bigotimes\limits_{j=0}^{\infty} \mathrm{Pot}(S)$.

3.) Sei $B \in \bigotimes\limits_{j=0}^{\infty} \mathrm{Pot}(S)$ "endlich dimensional", d.h. $B = \{ j_1 \} \times \ldots \times \{ j_{m+1} \} \times S \times \ldots$ mit $m \in \mathbb{N}_0$. Dann ist z.z.: $\int\limits_A \mathrm{P}_{X_n}[\,B\,]\, d\mathrm{P} = \mathrm{P}[\,\{ (X_n, X_{n+1}, \ldots) \in B \} \cap A\,] \quad \forall A \in \mathcal{F}_n$.

$$
\begin{aligned}
\mathrm{P}[\,\{ (X_n, X_{n+1}, \ldots) \in B \} \cap A\,] &= \mathrm{P}[\,\{ (X_n, \ldots, X_{n+m}) = (j_1, \ldots, j_{m+1}) \} \cap A\,] \\
&= \mathrm{P}[\,\{ X_n = j_1 \} \cap \ldots \cap \{ X_{n+m} = j_{n+m} \} \cap A\,] \\
&\overset{(5.6)}{=} \mathrm{P}[\,A \cap \{ X_n = j_1 \}\,] \cdot p_{j_1 j_2} \cdot \ldots \cdot p_{j_m j_{m+1}} \\
&\equiv \mathrm{P}_{j_1}[\,X_0^0 = j_1\,] \cdot \mathrm{P}[\,A \cap \{ X_n = j_1 \}\,] \cdot p_{j_1 j_2} \cdot \ldots \cdot p_{j_m j_{m+1}} \\
&\overset{(5.6)}{=} \mathrm{P}_{j_1}[\,X_0^0 = j_1, \ldots, X_m^0 = j_{m+1}\,] \cdot \mathrm{P}[\,A \cap \{ X_n = j_1 \}\,] \\
&= \sum_{j \in S} \mathrm{P}_j[\,X_0^0 = j_1, \ldots, X_m^0 = j_{m+1}\,] \cdot \mathrm{P}[\,A \cap \{ X_n = j_1 \}\,] \\
&\equiv \sum_{j \in S} \mathrm{P}_j[\,B\,] \cdot \mathrm{P}[\,A \cap \{ X_n = j \}\,] \equiv \int\limits_A \mathrm{P}_{X_n}[\,B\,]\, d\mathrm{P}\,.
\end{aligned}
$$

4.) Da $\sigma(\{\{j_1,\ldots,j_{m+1}\}\times S\times\ldots : j_i\in S, m\in\mathbb{N}\})\equiv\sigma(\mathcal{C}_2)=\bigotimes_{i=0}^{\infty}\operatorname{Pot}(S)$, folgt die Behauptung mit (1.5′). Man setze dort $\Omega_1 := \Omega$, $\Omega_2 := \bigtimes_{i=0}^{\infty} S$, $\mathcal{F}_1 := \mathcal{F}_n$, $\mathcal{F}_2 := \bigotimes_{i=0}^{\infty}\operatorname{Pot}(S)$, $X := \mathrm{id}_\Omega$ sowie $Y := (X_n, X_{n+1},\ldots)$.

q.e.d.

Die Verteilungsgleichheit zeigt

Korollar 5.8: Sei $(X,\mathbb{F})$ wie in (5.7), ν die Startverteilung von X. Dann gilt für alle $B\in\bigotimes_{j=0}^{\infty}\operatorname{Pot}(S)$:

$$\mathrm{P}[(X_0,X_1,\ldots)\in B] = \mathrm{P}_\nu[(X_0^0,X_1^0,\ldots)\in B]\equiv \mathrm{P}_\nu[B]$$

Beweis: $\mathrm{P}[(X_0,X_1,\ldots)\in B]\equiv \mathrm{E}\big[\mathrm{P}[(X_0,X_1,\ldots)\in B\mid\mathcal{F}_0]\big]$

$$= \mathrm{E}\big[\mathrm{P}_{X_0}[(X_0^0,X_1^0,\ldots)\in B]\big] \quad \text{nach (5.7)}$$

$$= \sum_{i\in S}\mathrm{P}[X_0=i]\cdot \mathrm{P}_i[(X_0^0,X_1^0,\ldots)\in B]$$

$$\equiv \sum_{i\in S}\nu(\{i\})\cdot \mathrm{P}_i[B].$$

Diese Rechnung gilt gemäß der Bemerkung nach Definition 3 analog für X^0 anstelle von X, da $(X^0)^0 = X^0$. Mithin hat man $\mathrm{P}_\nu[(X_0^0,X_1^0,\ldots)\in B]=\sum_{i\in S}\nu(\{i\})\cdot\mathrm{P}_i[B]$.

q.e.d.

Beim letzten Beweis fiel als wichtiges Nebenresultat ab:

(5.8′) $\mathrm{P}_\nu[B]=\sum_{i\in S}\nu(\{i\})\cdot\mathrm{P}_i[B]$ für $B\in\bigotimes_{j=0}^{\infty}\operatorname{Pot}(S)$.

Eine weitere Folgerung aus (5.7) ist die Einführung von n-Schritt-Übergangswahrscheinlichkeiten:

Korollar 5.9: Für $n\in\mathbb{N}_0$, $m\in\mathbb{N}$ und $i,j\in S$ hat man die Gleichungskette

$$\mathrm{P}[X_{m+n}=j\mid\mathcal{F}_n]=p_{ij}^{(m)}=\mathrm{P}_i[X_m^0=j] \quad \text{auf } \{X_n=i\}$$

Beweis: Folgende Betrachtung geschehe auf $\{X_n=i\}$:

$$\mathrm{P}[X_{m+n}=j\mid\mathcal{F}_n]=\sum_{j_1,\ldots,j_{m-1}\in S}\mathrm{P}[X_{n+1}=j_1,\ldots,X_{m+n-1}=j_{m-1},X_{n+m}=j\mid\mathcal{F}_n]$$

$$=\sum_{j_1,\ldots,j_{m-1}\in S} p_{ij_1}\cdot\ldots\cdot p_{j_{m-1}j} \quad \text{nach (5.6)}$$

$$\equiv p_{ij}^{(m)}$$

Außerdem nach (5.7) mit $B := \bigtimes_{k=0}^{m-1} S\times\{j\}\times S\times\ldots$:

$$\mathrm{P}[X_{m+n}=j\mid\mathcal{F}_n]\equiv\mathrm{P}[(X_n,X_{n+1},\ldots)\in B\mid\mathcal{F}_n]$$

$$=\mathrm{P}_i[B] \quad \text{nach (5.7), da } X_n=i$$

$$\equiv\mathrm{P}_i[X_m^0=j]$$

q.e.d.

Bemerkung: Dieses Korollar legt für die aus (5.3) her schon als Elemente von $\mathbb{P}^m$ bekannten $p_{ij}^{(m)}$ die Bezeichnung *m-Schritt-Übergangswahrscheinlichkeiten* nahe. $p_{ij}^{(m)}$ ist die Wahrscheinlichkeit, im Zeitpunkt $m+n$ den Zustand j erreicht zu haben, wenn man im Zeitpunkt $n \in \mathbb{N}_0$ im Punkt i gestartet ist. Die schon bekannte Beziehung $\mathbb{P}^{m+n} = \mathbb{P}^m \cdot \mathbb{P}^n$, d.h. $p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} \cdot p_{kj}^{(n)}$, wird als *Chapman-Kolmogoroff-Gleichung* bezeichnet.

Soweit die Motivation der Einführung des kanonischen Prozesses und die daraus resultierenden Eigenschaften. Nun wird eine wichtige Klasse von Funktionen vorgestellt, die für den weiteren Verlauf dieser Ausführung (speziell beim optimalen Stoppen, vgl. Abschnitt 6) noch von Bedeutung sein wird:

Definition 4: Seien S und $\mathbb{P}$ wie oben. Dann heißt eine Funktion $\varphi : S \to \mathbb{R}$ *superharmonisch* (*harmonisch*) $\iff$

$$\textbf{(5.10)} \qquad \varphi(i) \underset{(=)}{\geq} \sum_{j \in S} p_{ij} \cdot \varphi(j) \quad \forall\, i \in S$$

Bemerkung: Die Existenz der Summe in (5.10) wird mit vorausgesetzt.
Interpretiert man $\varphi(i)$ als Auszahlung im Zustand i, so ist die sofortige Auszahlung vorteilhafter, als auf die Auszahlung in einem späteren Schritt zu warten.

Mit Hilfe der superharmonischen Funktionen erhält man eine Verbindung zwischen der Martingal- und der Markoff-Kettentheorie:

Proposition 5.11: Sei $(X, \mathbb{F})$ eine Markoff-Kette mit Übergangsmatrix $\mathbb{P}$, Zustandsraum S und φ eine superharmonische Funktion auf S. Sind ferner $\varphi(X_0)$ und $\varphi^-(X_n)$, $n \in \mathbb{N}$, P-integrierbar, so ist $\varphi(X) = \big(\varphi(X_n)\big)_{n \in \mathbb{N}_0}$ ein Supermartingal bzgl. $\mathbb{F}$.

Beweis: $\varphi(X)$ ist adaptiert an $\mathbb{F}$, da φ automatisch meßbar. Sei nun $A \in \mathcal{F}_n$.

$$\begin{aligned}
\Longrightarrow -\infty < -\mathrm{E}[\,\varphi^-(X_{n+1})\,] &\leq \int_A \varphi(X_{n+1})\, d\mathrm{P} \quad \text{da } \varphi = \varphi^+ - \varphi^- \text{ und } -\varphi^- \leq 0 \\
&= \sum_{j \in S} \varphi(j) \cdot \mathrm{P}[\,A \cap \{\, X_{n+1} = j \,\}\,] \\
&= \sum_{j \in S} \varphi(j) \cdot \sum_{i \in S} \mathrm{P}[\,A \cap \{\, X_n = i \,\}\,] \cdot p_{ij} \quad \text{nach (5.5)} \\
&\overset{(*)}{=} \sum_{i \in S} \mathrm{P}[\,A \cap \{\, X_n = i \,\}\,] \cdot \sum_{j \in S} p_{ij} \cdot \varphi(j) \\
&\leq \int_A \varphi(X_n)\, d\mathrm{P} \quad \text{da } \varphi \text{ superharmonisch} \\
&\leq \ldots \leq \int_A \varphi(X_0)\, d\mathrm{P} < \infty \quad \text{nach Voraussetzung.}
\end{aligned}$$

Dabei gilt $(*)$ nach Zerlegung von φ in φ^+, φ^- zunächst für φ^-, φ^+. Damit aber auch für φ.

q.e.d.

Es folgen nun noch einige Beispiele aus der Praxis, die das Auftreten von Markoff-Ketten in demographischen Untersuchungen, EDV und anderen Gebieten aufzeigen. Zuvor noch

Definition 5: Sei S ein abzählbarer Raum, $\mathbb{P}$ eine Übergangsmatrix. Dann heißt ein Zustand $i \in S$ *absorbierend*, falls $p_{ii} = 1$.

Beispiel 1: Verzweigunsprozeß (Galton-Watson-Prozeß)

Historisch geht diese Art von Prozessen auf Untersuchungen Galtons zurück, wann ein Familienname ausstirbt, d.h. es keinen männlichen Nachfolger mehr gibt. Gesucht ist also ein mathematisches Modell, in dem ein stochastischer Prozeß $X = (X_n)_{n\in\mathbb{N}_0}$ jeweils die Anzahl der (z.B. männlichen) Mitglieder der n-ten Population beschreibt. Dabei sei die Vorgabe zu berücksichtigen, daß jedes Individuum unabhängig von den anderen eine zufällige Anzahl von Nachkommen gemäß der Zähldichte $b = (b_j)_{j\in\mathbb{N}_0}$ von $\mathbb{N}_0$ produziert.

Zur Modellbildung benutze man die folgende Terminologie:

i.) $b^{\star i}$, $i \in \mathbb{N}_o$, sei die i-fache Faltung von b mit sich selbst $\big($ vgl. W-Theorie 13.10 c.) $\big)$.
Setze $b^{\star 0}(j) := \delta_{0j}$, d.h. $b^{\star 0}$ ist die auf 0 konzentrierte Verteilung.

ii.) $g : [0,1] \to [0,1]$ mit $g(s) := \sum\limits_{j=0}^{\infty} b_j\, s^j$ sei die erzeugende Funktion von b. Dann ist g^i die erzeugende Funktion von $b^{\star i}$ $\big($ vgl. W-Theorie 15.6 c.) $\big)$

iii.) μ sei die mittlere Anzahl der Nachkommen eines Individuums, d.h. $\mu = \sum\limits_{j\in\mathbb{N}_o} j \cdot b_j$. Somit ist $\sum\limits_{j\in\mathbb{N}_o} j \cdot b_j^{\star i} = i \cdot \mu$ die mittlere Anzahl der Nachkommen von i Individuen, da die Summe von i gemäß b verteilter unabhängiger Zufallsvariablen die Verteilung $b^{\star i}$ besitzt.

iv.) Gemäß iii.) ist $\sigma^2 = \sum\limits_{j\in\mathbb{N}_o} (j-\mu)^2\, b_j$ die Varianz der Nachkommenszahl eines Individuums, wenn $\mu < \infty$ ist.

Man hat die folgenden zwei Modelle:

1. Modellannahme: Sei X eine Markoff-Kette bzgl. der kanonischen Filterung mit Zustandsraum $S = \mathbb{N}_o$, Übergangsmatrix $(p_{ij})_{i,j\in\mathbb{N}_o} = (b_j^{\star i})_{i,j\in\mathbb{N}_o}$ und der auf 1 konzentrierten Startverteilung ν bzgl. P (z.B. X der kanonische Prozeß und $\mathrm{P} := \mathrm{P}_\nu \equiv \mathrm{P}_1$ das nach Ionescu-Tulcea existierende Wahrscheinlichkeitsmaß).

2. Modellannahme: Sei $(Z_i)_{i\in\mathbb{N}_o}$ eine Folge unabhängiger, identisch verteilter Zufallsvariablen mit $Z_m \sim b$ $\Big($ z.B. Projektionen auf $\big(\mathop{\times}\limits_{i=0}^{\infty} \mathbb{N}_o, \mathrm{Pot}(\mathop{\times}\limits_{i=o}^{\infty} \mathbb{N}_o), \mathop{\times}\limits_{i=o}^{\infty} b \equiv \mathrm{P}_\infty \big) \Big)$.
Definiere $X_0 := 1$, $X_1 := Z_1$, $X_2 := Z_2 + \ldots + Z_{1+Z_1}$
und allgemein $\quad X_{n+1} := \sum\limits_{m=1}^{X_n} Z_{X_0+\ldots+X_{n-1}+m}$
Man ordnet also hier jedem Individuum der n-ten Generation dessen Nachkommen zu und addiert diese auf zu X_{n+1}.
Beh.: X ist unter P_∞ eine Markoff-Kette bzgl. der kanonischen Filterung mit Zustandsraum $S = \mathbb{N}_o$, Übergangsmatrix $(b_j^{\star i})_{i,j\in\mathbb{N}_o}$ und Startverteilung ν, welche auf 1 konzentriert ist.
Beweis: Der Beweis ist eine relativ leichte Übung.

□

Bemerkung: Dieses zweite Modell ist ein Warteschlangenmodell (vgl. Beispiel 2 unten).
Dabei faßt man die Nachkommen eines Kunden als die während seiner Bedienung eintreffenden Kunden auf. Der Begriff "Aussterben" bedeutet dann, daß die Warteschlange einmal abgebaut wird. Dort ist der Zustand 0 jedoch nicht absorbierend, da irgendwann neue Kunden eintreffen.

Für die nun folgenden Überlegungen verwende man Modell 1, denn aus Modell 2 folgt ja Modell 1.

Man interessiert sich für die Wahrscheinlichkeit, daß eine Familie ausstirbt. Definiere dafür die Kennzahl $\alpha := \mathrm{P}[\bigcup_{n\in\mathbb{N}_0} \{X_n = 0\}]$, die sogenannte <u>*Aussterbewahrscheinlichkeit*</u>.

Da der Zustand 0 absorbierend ist, gilt für alle $m, n \in \mathbb{N}_0$ mit $m \leq n$:

$\mathrm{P}[X_m = 0, X_n \neq 0] = \sum_{j\in\mathbb{N}} \mathrm{P}[X_m = 0, X_n = j] = \sum_{j\in\mathbb{N}} \mathrm{P}[X_m = 0] \cdot p_{0j}^{(n-m)} = 0$ nach (5.9),

da $\mathrm{P}[X_n = j \mid \mathcal{F}_m^X] = p_{0j}^{(n-m)}$ auf $\{X_m = 0\}$.

Somit erhält man
$$\begin{aligned}\alpha \equiv \mathrm{P}[\bigcup_{n\in\mathbb{N}_0} \{X_n = 0\}] &= \lim_{n\to\infty} \mathrm{P}[\bigcup_{m=1}^{n} \{X_m = 0\}] \\ &= \lim_{n\to\infty} \mathrm{P}[\bigcup_{m=1}^{n} \{X_m = 0, X_n = 0\}] \\ &= \lim_{n\to\infty} \mathrm{P}[X_n = 0] \equiv \lim_{n\to\infty} p_{10}^{(n)}\end{aligned}$$

Es soll nun α mit Hilfe einer Fixpunktgleichung gewonnen werden:

Sei dazu für $0 \leq s \leq 1$ $g_n(s) := \sum_{j\in\mathbb{N}_0} p_{1j}^{(n)} s^j \equiv \sum_{j\in\mathbb{N}_0} \mathrm{P}[X_n = j] \cdot s^j$ die erzeugende Funktion von X_n (d.h. von $\mathrm{P}X_n^{-1}$). Aus der Chapman-Kolmogoroff-Gleichung $p_{1j}^{(n+1)} = \sum_{i\in\mathbb{N}_0} p_{1i}^{(n)} \cdot p_{ij}$ folgt:

$g_{n+1}(s) = \sum_{i\in\mathbb{N}_0} p_{1i}^{(n)} \cdot \sum_{j\in\mathbb{N}_0} b_j^{\star i} \cdot s^j \equiv \sum_{i\in\mathbb{N}_0} p_{1j}^{(n)} g^i(s) \equiv g_n\big(g(s)\big)$, d.h. man hat eine Rekursion mit dem Startwert $g_1(s) \equiv \sum_{j\in\mathbb{N}_0} p_{1j} \cdot s^j \equiv \sum_{j\in\mathbb{N}_0} b_j \cdot s^j \equiv g(s)$. Dies zusammengefaßt ergibt:

<u>(5.12)</u> $g_{n+1} = g_n \circ g = \underbrace{g \circ \ldots \circ g}_{(n+1)\text{-mal}} = g \circ g_n$

Da nun $p_{10}^{(n+1)} \equiv g_{n+1}(0) \equiv g\big(g_n(0)\big) = g(p_{10}^{(n)})$ folgt aufgrund der Stetigkeit von g (g ist eine konvergente Potenzreihe): $\alpha \equiv \lim_{n\to\infty} p_{10}^{(n)} = \lim_{n\to\infty} p_{10}^{(n+1)} = \lim_{n\to\infty} g(p_{10}^{(n)}) = g(\alpha)$, d.h. α ist ein Fixpunkt von g.
Da man mit der Iteration im kleinsten Wert von $[0,1]$ beginnt $\big(p_{10} = g(0)\big)$ und g isoton auf $[0,1]$ ist, gilt:

<u>(5.13)</u> α ist die kleinste Lösung von $s = g(s)$ in $[0,1]$

Dabei ist $s = 1$ immer eine Lösung von $g(s) = s$ (diese Lösung ist nicht interessant) und $s = 0$ ist eine Lösung genau dann, wenn $b_0 = 0$ (d.h. es werden fast sicher Nachkommen produziert).

Die sonstige Existenz und Eindeutigkeit einer Lösung von $g(s) = s$ in $(0,1)$ behandelt

<u>Lemma 5.14:</u> Sei $0 < b_0 < 1$. Dann existiert eine Lösung von $g(s) = s$ in $(0,1)$ $\iff$ $\mu > 1$.
Diese Lösung ist dann eindeutig.

<u>Beweis:</u> Man sucht eine Nullstelle der Funktion $h(s) := g(s) - s$ im Intervall $(0,1)$. Es gelten:

$$h(0) = b_0 > 0\,,\; h(1) = 0\,,\; h'(1-0) = \mu - 1 \text{ sowie } h''(s) = g''(s) = \sum_{k=2}^{\infty} k \cdot (k-1) \cdot b_k \cdot s^{k-2}$$

Damit lassen sich folgende Fälle unterscheiden:

<u>Fall 1:</u> $b_0 + b_1 = 1 \Longrightarrow h(s) = b_0 + (b_1 - 1) \cdot s$, d.h. h ist linear und $\mu = b_1 < 1$.
In diesem Fall existiert keine Nullstelle.

<u>Fall 2:</u> $b_0 + b_1 < 1 \implies h''$ ist strikt positiv auf $(0,1)$
$\implies h$ ist strikt konvex über $(0,1)$

<u>Fall 2a:</u> $\mu - 1 \leq 0 \implies h' < 0$ auf $(0,1)$, da h' strikt isoton auf $(0,1)$
$\implies$ Es existiert keine Nullstelle in $(0,1)$, da $h(1) = 0$

<u>Fall 2b:</u> $\mu - 1 > 0 \implies h < 0$ in einer linksseitigen Umgebung von 1
$\implies$ Es existiert eine Nullstelle von h in $(0,1)$ nach dem Zwischenwertsatz, da $h(0) > 0$ und h stetig. Diese ist wegen der strikten Konvexität von h eindeutig.

q.e.d.

Sei $M := \sum_{n=0}^{\infty} X_n$ die Gesamtzahl aller Individuen, d.h. mit M gilt $\alpha = \mathrm{P}[\, M < \infty\,]$. Es ergibt sich mit Hilfe der Markoff-Eigenschaft:

$$\begin{aligned}\mathrm{E}[\, X_{n+1} \mid X_0, \dots, X_n\,] &= \sum_{j \in \mathbb{N}_0} j \cdot \mathrm{P}[\, X_{n+1} = j \mid X_0, \dots, X_n\,] \\ &= \sum_{j \in \mathbb{N}_0} j \cdot p_{ij} \quad \text{auf } \{\, X_n = i\,\} \text{ nach (5.4)} \\ &\equiv \sum_{j \in \mathbb{N}_0} j \cdot b_j^{\star i} \equiv i \cdot \mu \quad \text{auf } \{\, X_n = i\,\}\end{aligned}$$

Mit anderen Worten:

(<u>5.15</u>) $\mathrm{E}[\, X_{n+1} \mid X_0, \dots, X_n\,] = \mu \cdot X_n$

Insbesondere: $\mathrm{E}[\, X_{n+1}\,] = \mu \cdot \mathrm{E}[\, X_n\,]$. Da $\mathrm{E}[\, X_0\,] = 1$ folgt:

(<u>5.16</u>) $\mathrm{E}[\, X_n\,] = \mu^n$, d.h. $\mathrm{E}[\, M\,] = \sum_{n=0}^{\infty} \mu^n$.

Aufgrund von (5.15) hat man die im Bezug auf das Konvergenzverhalten wichtige Aussage:

(<u>5.17</u>) Im Fall $0 < \mu < \infty$ ist $\dfrac{X_n}{\mu^n}$ ein Martingal.

Mit diesen Resultaten ist es nun möglich, in den verschiedensten Fällen die Aussterbewahrscheinlichkeit sowie die erwartete Gesamtzahl aller Individuen zu berechnen:

<u>Satz 5.18:</u> a.) Ist $\mu < 1$ ($\implies b_0 > 0$), so ist $\alpha = 1$ und $\mathrm{E}[\, M\,] = \dfrac{1}{1-\mu} < \infty$.

b.) Ist $\mu = 1$ und $b_0 > 0$, so gilt: $\alpha = 1$, $\mathrm{E}[\, X_n\,] = 1$, $\mathrm{E}[\, M\,] = \infty$ sowie $\mathrm{Var}\,(X_n) = n \cdot \sigma^2$.

c.) Ist $\mu > 1$ und $b_0 > 0$, so ist $\alpha < 1$ und $\mathrm{E}[\, M\,] = \infty$. Genauer: α ist die einzige Lösung von $g(s) = s$ im Intervall $(0,1)$. Ist zusätzlich $1 < \mu < \infty$, so existiert eine integrierbare Zufallsvariable Y mit $Y = \lim_{n \to \infty} \dfrac{X_n}{\mu^n}$.

d.) Ist $b_0 = 0$, so ist $\alpha = 0$ und $M = \infty$ fast sicher.

<u>Beweis:</u> a.) $\alpha = 1$ folgt aus (5.14), $\mathrm{E}[\, M\,] = \dfrac{1}{1-\mu}$ aus $\mathrm{E}[\, M\,] = \sum_{n=0}^{\infty} \mu^n$ mit geometrischer Reihenformel.

b.) $\alpha = 1$ gilt nach Fall 2a, $\mathrm{E}[\,X_n\,] = 1$ nach (5.16) und damit $\mathrm{E}[\,M\,] = \sum_{n=0}^{\infty} 1 = \infty$

(aber $M < \infty$ fast sicher !). Ist g_n erzeugende Funktion von X_n, so gilt für $k \in \mathbb{N}$:

$g_n^{(k)}(s) = \sum_{n=k}^{\infty} (n)_k \cdot \mathrm{P}[\,X_n = m\,] \cdot s^{n-k}$, $0 < s < 1$, wo $(n)_k := n \cdot (n-1) \cdot \ldots \cdot (n-k+1)$.

Mit anderen Worten: $g_n^{(k)}(1-0) = \mathrm{E}[\,(X_n)_k\,]$

$$\Longrightarrow \mathrm{Var}\,(X_{n+1}) \equiv \mathrm{E}[\,X_n^2\,] - \mathrm{E}[\,X_n\,]^2 = \mathrm{E}[\,X_n(X_n-1)\,] + \mathrm{E}[\,X_n\,] - \mathrm{E}[\,X_n\,]^2$$
$$= g_n''(1-) + g_n'(1-) - \big(g_n'(1-)\big)^2$$

Hier: $\mathrm{Var}\,(X_{n+1}) = \sigma^2 \cdot \mathrm{E}[\,X_n\,] + \mu^2 \cdot \mathrm{Var}\,(X_n)$

Daraus folgt bei $\mu = 1$: $\mathrm{Var}\,(X_n) = n \cdot \sigma^2$.

c.) $\alpha > 0$ folgt aus Fall 2b von (5.14), $\mathrm{E}[\,M\,] = \infty$ ist klar. Nach (5.17) ist $(\frac{X_n}{\mu^n})_{n \in \mathbb{N}_0}$ ein Martingal. Da $\mathrm{E}[\,(\frac{X_n}{\mu^n})^+\,] = \mathrm{E}[\,\frac{X_n}{\mu^n}\,] = 1$ ergibt sich die Behauptung mit (4.9) c.).

d.) klar.

□

Bemerkung 5.19: Sind in der nullten Generation i Individuen vorhanden, so ist die Aussterbewahrscheinlichkeit α^i, d.h. $\lim_{n\to\infty} p_{i0}^{(n)} = \alpha^i$ (Übung)

Nachdem oben schon Teile eines Warteschlangenmodells vorgestellt wurden, nun eine Anwendung von obigem auf das

Beispiel 2: Ein Warteschlangenmodell

In einem Geschäft treffen nacheinander Kunden ein und beanspruchen Bedienung. Es kann jedoch stets nur ein Kunde bedient werden. Z_n beschreibe die Anzahl der Ankünfte von Kunden während der n-ten Bedienung, X_0 die Zahl der zu Beginn wartenden Kunden sowie X_n die Zahl der wartenden Kunden nach der n-ten Bedienung. Ist $X_n = 0$, so beginnt (fast sicher) eine neue Bedienung mit der Ankunft des nächsten Kunden. Somit hat man für X_{n+1}, $n \in \mathbb{N}_0$, und $i \in \mathbb{N}_0$ die Gleichung

(5.20) $X_0^{(i)} = i\,,\ X_{n+1}^{(i)} = (X_n^{(i)} - 1)^+ + Z_{n+1}$

Modellannahme: Sei $X_0^{(i)} = i$ konstant, $(Z_n)_{n\in\mathbb{N}}$ eine Folge identisch verteilter unabhängiger Zufallsvariablen auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit der Zähldichte $b = (b_j)_{j\in\mathbb{N}_0}$. Ferner sei $b_0 + b_1 < 1$ und $b_0 > 0$.

Nach (5.20) und (2.6) a.) ist $(X_n^{(i)})_{n\in\mathbb{N}_0}$ eine Markoff-Kette mit Übergangswahrscheinlichkeiten p_{lj} gegeben durch $p_{lj} = \mathrm{P}[\,(l-1)^+ + Z_1 = j\,] = b_{j-(l-1)^+}$. Setze dabei $b_j = 0$ für $j < 0$. Definiere ferner

(5.21) $\varrho := \mathrm{E}[\,Z_n\,] \equiv \sum_{j\in\mathbb{N}_0} j \cdot b_j$

als die Verkehrsrate (= mittlere Anzahl der Ankünfte pro Bedienung).
Man betrachtet nun den schon aus Beispiel 1 her bekannten eingebetteten Verzweigungsprozeß $(X_n')_{n\in\mathbb{N}_0}$. Dabei sind die Nachkomen eines Kunden die während seiner Bedienung eintreffenden Kunden, die Aussterbewahrscheinlichkeit die Abbauwahrscheinlichkeit der Warteschlange.

Definiere $X_0' := i \equiv 1$, $X_1' := Z_1$, $X_2' := Z_2 + \ldots + Z_{1+Z_1}$

und allgemein ($n \in \mathbb{N}$): $\qquad X_{n+1}' := \sum_{m=1}^{X_n'} Z_{X_0'+\ldots+X_{n-1}'+m}$ (vgl. oben)

Dann gilt für $T^{(1)} := \inf\{\, n \geq 1 : X_n^{(1)} = 0 \,\}$: $T^{(1)}$ = Anzahl der Bedienungen bis zum erstmaligen Abbau der Warteschlange = Anzahl der bis dahin eingetroffenen Kunden = Gesamtzahl M der Individuen im Verzweigungsprozeß (Übung) = $\sum\limits_{n \in \mathbb{N}_0} X'_n$.

Für den Warteschlangenprozeß hat man die Größen $f^*_{i0} := \mathrm{P}[\, T^{(i)} < \infty \,]$, $m_{i0} := \mathrm{E}[\, T^{(i)} \,]$ (diese Größen werden in Abschnitt 7 näher spezifiziert) und ist besonders an f^*_{00} (= Wahrscheinlichkeit für eine Rückkehr nach 0) und m_{00} (= mittlere Rückkehrzeit nach 0) interessiert.

Hier ist $f^*_{00} = f^*_{10}$ und $m_{00} = m_{10}$, da $p_{0j} = p_{1j} \quad \forall\, j \in \mathbb{N}_0$ und damit nach (5.6)
$\mathrm{P}(X_1^{(0)}, \ldots, X_n^{(0)})^{-1} = \mathrm{P}(X_1^{(1)}, \ldots, X_n^{(1)})^{-1} \quad \forall\, n \in \mathbb{N}$.

Aufgrund der Beziehung $T^{(1)} = M$ liefert (5.18) Aussagen über m_{00} und f^*_{00}: (hier ist $\mu = \varrho$)

Satz 5.22: Für die Markoff-Kette des Warteschlangenmodells gilt $f^*_{00} = 1$ und $m_{00} = \dfrac{1}{1-\varrho}$, falls $\varrho < 1$. Ist $\varrho = 1$, so ist $f^*_{00} = 1$ und $m_{00} = \infty$. Im Fall $\varrho > 1$ gilt $f^*_{00} = \alpha < 1$ sowie $m_{00} = \infty$, wobei α die einzige Lösung der Gleichung $s = \sum\limits_{j \in \mathbb{N}_0} b_j \cdot s^j$ in $(\,0, 1\,)$ ist.

Bemerkung 5.23: (analog zu oben) $f^*_{i0} = \alpha^i$ für $i \in \mathbb{N}$ wobei $\alpha = 1$ für $\varrho \leq 1$.

Bemerkung: Als Beispiel für obiges $b = (b_j)_{j \in \mathbb{N}_0}$ nehme man die Wahrscheinlichkeiten aus dem Beispiel in Abschnitt 1. Mit diesem spziellen b hat man ein sogenanntes M/G/1-Modell.

Notation: M $\hat{=}$ Poisson-Prozeß als Inputprozeß
G $\hat{=}$ beliebige Bedienungsverteilung
1 $\hat{=}$ Anzahl der Bediener.

Für das letzte Beispiel dieses Abschnitts benötigt man noch

Definition 6: Sei X' ein stochastischer Prozeß auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$, $\mathbb{F}$ eine Filterung in $\mathcal{F}$. Dann heißt $(X', \mathbb{F})$ eine *inhomogene Markoff-Kette*, falls X' an $\mathbb{F}$ adaptiert ist und Übergangsmatrizen $\mathbb{P}(n)$, $n \in \mathbb{N}$, existieren mit der Eigenschaft

$$\textbf{(5.24)} \quad \mathrm{P}[\, X'_{n+1} = j \mid \mathcal{F}_n \,] = p_{ij}(n+1) \ \text{ auf } \ \{\, X'_n = i \,\} \quad \forall\, i, j \in S\,,\ n \in \mathbb{N}_0$$

Bei der inhomogenen Markoff-Kette sind die Übergangswahrscheinlichkeiten zeitlichen Veränderungen unterworfen, d.h. die Umwelt- (Rahmen-) bedingungen des Prozesses ändern sich im Laufe der Zeit.
Mit einem Trick kann man jedoch aus einer inhomogenen eine homogene Markoff-Kette entwickeln, dies zeigt

Proposition 5.25: Ist $(X', \mathbb{F})$ eine inhomogene Markoff-Kette mit Übergangsmatrizen $\mathbb{P}(n)$, $n \in \mathbb{N}$, so gilt für den Prozeß $X = (X_n)_{n \in \mathbb{N}}$ mit $X_n := (X'_n, n)$:
$(X, \mathbb{F})$ ist eine (homogene) Markoff-Kette mit Übergangswahrscheinlichkeiten

$$p_{(i,n)\,(j,n+1)} = \begin{cases} p_{ij}(n+1) & \text{für } m = n\,,\ i, j \in S \\ 0 & \text{sonst} \end{cases}$$

Beweis: $\mathrm{P}[\, X_{n+1} = (j, m) \mid \mathcal{F}_n \,] = \begin{cases} \mathrm{P}[\, X'_{n+1} = j \mid \mathcal{F}_n \,] & \text{für } m = n+1 \\ \underline{0} & \text{sonst} \end{cases}$

□

Bemerkung: Obiger Trick hilft bei der Lösung diverser Probleme. Jedoch können keine Aussagen über das asymptotische Verhalten eines derartigen Prozesses gemacht werden, da ein System, welches die Zeit als Zustandskomponente enthält, nie wieder in den gleichen Zustand zurückkehrt.

Nun das angekündigte

<u>Beispiel 3:</u> Das Sekretärinnenproblem

Es werden $s \in \mathbb{N}$ Objekte (Bewerberinnen) in zufälliger Reihenfolge nacheinander vorgestellt. Bei der Vorstellung eines Objektes muß sofort entschieden werden, ob es die Stelle erhält oder nicht. Ein zunächst abgewiesenes Objekt steht später nicht mehr zur Verfügung. Jedem Objekt kann prinzipiell eine Rangzahl von 1 bis s zugeordnet werden ($1 \mathrel{\hat{=}}$ bestes Objekt, $s \mathrel{\hat{=}}$ schlechtestes Objekt). Die Information bei Inspektion eines Objektes ist sein relativer Rang bzgl. seinen Vorgängern. Ω sei Σ_s (= symmetrische Gruppe in s Elementen), d.h. die Menge aller Permutationen der Zahlen aus $\{1,\dots,s\}$. Ω wird mit der Gleichverteilung P versehen, $|\,\Omega\,| = s\,!$.

$Z_n : \Omega \to \{1,\dots,s\}$, $1 \le n \le s$, definiert durch $Z_n(\omega) = \omega(n)$ sei der absolute Rang des n-ten Objektes, d.h. Z_n ist die n-te Projektion. Die Information bei der Vorstellung des n-ten Objektes ist die relative Rangfolge $R^n = (R^n_1,\dots,R^n_n) \in \Sigma_n$, d.h. R^n_m ist der Rang von Z_m in $(Z_1,\dots,Z_n)$, $1 \le m \le n$.

<u>Beispiel:</u> $s = 5\,,\ \omega \equiv \big(\omega(1),\dots,\omega(5)\big) = (4,2,3,1,5)$
$\Longrightarrow R^3(\omega) = (3,1,2)\,,\ R^4(\omega) = (4,2,3,1)$

Definiere $\mathcal{F}_n := \sigma(R^n)$. $\mathcal{F} = (\mathcal{F}_n)_{1\le n\le s}$ ist eine Filterung, da R^n eine meßbare Funktion von R^{n+1} ist.

Es gilt die Gleichung $\mathrm{P}[\,R^n = r,\ Z_{n+1} = z_{n+1},\dots,Z_s = z_s\,] = \dfrac{1}{s\,!}\quad \forall\, r \in \Sigma_n,\ 1 \le z_m \le s,\ z_l \ne z_k,\ k \ne l,$
da dieses Ereignis für genau ein $\omega \in \Omega$ eintritt. Daraus resultiert:

(<u>5.26</u>) $\mathrm{P}[\,R^n = r\,] = \dfrac{1}{n\,!}\quad \forall\, r \in \Sigma_n$ <u>Beweis:</u> $\dfrac{1}{n\,!} = \dfrac{1}{s\,!}\cdot s\cdot(s-1)\cdot\ldots\cdot(s-n-1)$

Definiere $X'_n := 1_{\{R^n_n = 1\}}$, d.h. X'_n zeigt an, ob das n-te Objekt das bis dahin beste ist. X'_n ist $\mathcal{F}_n$-meßbar, da R^n_n eine Projektion von R^n ist. Es ergibt sich:

(<u>5.27</u>) R^n und X'_{n+1} sind unabhängig sowie $\mathrm{P}[\,X'_n = 1\,] = \dfrac{1}{n}\,,\ 1 \le n \le s.$

<u>Beweis:</u>
$$\begin{aligned}\mathrm{P}[\,R^n = (r_1,\dots,r_n)\,,\ X'_{n+1} = 1\,] &\equiv \mathrm{P}[\,R^{n+1} = (r_1+1,\dots,r_n+1,1)\,]\\ &= \frac{1}{(n+1)\,!}\quad \text{nach (5.26)}\\ &= \frac{1}{n+1}\cdot \mathrm{P}[\,R^n = (r_1,\dots,r_n)\,]\quad \text{ebenfalls nach (5.26)}\end{aligned}$$

Durch Summation über $(r_1,\dots,r_n) \in \Sigma_n$ ($|\,\Sigma_n\,| = n\,!$) erhält man: $\mathrm{P}[\,X'_{n+1} = 1\,] = \dfrac{1}{n+1}$.

Es bleibt noch der Fall $X'_{n+1} = 0$ zu untersuchen:

$$\begin{aligned}\mathrm{P}[\,R^n = (r_1,\dots,r_n)\,,\ X'_{n+1} = 0\,] &= \mathrm{P}[\,R^n = (r_1,\dots,r_n)\,] - \mathrm{P}[\,R^n = (r_1,\dots,r_n)\,,\ X'_{n+1} = 1\,]\\ &= \mathrm{P}[\,R^n = (r_1,\dots,r_n)\,]\cdot(1 - \mathrm{P}[\,X'_{n+1} = 1\,])\end{aligned}$$

q.e.d.

Aus (5.27) folgt unmittelbar

(<u>5.28</u>) $X'_1,\dots,X'_s$ sind unabhängig (da $(X'_1,\dots,X'_n)$ eine Funktion von R^n ist $\forall\, n \in \{1,\dots,s\}$)

Mit den Setzungen $X'_n := 0\,,\ n \notin \{1,\dots,s\}$ und $\mathcal{F}_0 := \{\emptyset,\Omega\}\,,\ \mathcal{F}_n := \mathrm{Pot}(\Omega)$ für $n > s$ sowie $X_n := \big(X'_n, n \wedge (s+1)\big)$ erhält man die Aussagen

(5.29) $(X', \mathbb{F})$ ist eine inhomogene und $(X, \mathbb{F})$ eine homogene Markoff-Kette mit den Übergangswahrscheinlichkeiten ($i = 0\,,\,1$)

$p_{(i,m)\,(1,m+1)} = \dfrac{1}{n+1}\,,\ 0 \leq m \leq s$ (die Wahrscheinlichkeit, daß das neuvorgestellte Objekt das bisher beste ist)

$p_{(i,m)\,(0,m+1)} = \dfrac{m}{m+1}\,,\ 0 \leq m < s$ (die Komplementwahrscheinlichkeit für das Komplementereignis)

sowie $p_{(i,k)\,(0,s+1)} = 1\,,\ k = s, s+1\,,\ i = 0,1$ (klar nach obigen Setzungen, der Zustand $(0, s+1)$ ist absorbierend)

Als Zustandsraum von $(X, \mathbb{F})$ hat man daher $S = \{\,(0,0)\,\} \cup (\{\,0,1\,\} \times \{\,1,\ldots,s\,\}) \cup \{\,(0,s+1)\,\}$.

Beweis: Sei o.E. $n < s$ (der Fall $n \geq s$ ist klar)

Sei $A \in \mathcal{F}_n \equiv \sigma(R^n)$, d.h. $A = \{\,R^n \in B^n\,\}$ mit $B^n \subset \Sigma_n$.

$$\Longrightarrow \mathrm{P}[\,A \cap \{\,X'_{n+1} = j\,\}\,] = \mathrm{P}[\,A\,] \cdot \mathrm{P}[\,X'_{n+1} = j\,] \quad \text{wegen (5.27)}$$

$$\equiv \int_A \mathrm{P}[\,X'_{n+1} = j\,]\,d\mathrm{P}$$

Setze deshalb $p_{X'_n(\,\cdot\,),j}(n+1) := \mathrm{P}[\,X'_{n+1} = j\,]$.

Damit wird $(X', \mathbb{F})$ zur inhomogenen Markoff-Kette. Die zweite Behauptung folgt mit (5.25).

q.e.d.

Nachdem in Abschnitt 4 schon der Begriff der Stoppzeit vorgestellt und erläutert wurde, werden im folgenden Abschnitt Kriterien für das Auffinden von optimalen Stoppzeiten gegeben und an einigen Beispielen (z.B. dem Sekretärinnenproblem) erläutert.

Abschnitt 6: Optimales Stoppen

Sei $(X, \mathbb{F})$ eine Markoff-Kette auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Übergangsmatrix $\mathbb{P}$, Startverteilung ν und Zustandsraum S.
Eine beschränkte *Nutzenfunktion* $u : S \rightarrow \mathbb{R}$ bewerte die Zustände aus S.
Man sucht einen Zeitpunkt τ, in dem man sich aufgrund der bis dahin gemachten Erfahrungen/Beobachtungen für den vorliegenden Zustand X_τ und den damit verbundenen Nutzen $u(X_\tau)$ entscheidet. Dabei soll der erwartete Nutzen maximiert werden. Der Entscheidungszeitpunkt τ wird (analog wie in Abschnitt 4) durch eine Stoppzeit beschrieben (hier wieder mit τ notiert), d.h. man interessiert sich für den gestoppten Prozeß X_τ. Mache dazu folgende

Definition 1: $\tau \in \mathcal{T} :\Longleftrightarrow \tau$ ist eine Stoppzeit bzgl. $\mathbb{F}$ mit der Eigenschaft $\mathrm{P}[\,\tau < \infty\,] = 1$.
Setzt man noch $X_\infty := e \notin S$, so ist X_τ auch für $\tau = \infty$ definiert und meßbar.

Gemäß obiger Motivation sucht man ein $\tau_* \in \mathcal{T}$ derart, daß $\mathrm{E}[\,u(X_{\tau_*})\,] = \sup\limits_{\tau \in \mathcal{T}} \mathrm{E}[\,u(X_\tau)\,]$.
Zur Lösung dieses Problems geht man zur internen Markoff-Familie $(X^0, \mathbb{F}^0, \{\,\mathrm{P_i}\,,\ \mathrm{i} \in \mathrm{S}\,\})$ über und definiert:

Definition 2: $\tau \in \mathcal{T}_i :\Longleftrightarrow \tau$ ist eine Stoppzeit bzgl. $\mathbb{F}^0$ mit der Eigenschaft $\mathrm{P}_i[\,\tau < \infty\,] = 1$.
$v : S \rightarrow \mathbb{R}$ heißt *Wertfunktion* $:\Longleftrightarrow v(i) = \sup\limits_{\tau \in \mathcal{T}_i} \mathrm{E}_i[\,u(X^0_\tau)\,]$

Eine Wertfunktion gibt bei einem Start in $i \in S$ den bei einem optimalen Stoppzeitpunkt erreichbaren mittleren Nutzen an.

Eine Charakterisierung von Stoppzeiten aus $\mathcal{T}_i$, $i \in S$, liefert

Lemma 6.1: a.) Sind τ^1, τ^2 Stoppzeiten, so sind auch $\tau^1 \wedge \tau^2$ und $\tau^1 \vee \tau^2$ Stoppzeiten.

b.) Mit τ sind auch $1+\tau$ sowie $1+\tau(X_1^0, X_2^0, \ldots)$ Stoppzeiten.

Bemerkung: $\tau(X_1^0, X_2^0, \ldots)$ ist i.a. keine Stoppzeit. Dies gilt z.B. für $\tau := \inf\{\, n \in \mathbb{N}_0 : X_n^0 \in B \,\}$ mit $B \in \mathrm{Pot}\,(S)$, denn es ist dann $\tau(X_1^0, X_2^0, \ldots) = \inf\{\, n \in \mathbb{N}_0 : X_{n+1}^0 \in B \,\}$.

c.) Sind $\tau_j \in \mathcal{T}_j \quad \forall\, j \in S \implies \tau := \tau_{X_0^0}(X_0^0, X_1^0, \ldots) \equiv \tau_{X_0^0} \in \mathcal{T}_j \quad \forall\, i \in S$

d.) Sind $\tau_j \in \mathcal{T}_j \quad \forall\, j \in S \implies \tau' := 1 + \tau_{X_1^0}(X_1^0, X_2^0, \ldots) \in \mathcal{T}_i \quad \forall\, i \in S$

und es gilt: $\mathrm{E}_i[\, u(X_{\tau'}^0)\,] = \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\, u(X_{\tau_j}^0)\,] \quad \forall\, i \in S.$

e.) Für $\tau \in \mathcal{T}_i$ mit $\{\, X_0^0 = i \,\} \subset \{\, \tau > 0 \,\}$ definiere $\tau_{[i]} := \tau(i, X_0^0, X_1^0, \ldots) - 1$ ($\tau_{[i]}$ wird auch als *bedingte Stoppzeit* bezeichnet). Dann gelten:

$\tau_{[i]} \in \mathcal{T}_j$ für alle $j \in S$ mit $p_{ij} > 0$ und $\mathrm{E}_i[\, u(X_\tau^0)\,] = \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\, u(X_{\tau_{[i]}}^0)\,]$.

Beweis: a.) $\{\, \tau^1 \wedge \tau^2 \le n \,\} = \{\, \tau^1 \le n \,\} \cup \{\, \tau^2 \le n \,\}$; $\{\, \tau^1 \vee \tau^2 \le n \,\} = \{\, \tau^1 \le n \,\} \cap \{\, \tau^2 \le n \,\}$.

b.) $\{\, 1+\tau \le n \,\} = \{\, \tau \le n-1 \,\} = \{\, \tau \le n \,\} - \{\, \tau = n \,\}$.

$$\begin{aligned}
\{\, \tau(X_1^0, X_2^0, \ldots) \le n \,\} &= (X_1^0, X_2^0 \ldots)^{-1}\{\, \tau \le n \,\} \\
&= (X_1^0, X_2^0 \ldots)^{-1}[\, B_n \times S \times \ldots] \quad \text{mit } B_n \subset \mathop{\times}_{j=0}^{n} S \\
&= (X_1^0, \ldots, X_{n+1}^0)^{-1}(B_n) \in \sigma(X_1^0, \ldots, X_{n+1}^0) \subset \mathcal{F}_{n+1}^0
\end{aligned}$$

$\implies \{\, 1 + \tau(X_1^0, X_2^0, \ldots) \le n \,\} \in \mathcal{F}_n^0 \quad \forall\, n \in \mathbb{N}_0$.

c.) $\tau_{X_0^0} = \sum_{i \in S} \tau_i \cdot 1_{\{\, X_0^0 = i \,\}}$ ist eine Stoppzeit. Ferner: $\mathrm{P}_j[\, \tau_{X_0^0} < \infty\,] \equiv \mathrm{P}_j[\, \tau_j < \infty\,] = 1$.

d.) Nach c.), b.) ist τ' eine Stoppzeit, da
$$\begin{aligned}
\tau' &= 1 + \tau_{X_1^0}(X_1^0, X_2^0, \ldots) \\
&= 1 + \big(\tau_{X_0^0}(X_0^0, X_1^0, \ldots)\big)(X_1^0, X_2^0, \ldots).
\end{aligned}$$

Zum Beweis von $\tau' \in \mathcal{T}_i \quad \forall\, i \in S$ definiere $u(e) := -\infty$ und $w : \mathop{\times}_{j=0}^{\infty} S \to \overline{\mathbb{R}}$ durch $w := u(X_{\tau_{X_0^0}}^0)$. Nach c.) ist w P_i-fast sicher endlich $\forall\, i \in S$. Da außerdem $w(X_1^0, X_2^2, \ldots) \equiv \big(u(X_{\tau_{X_0^0}}^0)\big)(X_1^0, X_2^0, \ldots) = u(X_{1+\tau_{X_1^0}(X_1^0, X_2^0, \ldots)}^0) \equiv u(X_{\tau'}^0)$ hat man $\tau' \in \mathcal{T}_i$ für $i \in S$. Ferner gilt mit Hilfe des Beweisprinzips für Integrale (auf w angewendent):

$$\begin{aligned}
\mathrm{E}_i[\, u(X_{\tau'}^0)\,] &\equiv \mathrm{E}_i[\, w(X_1^0, X_2^0, \ldots)\,] \\
&= \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\, w(X_0^0, X_1^0, \ldots)\,] \quad \text{nach (5.7)} \\
&\equiv \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\, u(X_{\tau_j}^0)\,]
\end{aligned}$$

e.) Der Beweis wird dem Leser überlassen.

□

Um die Ergebnisse für superharmonische Funktionen aus Abschnitt 5 auf die Wertfunktion anwenden zu können, nun das

Lemma 6.2: Die Wertfunktion ist superharmonisch.

Beweis: Zu $\epsilon > 0$, $j \in S$ wähle man $\tau_j \in \mathcal{T}_j$ mit $\mathrm{E}_j[\,u(X^0_{\tau_j}\,] \geq v(j) - \epsilon$. τ_j existiert aufgrund der Supremumseigenschaft von v. Zu τ_j, $j \in S$, definiere τ' wie in (6.1) d.).

$$\begin{aligned} \Longrightarrow v(i) &\geq \mathrm{E}_i[\,u(X^0_{\tau'})\,] \quad \text{da } \tau' \in \mathcal{T}_i \\ &= \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\,u(X^0_{\tau_j})\,] \quad \text{nach (6.1) d.)} \\ &\geq \sum_{j \in S} p_{ij} \cdot \big(v(j) - \epsilon\big) \equiv \sum_{j \in S} p_{ij} \cdot v(j) - \epsilon \end{aligned}$$

$\epsilon \to 0$ liefert die Behauptung.

q.e.d.

Eine Verallgemeinerung der Eigenschaft "superharmonisch" zeigt

Lemma 6.3: Ist $\varphi : S \to \mathbb{R}$ superharmonisch und nach unten beschränkt, so gilt:

$$\varphi(i) \geq \mathrm{E}_i[\,\varphi(X^0_\tau)\,] \quad \forall\, i \in S\,,\ \tau \in \mathcal{T}_i$$

Beweis: Nach (5.11) ist $\big(-\varphi(X_n)\big)_{n \in \mathbb{N}_0}$ ein Submartingal. Da φ nach unten beschränkt ist, gilt $\big(-\varphi(X_n)\big)^+ \leq c \quad \forall\, n \in \mathbb{N}_0$ für ein $c \geq 0$. Daher ist $\Big(\big(-\varphi(X_n)\big)^+\Big)_{n \in \mathbb{N}_0}$ gleichmäßig integrierbar, somit die Voraussetzungen vom Stoppsatz (4.6) b.) erfüllt. Wegen $\mathrm{E}_i[\,\varphi(X^0_0) = \varphi(i)$ liefert dieser die Behauptung.

q.e.d.

Die Beziehung der Wertfunktion v zur Nutzenfunktion u und deren Berechnungsmöglichkeit werden im folgenden aufgezeigt:

Lemma 6.4: Ist $\varphi : S \to \mathbb{R}$ superharmonisch, so impliziert $\varphi \geq u$ die Relation $\varphi \geq v$.

Beweis: Da $\varphi \geq u$ ist φ nach unten beschränkt und daher (6.3) anwendbar (mit $i \in S$, $\tau \in \mathcal{T}_i$):

$$\begin{aligned} \mathrm{E}_i[\,u(X^0_\tau)\,] &\leq \mathrm{E}_i[\,\varphi(X^0_\tau)\,] \quad \text{nach Voraussetzung} \\ &\leq \varphi(i) \quad \text{nach (6.3)} \end{aligned}$$

$$\Longrightarrow v(i) \equiv \sup_{\tau \in \mathcal{T}_i} \mathrm{E}_i[\,u(X^0_\tau)\,] \leq \varphi(i)$$

q.e.d.

Satz 6.5: a.) v ist die kleinste superharmonische Majorante von u.

b.) Es gilt die Optimalitätsgleichung **(6.6)** $\quad v(i) = \max\{u(i), \sum_{j \in S} p_{ij} \cdot v(j)\}$

Beweis: a.) $\tau = \underline{0} \in \mathcal{T}_i \quad \forall\, i \in S \Longrightarrow u(i) \equiv \mathrm{E}_i[\,u(X^0_\tau)\,] \leq v(i)$, da v das Supremum ist. Nach (6.2) ist v superharmonisch und nach (6.4) sogar die kleinste superharmonische Majorante.

b.) Nach a.) ist "$\geq$" klar, da $v \geq u$ und v superharmonisch. Es bleibt die Relation "$\leq$" zu zeigen. Zu jeder Stoppzeit $\tau \in \mathcal{T}_i$ existiert eine Menge $B^0_\tau \subset S$ mit $\{\,\tau = 0\,\} = \{\,X^0_0 \in B^0_\tau\,\}$, da τ an $\mathbb{F}^0$ adaptiert.

$$i \in B^0_\tau \Longrightarrow \mathrm{E}_i[\,u(X^0_\tau)\,] = \mathrm{E}_i[\,u(X^0_\tau) \cdot 1_{\{\,X^0_0 = i\,\}}\,] = \mathrm{E}_i[\,u(X^0_0)\,] \equiv u(i)$$

$$\begin{aligned} i \notin B^0_\tau \Longrightarrow \mathrm{E}_i[\,u(X^0_\tau)\,] &= \sum_{j \in S} p_{ij} \cdot \mathrm{E}_j[\,u(X^0_{\tau_{[i]}})\,] \quad \text{nach (6.1) e.)} \\ &\leq \sum_{j \in S} p_{ij} \cdot v(j)\,, \text{ da } \tau_{[i]} \in \mathcal{T}_j \text{ für } p_{ij} > 0 \end{aligned}$$

$$\Longrightarrow v(i) \leq \max\{u(i), \sum_{j \in S} p_{ij} \cdot v(j)\}$$

q.e.d.

Man nähert sich der Lösung des Problems des optimalen Stoppens durch

Definition 3: Eine Stoppzeit τ heißt <u>*optimal in $i \in S$*</u>, falls $\tau \in \mathcal{T}_i$ und $\mathrm{E}_i[\, u(X_\tau^0)\,] = v(i)$.

Setze $M^0 := \{\, i \in S : u(i) = v(i)\,\}$

$$M^+ := \{\, i \in S : v(i) = \sum_{j \in S} p_{ij} \cdot v(j)\,\}$$

sowie $\tau^0 := \inf\{\, m \in \mathbb{N}_0 : X_n^0 \in M^0\,\}$, $\inf \emptyset := \infty$.

Bemerkung: In den Zuständen aus M^0 ist der gegenwärtige Nutzen gleich dem optimalen zu erwartenden Nutzen, es lohnt sich nicht, weiterzumachen (bzw. Stoppen lohnt sich).
In M^+ ist der zu erwartende Nutzen gleich dem, wenn man vom Startpunkt i aus mindestens eine Stufe weitergeht und sich dann vom nächsten Zustand j an optimal verhält:
Im Beweis von (6.2) hatte man für $i \in S$ zu jedem $\epsilon > 0$ ein $\tau' \in \mathcal{T}_i$ mit $\tau' \geq 1$ und der Eigenschaft $\sum_{j \in S} p_{ij} \cdot v(j) - \epsilon \leq \mathrm{E}_i[\, u(X_{\tau'}^0)\,] \Longrightarrow \sum_{j \in S} p_{ij} \cdot v(j) \leq \sup_{\substack{\tau \in \mathcal{T}_i \\ \tau \geq 1}} \mathrm{E}_i[\, u(X_\tau^0)\,]$.

Andererseits gilt für $\tau \in \mathcal{T}_i$ mit $\tau \geq 1$ gemäß dem Beweis von (6.5) b.) wegen $\{\, X_0^0 = i\,\} \subset \{\, \tau > 0\,\} = \bigtimes_{j=0}^{\infty} S$: $\sup_{\substack{\tau \in \mathcal{T}_i \\ \tau \geq 1}} \mathrm{E}_i[\, u(X_\tau^0)\,] \leq \sum_{j \in S} p_{ij} \cdot v(j)$ und mithin

$$\textbf{(6.7)} \quad \sum_{j \in S} p_{ij} \cdot v(j) = \sup_{\substack{\tau \in \mathcal{T}_i \\ \tau \geq 1}} \mathrm{E}_i[\, u(X_\tau^0)\,].$$

Dies bedeutet gerade, daß es in einem Zustand $i \in M^+$ kein Fehler ist, noch einen Schritt weiterzugehen, ehe man stoppt.

Man kann mit Hilfe der beiden Mengen M^+ und M^0 jeden Zustand aus S charakterisieren. Es gilt nämlich als direkte Folgerung aus (6.6):

Korollar 6.8: $S = M^0 \cup M^+$. □

Eine weitere Folgerung ist

Korollar 6.9: Es gilt für alle $i \in S$ und $\tau \in \mathcal{T}_i$: $v(i) \overset{(1)}{\geq} \mathrm{E}_i[\, v(X_\tau^0)\,] \overset{(2)}{\geq} \mathrm{E}_i[\, u(X_\tau^0)\,]$.

Beweis: (2) folgt aus $v \geq u$, (1) aus (6.2) und (6.3).

□

Diese Ungleichungskette motiviert im Hinblick auf optimales Stoppen die

Definition 4: Eine Stoppzeit $\tau \in \mathcal{T}_i$ heißt <u>*wertkonservierend in $i \in S$*</u>, falls

$$\textbf{(6.10)} \quad v(i) = \mathrm{E}_i[\, v(X_\tau^0)\,].$$

Ferner heißt $\tau \in \mathcal{T}_i$ <u>*wertegalisierend in $i \in S$*</u>, falls

$$\textbf{(6.11)} \quad \mathrm{E}_i[\, v(X_\tau^0)\,] = \mathrm{E}_i[\, u(X_\tau^0)\,].$$

Mit diesen beiden Begriffen hat man einen Hauptsatz dieses Abschnitts:

Satz 6.12: Eine Stoppzeit $\tau \in \mathcal{T}_i$ ist genau dann optimal in $i \in S$, wenn sie wertkonservierend und wertegalisierend in i ist. Dabei ist τ genau dann wertkonservierend in i, wenn $X_n^0 \in M^+$, $0 \leq n < \tau$, P_i-fast sicher. Weiter ist τ wertegalisierend in i genau dann, wenn $X_\tau^0 \in M^0$ P_i-fast sicher.

Beweis: Die erste Aussage ist mit (6.9) klar.

Ist $X^0_\tau \in M^0$ P_i-fast sicher $\Longrightarrow$ $\mathrm{E}_i[\,v(X^0_\tau)\,] = \mathrm{E}_i[\,u(X^0_\tau)\,]$, d.h. (6.11) ist erfüllt.

Gilt umgekehrt (6.11), so ist wegen $v \geq u$ $\mathrm{P}_i[\,v(X^0_\tau) > u(X^0_\tau)\,] = 0$, also $X^0_\tau \in M^0$ P_i-f.s.

Aufgrund des Stoppsatzes (4.6) b.) hat man, da v superharmonisch, die Ungleichungskette $\mathrm{E}_i[\,v(X^0_\tau)\,] \leq \mathrm{E}_i[\,v(X^0_\tau)\,] \leq v(i) \quad \forall\, n \in \mathbb{N}_0$ (vgl. dazu (6.3)).

Es gelte nun (6.10), d.h. $v(i) = \mathrm{E}_i[\,v(X^0_\tau\,]$

$$\overset{(*)}{\Longleftrightarrow} \mathrm{E}_i[\,v(X^0_{\tau\wedge n})\,] = \mathrm{E}_i[\,v(X^0_{\tau\wedge n+1})\,] \quad \forall\, n \in \mathbb{N}_0$$

$$\Longleftrightarrow \mathrm{E}_i[\,v(X^0_n)\cdot 1_{\{\tau>n\}}\,] = \mathrm{E}_i[\,v(X^0_{n+1})\cdot 1_{\{\tau>n\}}\,] \quad \forall\, n \in \mathbb{N}_0$$

$$\Longleftrightarrow \int\limits_{\{\tau>n\}} v(X^0_{n+1}) - v(X^0_n)\; d\mathrm{P}_i = 0 \quad \forall\, n \in \mathbb{N}_0$$

$$\Longleftrightarrow \int\limits_{\{\tau>n\}} v(X^0_n) - \sum_{j\in S} \mathrm{P}_i[\,X^0_{n+1} = j \mid \mathcal{F}^0_n\,]\cdot v(j)\; d\mathrm{P}_i = 0 \quad \forall\, n \in \mathbb{N}_0$$

$$\Longleftrightarrow \int\limits_{\{\tau>n\}} v(X^0_n) - \sum_{j\in S} p_{X^0_n\, j}\cdot v(j)\; d\mathrm{P}_i = 0 \quad \forall\, n \in \mathbb{N}_0 \quad \text{nach (5.9)}$$

$$\Longleftrightarrow v(X^0_n) = \sum_{j\in S} p_{X^0_n\, j}\cdot v(j) \quad \mathrm{P}_i\text{-fast sicher auf } \{\,\tau > n\,\} \quad \forall\, n \in \mathbb{N}_0$$

$$\text{da } v(i) \geq \sum_{j\in S} p_{ij}\cdot v(j) \quad \forall\, i \in S$$

$$\Longleftrightarrow X^0_n \in M^+ \quad \mathrm{P}_i\text{-fast sicher auf } \{\,\tau > n\,\} \quad \forall\, n \in \mathbb{N}_0.$$

Dabei ist in $(*)$ "$\Rightarrow$" klar, "$\Leftarrow$" folgt mit (4.3) (v ist beschränkt).

q.e.d.

Es stellt sich natürlich die Frage nach der Existenz optimaler Stoppzeiten. Dies zeigt

Korollar 6.13: Eine in $i \in S$ optimale Stoppzeit existiert genau dann, wenn $\tau^0 \in \mathcal{T}_i$ ist. Dann ist schon τ^0 optimal in i.

Beweis: "$\Leftarrow$": $\tau^0 \in \mathcal{T}_i \Longrightarrow X^0_{\tau^0} \in M^0$ P_i-fast sicher

$\Longrightarrow \tau^0$ ist wertegalisierend in i nach (6.12).

Ferner gilt nach (6.8): $X^0_n \in M^+$ für $0 \leq n < \tau$ P_i-fast sicher

$\Longrightarrow \tau^0$ ist wertkonservierend in i nach (6.12). Mithin ist τ^0 gemäß (6.12) optimal in i.

"$\Rightarrow$": $\tau \in \mathcal{T}_i$ optimal in $i \Longrightarrow \tau$ ist wertkonservierend und wertegalisierend in i

$\Longrightarrow X^0_\tau \in M^0$ P_i-fast sicher

$\Longrightarrow \tau^0 \leq \tau$ P_i-fast sicher, da τ^0 das Infimum aller n mit dieser Eigenschaft ist

$\Longrightarrow \tau^0 \in \mathcal{T}_i$.

q.e.d.

Eine optimale Stoppzeit braucht nicht zu existieren. Dies zeigt

Beispiel 1: Sei $S = \mathbb{N}$, $p_{i\,(i+1)} := 1$ und $u(i) := \dfrac{i}{i+1}$. Dann ist $v = \underline{1}$ (v ist antiton und kleinste Majorante von u, welches isoton ist) und somit $M^0 = \emptyset$. Daher existiert keine optimale Stoppzeit.

Dieses Beispiel motiviert eine Abschwächung des Optimalitätsbegriffes:

Definition 5: Eine Stoppzeit τ heißt <u>ϵ-optimal in $i \in S$</u> ($\epsilon > 0$) $\Longleftrightarrow \tau \in \mathcal{T}_i$ und $\mathrm{E}_i[\,u(X^0_\tau)\,] \geq v(i) - \epsilon$

Konsistent zum schon bekannten Fall $\epsilon = 0$ definiert man sich

$M^\epsilon := \{\, i \in S : u(i) \geq v(i) - \epsilon\,\}$ sowie $\tau^\epsilon := \inf\{\, n \in \mathbb{N}_0 : X^0_n \in M^\epsilon\,\}$.

Daß diese Definition ausreicht, um das Problem $M^\epsilon = \emptyset$ zu vermeiden, zeigt

Lemma 6.14: Sei $\epsilon > 0$. Dann ist $M^\epsilon \neq \emptyset$ und $\tau^\epsilon \in \mathcal{T}_i \quad \forall\, i \in S$.

Beweis: Es ist $v(i) \leq \sup\limits_{j\in S} u(j)$. Daher gilt: $\emptyset \neq \{\, i \in S : u(i) > \sup\limits_{j\in S} u(j) - \epsilon \,\} \subset M^\epsilon$.

Weiter ist $\mathrm{P}_i[\,\tau^\epsilon < \infty\,] \geq \mathrm{P}_i[\,X^0_\tau \in M^\epsilon, \tau < \infty\,] = \mathrm{P}_i[\,X^0_\tau \in M^\epsilon\,] \quad \forall\, \tau \in \mathcal{T}_i$, da τ^ϵ das Infimum.

$$\begin{aligned}
\Longrightarrow \mathrm{P}_i[\,\tau^\epsilon = \infty\,] &\leq \mathrm{P}_i[\,X^0_\tau \notin M^\epsilon\,] = \mathrm{P}_i[\,v(X^0_\tau) - u(X^0_\tau) > \epsilon\,] \\
&\leq \frac{1}{\epsilon} \cdot \mathrm{E}_i[\,v(X^0_\tau) - u(X^0_\tau)\,] \quad \text{mit der Markoff-Ungleichung} \\
&\leq \frac{1}{\epsilon} \cdot \big(v(i) - \mathrm{E}_i[\,u(X^0_\tau)\,]\big) \quad \forall\, \tau \in \mathcal{T}_i \quad \text{nach (6.9)}
\end{aligned}$$

$$\Longrightarrow \mathrm{P}_i[\,\tau^\epsilon = \infty\,] \leq \inf_{\tau\in\mathcal{T}_i} \frac{1}{\epsilon} \cdot \big(v(i) - \mathrm{E}_i[\,u(X^0_\tau)\,]\big) = 0$$

q.e.d.

Da $\tau^\epsilon \in \mathcal{T}_i$ für jedes $i \in S$, hat man analog zu (6.13):

Satz 6.15: τ^ϵ ist ϵ-optimal in $i \quad \forall\, i \in S \quad \forall\, \epsilon > 0$.

Beweis: Die Bedingung $\tau^\epsilon \in \mathcal{T}_i$ für $i \in S$ ist nach obigem klar. Ferner gilt:

$$\begin{aligned}
\{\,\tau^\epsilon > n\,\} &= \{\,X^0_m \notin M^\epsilon \quad \forall\, m \leq n\,\} \subset \{\,X^0_m \notin M^0 \quad \forall\, m \leq n\,\} \\
&= \{\,X^0_m \in M^+ \quad \forall\, m \leq n\,\} \quad \text{nach (6.8)}
\end{aligned}$$

$\Longrightarrow \tau^\epsilon$ ist wertkonservierend in $i \Longrightarrow v(i) = \mathrm{E}_i[\,v(X^0_{\tau^\epsilon})\,] \quad \forall\, \epsilon > 0 \quad$ nach (6.10)

$\Longrightarrow v(i) \leq \mathrm{E}_i[\,u(X^0_{\tau^\epsilon})\,] + \epsilon$ nach Definition von τ^ϵ.

q.e.d.

Ein Spezialfall zeigt das nächste Korollar auf:

Korollar 6.16: Ist S endlich, so ist τ^0 optimal in $i \quad \forall\, i \in S$.

Beweis: S endlich $\Longrightarrow \exists\, \delta > 0$, so daß $\forall\, \epsilon > 0$ mit $\epsilon < \delta$ gilt: $M^\epsilon = M^0$, $\tau^\epsilon = \tau^0$.

(Setze z.B. $\delta := \min\{\, v(i) - u(i) : i \in S,\ v(i) > u(i)\,\}$, $\min \emptyset := \infty$)

Damit ist also τ^0 ϵ-optimal in $i \quad \forall\, 0 < \epsilon \leq \delta \quad \forall\, i \in S \Longrightarrow \tau^0$ ist optimal in $i \quad \forall\, i \in S$.

q.e.d.

Ein Fall, in dem die Wertfunktion v bekannt ist, beschreibt

Beispiel 2: Symmetrische Irrfahrt mit absorbierenden Schranken

Sei dazu $S = \{\,0, \ldots, c\,\}$, $p_{00} = p_{cc} := 1$ sowie $p_{i,i+1} = p_{i,i-1} = \dfrac{1}{2}$. Dann ist eine Funktion $\varphi : S \to \mathbb{R}$ genau dann superharmonisch, wenn $\varphi(i) \geq \dfrac{1}{2} \cdot \big(\varphi(i-1) + \varphi(i+1)\big)$ gilt, $0 < i < c$.

Mit anderen Worten: φ ist konkav auf S. Demnach ist v gemäß (6.5) a.) die kleinste konkave Majorante von u und kann graphisch ermittelt werden.

$\big($ Ist z.B. $u = 1_{\{c\}}$ (vgl. Ruinproblem in Abschnitt 4), so ist $v(i) = \dfrac{i}{c}$ $\big)$.

Man hat nun das Problem des optimalen Stoppens durch den Übergang zur internen Markoff-Familie für diese gelöst. Nun wird der Zusammenhang zwischen dem Hilfsproblem und der Ausgangssituation aufgezeigt:

Es war der Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ und die Markoff-Kette $(X, \mathbb{F})$ mit Übergangsmatrix $\mathbb{P}$ und Startverteilung ν gegeben. Mit $\tau \in \mathcal{T}$ gilt analog zu (6.9):

$$\begin{aligned}
\mathrm{E}[\,u(X_\tau)\,] &\leq \mathrm{E}[\,v(X_\tau)\,] \quad \text{da } u \leq v \\
&\leq \mathrm{E}[\,v(X_0)\,] \quad \text{nach (4.6)} \\
&= \sum_{i\in S} \nu(\{\,i\,\}) \cdot v(i) \;=:\; v(\nu)
\end{aligned}$$

Mithin ist also **(6.17)** $\sup_{\tau \in \mathcal{T}} \mathrm{E}[\, u(X_\tau)\,] \le v(\nu)$.

Da jedes $\tau \in \mathcal{T}_i$ eine Funktion auf $\bigtimes_{j=0}^{\infty} S$ ist, ist $\hat\tau := \tau(X_0, X_1, \ldots)$ definiert. Hat man z.B. die ϵ-optimale Stoppzeit $\tau \equiv \tau^\epsilon = \inf\{\, n : X_n^0 \in M^\epsilon \,\}$, so folgt $\hat\tau = \inf\{\, n : X_n \in M^\epsilon \,\}$.

Es ergeben sich die folgenden Zusammenhänge zwischen Hilfsproblem und Ausgangssituation:

Lemma 6.18: Für $\tau \in \bigcap_{\substack{i\in S\\ \nu(\{i\})>0}} \mathcal{T}_i$ gilt: $\hat\tau := \tau(X_0, X_1, \ldots) \in \mathcal{T}$ und $\mathrm{E}[\, u(X_{\hat\tau})\,] = \mathrm{E}_\nu[\, u(X^0_\tau)\,]$.

Beweis:
$$\begin{aligned}\{\, \hat\tau = n \,\} &= (X_0, X_1, \ldots)^{-1}\{\, \tau = n \,\}\\ &\equiv (X_0, X_1, \ldots)^{-1}[\, B_n \times S \times \ldots \,] \text{ mit } B_n \subset \bigtimes_{j=0}^{n} S, \text{ da } \tau \text{ Stoppzeit bzgl. } \mathbb{F}^0 \text{ ist}\\ &= (X_0, \ldots, X_n)^{-1}[\, B_n \,] \in \sigma(X_0, \ldots, X_n) \subset \mathcal{F}_n \text{ , da } X \text{ adaptiert.}\end{aligned}$$

Ferner gilt:
$$\begin{aligned}\mathrm{P}[\, \tau(X_0, X_1, \ldots) < \infty \,] &= \mathrm{P}_\nu[\, \tau(X^0_0, X^0_1, \ldots) < \infty \,] \quad \text{nach (5.8)}\\ &\equiv \mathrm{P}_\nu[\, \tau < \infty \,]\\ &= \sum_{i\in S} \nu(\{\, i \,\}) \cdot \mathrm{P}_i[\, \tau < \infty \,] \quad \text{nach (5.8}'\text{)}\\ &= 1\end{aligned}$$

$$\begin{aligned}\mathrm{E}[\, u(X_{\tau(X_0,X_1,\ldots)})\,] &= \sum_{n\in\mathbb{N}_0} \sum_{i\in S} u(i) \cdot \mathrm{P}[\, X_n = i\,,\ \tau(X_0, X_1, \ldots) = n \,]\\ &\equiv \sum_{n\in\mathbb{N}_0} \sum_{i\in S} u(i) \cdot \mathrm{P}[\, (X_0, X_1, \ldots) \in \{\, \tau = n \,\} \cap \{\, \underbrace{S \times \ldots \times S}_{n\text{-fach}} \times i \times S \times \ldots \,\}\,]\\ &= \sum_{n\in\mathbb{N}_0} \sum_{i\in S} u(i) \cdot \mathrm{P}_\nu[\, \{\, \tau = n \,\} \cap \{\, S \times \ldots \times S \times i \times S \times \ldots \,\}\,] \quad \text{nach (5.8)}\\ &\equiv \sum_{n\in\mathbb{N}_0} \sum_{i\in S} u(i) \cdot \mathrm{P}_\nu[\, \tau = n\,,\ X^0_n = i \,] \equiv \mathrm{E}_\nu[\, u(X^0_\tau)\,]\end{aligned}$$

q.e.d.

Diese Resultate führen zu

Satz 6.19: a.) Es gilt $v(\nu) = \sup_{\tau\in\mathcal{T}} \mathrm{E}[\, u(X_\tau)\,]$.

b.) Gilt für $\epsilon > 0$ und $\tau \in \bigcap_{\substack{i\in S\\ \nu(\{i\})>0}} \mathcal{T}_i$ die Ungleichung $\mathrm{E}_i[\, u(X^0_\tau)\,] \ge v(i) - \epsilon$,

so folgt für $\hat\tau = \tau(X_0, X_1, \ldots)$: $\mathrm{E}[\, u(X_{\hat\tau})\,] \ge v(\nu) - \epsilon$.

Beweis: In a.) gilt "$\ge$" nach (6.17). Für "$\le$" wähle man τ wie in b.) beschrieben. (Diese Wahl ist nach (6.15) für $\epsilon > 0$ stets möglich, z.B. τ als ϵ-optimale Stoppzeit.)

Dann gilt:
$$\begin{aligned}\mathrm{E}[\, u(X_{\hat\tau})\,] &\equiv \mathrm{E}_\nu[\, u(X^0_\tau)\,] \quad \text{nach (6.18)}\\ &= \sum_{i\in S} \nu(\{\, i \,\}) \cdot \mathrm{E}_i[\, u(X^0_\tau)\,] \quad \text{nach (5.8}'\text{)}\\ &\ge \sum_{i\in S} \nu(\{\, i \,\}) \cdot v(i) - \epsilon = v(\nu) - \epsilon\end{aligned}$$

Damit ist b.) bewiesen und $\epsilon \to 0$ liefert "$\le$" in a.), da $\hat\tau \in \mathcal{T}$ gemäß (6.18).

q.e.d.

Bemerkung: (6.19) zeigt also, daß der optimale erreichbare mittlere Nutzen $v(\nu)$ ist. Hat man ferner im Hilfsproblem eine ϵ-optimale Stoppzeit τ gefunden, so verhält sich die zusammengesetzte Stoppzeit $\hat\tau$ im Bezug auf $v(\nu)$ ebenfalls ϵ-optimal.

Diese Theorie wird nun zur Lösung des schon bekannten Sekretärinnenproblems verwendet:

Beispiel 3: Wie im Beispiel 3 von Abschnitt 5 sei $S = \{(0,0)\} \cup (\{0,1\} \times \{1,\ldots,s\}) \cup \{(0,s+1)\}$. Man sucht eine Stoppzeit τ bzgl. $\mathbb{F}$, so daß die Größe $\mathrm{P}[X'_\tau = 1, X'_l = 0, \tau < l \leq s]$ maximal wird. Dabei bedeutet der obige Ausdruck gerade, daß die τ-te Sekretärin die absolut beste von allen ist, man also die beste Sekretärin einstellt. Um eine Nutzenfunktion zu gewinnen, schreibe:

$$\begin{aligned}\mathrm{P}[X'_\tau = 1, X'_l = 0, \tau < l \leq s] &= \sum_{n=0}^{s+1} \mathrm{P}[\tau = n, X'_\tau = 1, X'_l = 0, \tau < l \leq s] \\ &= \sum_{n=0}^{s+1} \mathrm{P}[\tau = n, X'_n = 1] \cdot \prod_{l=n+1}^{s} \mathrm{P}[X'_l = 0] \quad \text{nach (5.27)} \\ &\equiv \sum_{n=0}^{s+1} \mathrm{E}[1_{\{\tau=n\}} \cdot u(X'_n, n)] \equiv \mathrm{E}[u(X_\tau)]\end{aligned}$$

mit $u(i,n) := 1_{\{1\}}(i) \cdot \prod_{l=n+1}^{s} \frac{l-1}{l} = \begin{cases} \frac{n}{s} & \text{für } i = 1 \\ 0 & \text{für } i = 0 \end{cases}$

Damit man die Ergebnisse dieses Abschnittes nutzen kann, gehe man nun zur internen Markoff-Familie über. Der Zustand $(0, s+1)$ ist absorbierend, man verläßt $(0, s+1)$ fast sicher nicht und der erwartete Nutzen ist 0. Folglich: $v(0, s+1) = 0$.
Die Optimalitätsgleichung (6.6) führt zu $(0 \leq m < s)$:

$$\textbf{(6.20)} \quad \begin{cases} v(0,s) = \max\{u(0,s), v(0,s+1)\} = 0 = v(0,s+1) \\ v(1,s) = \max\{u(1,s), v(0,s+1)\} = 1 \\ v(0,m) = \max\{u(0,m), \frac{v(1,m+1)}{m+1} + \frac{m \cdot v(0,m+1)}{m+1}\} \\ \qquad = \frac{v(1,m+1)}{m+1} + \frac{m \cdot v(0,m+1)}{m+1} \\ v(1,m) = \max\{u(1,m), v(0,m)\} \equiv \max\{\frac{m}{s}, v(0,m)\} \end{cases}$$

Setze nun $f_s(m) \equiv f(m) := \frac{1}{m} + \frac{1}{m+1} + \ldots + \frac{1}{s-1}$, $0 \leq m \leq s$. f ist antiton und es gilt: $f(0) = \infty$, $f(1) > 1$ (im Fall $s > 2$, die Fälle $s = 1,2$ sind nicht interessant) sowie $f(s) = 0$. Wähle nun ein $k = k(s)$ derart, daß

(6.21) $\quad s > k = k(s) \geq 0$, $f(k) > 1 \geq f(k+1)$ (also $k \geq 1$ bei $s > 2$).

Damit ergibt sich

$$\textbf{(6.22)} \quad \begin{cases} v(0,m) = \frac{m}{s} \cdot f(m), \; v(1,m) = \frac{m}{s} & \text{falls } s \geq m > k \\ v(i,m) = \frac{k}{s} \cdot f(k) & \text{falls } i = 0,1 \text{ und } m \leq k \end{cases}$$

<u>Beweis</u> durch Rückwärtsinduktion: Der Fall $m = s$ ist klar.

$m+1 \to m \geq k$: $\quad v(0,m) = \frac{1}{m+1} \cdot \frac{m+1}{s} + \frac{m}{m+1} \cdot \frac{m+1}{s} \cdot f(m+1)$

$$= \frac{m}{s} \cdot \left(f(m+1) + \frac{1}{m}\right) \equiv \frac{m}{s} \cdot f(m).$$

$$v(1,m) = \max\{\frac{m}{s}, \frac{m}{s} \cdot f(m)\} = \begin{cases} \frac{m}{s} & \text{für } m > k \\ \frac{k}{s} \cdot f(k) & \text{für } m = k \end{cases}$$

$k \geq m+1 \to m \geq 0: \quad v(0,k) = \frac{k}{s} \cdot f(k) = v(1,k)$ nach obigem

$$\Longrightarrow v(0,m) = \frac{1}{m+1} \cdot \frac{k}{s} \cdot f(k) + \frac{m}{m+1} \cdot \frac{k}{s} \cdot f(k) = \frac{k}{s} \cdot f(k)$$

und $v(1,m) = \max\{\frac{m}{s}, \frac{k}{s} \cdot f(k)\} = \frac{k}{s} \cdot f(k)$ weil $\frac{m}{s} \leq \frac{k}{s} \leq \frac{k}{s} \cdot f(k)$

Dies liefert als Resultat **(6.23)** $M^0 = \{(0,s), (0,s+1)\} \cup \{(1,m),\ m > k\}$

als die Menge der Zustände, wo sich ein Aufhören lohnt, d.h. man eine Sekretärin einstellen sollte (dabei sind die ersten beiden Punkte natürlich uninteressant). Die Erfolgswahrscheinlichkeit wird charakterisiert durch

(6.24) $v(0,0) = \frac{k}{s} \cdot f(k) \equiv v(\nu)$, wo ν die auf $(0,0)$ konzentrierte Startverteilung ist.

Als Zusammenfassung:

Satz 6.25: Für das Problem der besten Wahl ist bei $s > 2$ die optimale Lösung:
Schaue die ersten k an, ohne sie zu nehmen, wobei sich $k = k(s)$ gemäß (6.21) ergibt.
Danach nimm die erste, die besser ist als die k vorangehenden. Die Wahrscheinlichkeit, dabei die insgesamt beste Wahl zu treffen, ist $\frac{k}{s} \cdot f(k)$.

Beispiel: $s = 3 \implies k = 1,\ f(k) = \frac{1}{1} + \frac{1}{2} = \frac{3}{2} \implies v(\nu) = \frac{1}{3} \cdot \frac{3}{2} = \frac{1}{2}.$

Bemerkung: Für großes s kann man $k(s)$ durch Integrale approximieren:

Es gilt $\int_m^s \frac{1}{x}\,dx \leq f(m) \leq \int_{m-1}^{s-1} \frac{1}{x}\,dx$, da $f(m)$ jeweils eine Ober- bzw. Untersumme von dem jeweiligen Integral ist. Mit anderen Worten:

$\ln \frac{s}{m} \leq f(m) \leq \ln \frac{s-1}{m-1} \implies \ln \frac{s-1}{k-1} > 1 \geq \ln \frac{s}{k+1}$ nach (6.21). Daraus folgt

(6.26) $\frac{s}{e} - 1 \leq k < \frac{s}{e} + \frac{e-1}{e}$ und daher $\frac{k(s)}{s} \longrightarrow \frac{1}{e}$, $(s \to \infty)$, dabei $\frac{1}{e} \approx \frac{1}{3}$.

Aus $1 < f(k) = \frac{1}{k} + f(k+1) \leq \frac{1}{k} + 1$ und (6.24) folgt: $v(\nu) \longrightarrow \frac{1}{e}$, $(s \to \infty)$.

Es ist erstaunlich, daß dies ein Verfahren ist, bei dem die Erfolgswahrscheinlichkeit bei wachsendem s nicht gegen 0 konvergiert.

Im letzten Abschnitt dieses Kapitels werden, wie im Abschnitt 5 schon angedeutet, Aussagen über das Verhalten einer Markoff-Kette während eines unendlichen Zeitraums gewonnen, mit anderen Worten

Abschnitt 7: Asymptotisches Verhalten von Markoff-Ketten

Für die nun folgenden Betrachtungen treffe man folgende Konvention:
S sei ein abzählbarer Raum, $(X, \{\mathrm{P}_i\}_{i \in S})$ interne Markoff-Familie zu einer Markoff-Kette mit Übergangsmatrix $\mathbb{P}$ und $\mathbb{F}$ die kanonische Filterung in $\bigotimes_{j=0}^{\infty} \mathrm{Pot}(S)$ (vgl. Definition 3 in Abschnitt 5). ν bezeichne jeweils eine Startverteilung auf S.

Diese Konvention stellt nach den Ergebnissen aus Abschnitt 5 keine Einschränkung der Allgemeinheit dar: Hat man zu einer Übergangsmatrix $\mathbb{P}$ und einer Startverteilung ν eine Markoff-Kette $(X', \mathbb{F}')$ mit Zustandsraum S auf einem beliebigen Wahrscheinlichkeitsraum $(\Omega', \mathcal{F}', \mathrm{P}')$, so stimmt nach Abschnitt 5 die Verteilung von X unter P_ν mit der von X' unter P' überein. Die folgenden Resultate/Überlegungen gelten daher mit X auch für X'.

Zur Behandlung der Frage wann, wie oft und mit welcher Wahrscheinlichkeit die Markoff-Kette X einen Zustand $h \in S$ nach einem Start in $i \in S$ erreicht, mache die

Definition 1: Für $h \in S$ setze man $T_0 := 0$

$$T_1 \equiv T := \inf\{ n > 0 : X_n = h \}$$

und allgemein: $T_{n+1} := \inf\{ n > T_n : X_n = h \}$

Dabei werde jeweils $\inf \emptyset := \infty$ gesetzt. T_n gibt den Zeitpunkt des n-ten Besuchs von X in h an. Der Prozeß $Z = (Z_n)_{n \in \mathbb{N}}$ beschreibe die Zeitabstände zwischen den einzelnen Besuchen in h, d.h. setze $Z_n := T_n - T_{n-1}$ (wo $\infty - \infty := \infty$).

Ferner sei noch für $i \in S$ und $n \in \mathbb{N}$ $\quad f_{ih}^{(n)} := \mathrm{P}_i[T = n] \equiv \sum\limits_{i_1 \neq h, \ldots, i_{n-1} \neq h} p_{i i_1} \cdot \ldots \cdot p_{i_{n-1} h}$

die Wahrscheinlichkeit, bei einem Start in i nach genau n Schritten zum ersten Male nach h zu kommen bzw. für $i = h$ nach h zurückzukehren.

Für den Prozeß Z hat man

Lemma 7.1: Für $m_1, \ldots, m_n \in \mathbb{N}$ mit $n \in \mathbb{N}$ gilt:

$$\mathrm{P}_\nu[Z_1 = m_1, \ldots, Z_n = m_n] = \mathrm{P}_\nu[Z_1 = m_1] \cdot f_{hh}^{(m_2)} \cdot \ldots \cdot f_{hh}^{(m_n)}$$

Beweis: Setze für $t \in \{ 1, \ldots, n \}$ $\quad A_t := \{ X_{m_1 + \ldots + m_{t-1} + l} \neq h \,,\ 0 < l < m_t \,,\ X_{m_1 + \ldots + m_t} = h \}$.

Dann folgt mit $m_0 := 0$:

$$\begin{aligned}
\mathrm{P}_\nu[Z_1 = m_1, \ldots, Z_n = m_n] &= \mathrm{P}_\nu[\bigcap_{t=1}^{n-1} A_t \cap A_n] \\
&\equiv \int\limits_{\bigcap_{t=1}^{n-1} A_t} 1_{A_n} \, d\mathrm{P}_\nu \\
&\equiv \int\limits_{\bigcap_{t=1}^{n-1} A_t} \mathrm{P}_\nu[A_n \mid \mathcal{F}_{m_1 + \ldots + m_{n-1}}] \, d\mathrm{P}_\nu \\
&\overset{(5.7)}{=} \int\limits_{\bigcap_{t=1}^{n-1} A_t} \mathrm{P}_{X_{m_1 + \ldots + m_{n-1}}}[X_l \neq h \,,\ 0 < l < m_n \,,\ X_{m_n} = l] \, d\mathrm{P}_\nu \\
&\overset{(*)}{=} \int\limits_{\bigcap_{t=1}^{n-1} A_t} \mathrm{P}_h[X_l \neq h \,,\ 0 < l < m_n \,,\ X_{m_n} = l] \, d\mathrm{P}_\nu \\
&\equiv \mathrm{P}_\nu[\bigcap_{t=1}^{n-1} A_t] \cdot f_{hh}^{(m_n)}
\end{aligned}$$

$(*)$ besteht, da $X_{m_1 + \ldots + m_{n-1}} = h$ auf $\bigcap\limits_{t=1}^{n-1} A_t$. Die Behauptung folgt nun mit Induktion.

q.e.d.

Nun eine Erweiterung der zu Verfügung stehenden Begriffe:

Definition 2: Sei wieder $h \in S$ fest. Für $i \in S$ definiere man $f^*_{ih} := \sum\limits_{m=1}^{\infty} f^{(m)}_{ih} \equiv \mathrm{P}_i[\,T < \infty\,]$ als die Wahrscheinlichkeit, jemals nach h (zurück-) zu kommen.

Setze ferner $m_{ih} := \left\{ \begin{array}{ll} \sum\limits_{m=1}^{\infty} m \cdot f^{(m)}_{ih} & \text{falls } f^*_{ih} = 1 \\ & \\ \infty & \text{sonst} \end{array} \right\} \equiv \mathrm{E}_i[\,T\,]$ als die mittlere Zeit,

von i nach h zu kommen (d.h. m_{hh} ist die <u>*mittlere Rückkehrzeit*</u>).

Damit unterscheidet man: h heißt $\begin{cases} \underline{\textit{rekurrent}}, \text{ falls } f^*_{hh} = 1. \\ \underline{\textit{transient}}, \text{ falls } f^*_{hh} < 1. \\ \underline{\textit{positiv rekurrent}}, \text{ falls } m_{hh} < \infty. \\ \underline{\textit{null-rekurrent}}, \text{ falls } m_{hh} = \infty \text{ und } f^*_{hh} = 1. \end{cases}$

Für $n \in \overline{\mathbb{N}}$ sei $N(n) := \begin{cases} |\{\, l \in \mathbb{N} : T_l \leq n \,\}| \equiv |\{\, 1 \leq m \leq n : X_m = h \,\}| & \text{für } n < \infty \\ |\{\, l \in \mathbb{N} : T_l < \infty \,\}| & \text{für } n = \infty \end{cases}$

$N(n)$ ist die Anzahl der Besuche in h zwischen den Zeitpunkten 1 und n.

Als letzte Kennzahl definiere man noch $g_{ih} := \mathrm{P}_i[\,X_n = h \text{ für unendlich viele } n\,]$, $i \in S$.

Bemerkungen: 1.) Die Bezeichnung "positiv rekurrent" stammt von der Beziehung $\dfrac{1}{m_{hh}} > 0$, der Begriff "null-rekurrent" von $\dfrac{1}{m_{hh}} = 0$.

2.) Summiert man in (7.1) über m_t , $1 \leq t \leq n$, so erhält man

(<u>7.2′</u>) $\mathrm{P}_\nu[\,T_n < \infty\,] = \mathrm{P}_\nu[\,T < \infty\,] \cdot (f^*_{hh})^{n-1}$, $n \in \mathbb{N}$.

Summiert man in (7.1) über m_t , $1 \leq t < n$, so erhält man

(<u>7.2″</u>) $\mathrm{P}_\nu[\,Z_n = m_n\,] = \mathrm{P}_\nu[\,T < \infty\,] \cdot (f^*_{hh})^{n-2} \cdot f^{(m_n)}_{hh}$, $n > 2$.

Mit Hilfe von (7.1) und (7.2″) erkennt man sofort: $(Z_n)_{n \in \mathbb{N}}$ bildet unter P_h eine Folge von unabhängigen, identisch verteilten, fast sicher endlichen Zufallsvariablen, welche gemäß der Zähldichte $(f^{(k)}_{hh})_{k \in \mathbb{N}}$ verteilt sind, falls h rekurrent ist.

Für $N(n)$ hat man das

Lemma 7.3: a.) $\mathrm{E}_\nu[\,N(n)\,] = \sum\limits_{l=1}^{n} \mathrm{P}_\nu[\,T_l \leq n\,] = \sum\limits_{m=1}^{n} \mathrm{P}_\nu[\,X_m = h\,]$

b.) $\mathrm{E}_\nu[\,N(\infty)\,] = \sum\limits_{m=1}^{\infty} \mathrm{P}_\nu[\,X_m = h\,] = \dfrac{\mathrm{P}_\nu[\,T < \infty\,]}{1 - f^*_{hh}}$

Beweis: a.) $N(n) = \sum\limits_{l=1}^{n} 1_{\{\,T_l \leq n\,\}} = \sum\limits_{m=1}^{n} 1_{\{\,X_m = h\,\}}$ nach Definition 2.

b.) Nach Definition 2 gilt: $N(\infty) = \sum_{l=1}^{\infty} 1_{\{T_l < \infty\}}$. Daher mit monotoner

Konvergenz: $\mathrm{E}_\nu[N(\infty)] \equiv \mathrm{E}_\nu[\sum_{l=1}^{\infty} 1_{\{T_l<\infty\}}] = \sum_{l=1}^{\infty} \mathrm{P}_\nu[T_l < \infty]$

$$= \sum_{l=1}^{\infty} \mathrm{P}_\nu[T_l < \infty] \cdot (f_{hh}^*)^{n-1} \quad \text{nach } (7.2')$$

$$= \frac{\mathrm{P}_\nu[T < \infty]}{1 - f_{hh}^*} \quad \text{nach geometrischer Reihenformel}$$

q.e.d.

Wie man bei der Definition von g_{ih} vermuten konnte, erfüllt g_{hh} ein 0-1-Gesetz. Dieses rechtfertigt die Begriffe "rekurrent" und "transient":

Satz 7.4: (0-1-Gesetz) Für $h \in S$ gilt: $g_{hh} = \begin{cases} 1 & \text{falls } h \text{ rekurrent} \\ 0 & \text{falls } h \text{ transient} \end{cases}$

Beweis: $g_{hh} \equiv \mathrm{P}_h[T_n < \infty \quad \forall n \in \mathbb{N}] = \mathrm{P}_h[\lim_{n\to\infty}\{T_n < \infty\}]$ da $\{T_{l+1} < \infty\} \subset \{T_l < \infty\}$

$= \lim_{n\to\infty} \mathrm{P}_h[T_n < \infty]$ da P_h ein Wahrscheinlichkeitsmaß

$= \lim_{n\to\infty} (f_{hh}^*)^n$ nach $(7.2')$

q.e.d.

Ein nützliches notwendiges und hinreichendes Kriterium für Transienz (und damit für Rekurrenz) liefert

Satz 7.5: Es gelten die folgenden Äquivalenzen:

$$h \text{ ist transient} \iff \sum_{n=1}^{\infty} p_{hh}^{(n)} < \infty \iff \sum_{n=1}^{\infty} p_{ih}^{(n)} < \infty \quad \forall i \in S.$$

Beweis: Nach (7.3) b.) mit i statt ν gilt: $\sum_{m=1}^{\infty} p_{ih}^{(m)} \equiv \sum_{m=1}^{\infty} \mathrm{P}_i[X_m = h] = \frac{f_{ih}^*}{1 - f_{hh}^*}$.

Diese Formel und ein Ringschluß ergeben obige Äquivalenzen.

□

Aus der Wahrscheinlichkeitstheorie kennt man schon den Begriff des Erneuerungsprozesses und die damit verbundenen Ergebnisse (vgl. § 32). Mit deren Hilfe wird nun ein Satz gewonnen, der erste Aussagen über das asymptotische Verhalten von X, nämlich das Grenzwertverhalten des arithmetischen Mittels der Eintrittswahrscheinlichkeit von X in $\{h\}$, zulässt.

Satz 7.6: Ist h rekurrent, so gilt:

$$\frac{1}{n}\,\mathrm{E}_\nu[N(n)] \;\Big(\overset{(7.3)}{\equiv} \frac{1}{n}\sum_{m=1}^{n} \mathrm{P}_\nu[X_m = h]\Big) \longrightarrow \frac{\mathrm{P}_\nu[T < \infty]}{m_{hh}},\ (n \to \infty)$$

Beweis: Ist $\mathrm{P}_\nu[T = k] \equiv \mathrm{P}_\nu[Z_1 = k] > 0$, so gilt nach (7.1) für $m_j \in \mathbb{N}$:

$$\mathrm{P}_\nu[Z_2 = m_2, \ldots, Z_n = m_n \mid T = k] = \prod_{j=2}^{n} f_{hh}^{(m_j)}$$

Analog zu Bemerkung 2 sind unter $\mathrm{P}_\nu^k[\,\cdot\,] := \mathrm{P}_\nu[\,\cdot \mid T = k]$ $Z_2, Z_3, \ldots$ unabhängig und identisch verteilt gemäß der Zähldichte $(f_{hh}^{(m)})_{m\in\mathbb{N}}$. Mithin bildet $(T_l - T)_{l\geq 2}$ unter P_ν^k einen Erneuerungsprozeß mit Erneuerungsfunktion $H(n) \equiv \sum_{l=2}^{\infty} \mathrm{P}_\nu^k[T_l - T \leq n] \overset{(*)}{=} \mathrm{E}_h[N(n)]$.

Beweis von (∗):

$$\sum_{l=2}^{\infty} \mathrm{P}_\nu^k[\,T_l - T \le n\,] \equiv \sum_{l=2}^{\infty} \mathrm{P}_\nu^k[\sum_{j=2}^{l} Z_j \le n\,] = \sum_{l=2}^{\infty} \sum_{\substack{j_2+\ldots+j_l \le n \\ j_i \in \mathbb{N}}} \mathrm{P}_\nu^k[\,Z_2 = j_2, \ldots, Z_l = j_l\,]$$

$$= \sum_{l=2}^{\infty} \sum_{\substack{j_2+\ldots+j_l \le n \\ j_i \in \mathbb{N}}} \prod_{s=2}^{l} f_{hh}^{(j_s)}$$

$$= \sum_{l=2}^{\infty} \sum_{\substack{j_2+\ldots+j_l \le n \\ j_i \in \mathbb{N}}} \mathrm{P}_h[\,Z_1 = j_2, \ldots, Z_{l-1} = j_l\,] \quad \text{nach Bemerkung 2}$$

$$\equiv \sum_{l=1}^{\infty} \mathrm{P}_h[\,T_l \le n\,] = \mathrm{E}_h[\,N(n)\,] \quad \text{nach (7.3) a.)}$$

Für $l \ge 2$ gilt weiter: $\mathrm{P}_\nu[\,T_l \le n\,] = \mathrm{P}_\nu[\,1 \le T < T_l \le n\,] = \sum_{t=1}^{n-1} \mathrm{P}_\nu[\,T = t, T_l - T \le n - t\,]$

$$\equiv \sum_{t=1}^{n-1} \mathrm{P}_\nu[\,T = t\,] \cdot \mathrm{P}_\nu^t[\,T_l - T \le n - t\,]$$

$$\Longrightarrow \mathrm{E}_\nu[\,N(n)\,] \equiv \sum_{l=1}^{\infty} \mathrm{P}_\nu[\,T_l \le n\,] = \mathrm{P}_\nu[\,T \le n\,] + \sum_{l=2}^{\infty} \mathrm{P}_\nu[\,T_l \le n\,]$$

$$= \mathrm{P}_\nu[\,T \le n\,] + \sum_{t=1}^{n-1} \mathrm{P}_\nu[\,T = t\,] \cdot \sum_{l=2}^{\infty} \mathrm{P}_\nu^t[\,T_l - T \le n - t\,] \quad \text{nach vorigem}$$

Zusammen mit (∗) ergibt dies:

(7.7a) $\mathrm{E}_\nu[\,N(n)\,] = \mathrm{P}_\nu[\,T \le n\,] + \sum_{t=1}^{n-1} \mathrm{P}_\nu[\,T = t\,] \cdot \mathrm{E}_h[\,N(n-t)\,]$

Mit $H(m) = 0$ für $m \le 0$ folgt aus (7.7a) sofort (dabei ist O Landau'sches Symbol)

(+) $\quad \frac{1}{n}\, \mathrm{E}_\nu[\,N(n) = \mathrm{O}(\frac{1}{n}) + \sum_{t=1}^{\infty} \mathrm{P}_\nu[\,T = t\,] \cdot \frac{H(n-t)}{n}$

Nun gilt nach (7.3) $H(n) \equiv \mathrm{E}_h[\,N(n)\,] = \sum_{l=1}^{n} \mathrm{P}_\nu[\,T_l \le n\,] \le n$ sowie damit für alle $t \in \mathbb{N}$ gemäß dem elementaren Erneuerungssatz:

$$1 \ge \frac{H(n-t)}{n} \equiv \frac{n-t}{n} \cdot \frac{H(n-t)}{n-t} \longrightarrow \frac{1}{\mathrm{E}_h[\,Z_1\,]} \equiv \frac{1}{\mathrm{E}_h[\,T\,]} \equiv \frac{1}{m_{hh}}\,, \ (n \to \infty)$$

Wegen $0 \le \sum_{t=1}^{\infty} \mathrm{P}_\nu[\,T = t\,] \cdot \frac{H(n-t)}{n} \le \mathrm{P}_\nu[\,T < \infty\,] \le 1 \quad \forall\, n \in \mathbb{N}$ folgt daher die Behauptung aus (+) mit majorisierter Konvergenz.

q.e.d.

Aus (7.7a) erhält man durch die Betrachtung von $\mathrm{E}_\nu[\,N(n)\,] - \mathrm{E}_\nu[\,N(n-1)\,]$ das äquivalente Nebenresultat

(7.7b) $\mathrm{P}_\nu[\,X_n = h\,] = \sum_{t=1}^{n} \mathrm{P}_\nu[\,T = t\,] \cdot p_{hh}^{(n-t)}$

Als Verallgemeinerung lassen sich noch leicht weitere Wahrscheinlichkeiten berechnen, so z.B.:

(7.8) $\mathrm{P}_\nu[\,X_n \in B, T \le n\,] = \sum_{t=1}^{n} \mathrm{P}_\nu[\,T = t\,] \cdot \mathrm{P}_h[\,X_{n-t} \in B\,]$ für $B \subset S$

Dies bedeutet, daß die Wahrscheinlichkeit eines Eintritts von X in die Menge B im Zeitpunkt n unter der Bedingung, vorher in h gewesen zu sein, gleich der Summe der jeweiligen Wahrscheinlichkeiten ist, nach Erreichen von h im Zeitpunkt $t \le n$ den Prozeß erneut zu starten und dann im Zeitpunkt $n-t$ B zu erreichen.

<u>Beweis:</u>
$$\begin{aligned} \mathrm{P}_\nu[\,X_n \in B, T \le n\,] &\equiv \sum_{t=1}^{n} \mathrm{P}_\nu[\,X_n \in B, T = t\,] = \sum_{t=1}^{n} \int_{\{T=t\}} 1_{\{X_n \in B\}}\, d\mathrm{P}_\nu \\ &\equiv \sum_{t=1}^{n} \int_{\{T=t\}} \mathrm{P}_\nu[\,X_n \in B \mid \mathcal{F}_t\,]\, d\mathrm{P}_\nu \overset{(5.7)}{=} \sum_{t=1}^{n} \int_{\{T=t\}} \mathrm{P}_{X_t}[\,X_{n-t} \in B\,]\, d\mathrm{P}_\nu \\ &\equiv \sum_{t=1}^{n} \int_{\{T=t\}} \mathrm{P}_h[\,X_{n-t} \in B\,]\, d\mathrm{P}_\nu\,, \text{ da } X_t = h \text{ auf } \{T=t\}. \end{aligned}$$

q.e.d.

Wie schon angekündigt, führt (7.6) zu einer ersten Aussage über das asymptotische Verhalten der n-Schritt-Übergangswahrscheinlichkeiten, nämlich die Konvergenz im arithmetischen Mittel (Cesaro-Konvergenz). Mache dazu die

<u>Definition 3:</u> Definiere $\overline{p}_{ih}^{(n)} := \dfrac{1}{n} \sum_{m=1}^{n} p_{ih}^{(m)}$ sowie $p_{ih}^{(\infty)} := \dfrac{f_{ih}^*}{m_{hh}}$.

Mit diesen Setzungen formuliert man

<u>Korollar 7.9:</u> a.) $\overline{p}_{ih}^{(n)} \longrightarrow p_{ih}^{(\infty)}$, $(n \to \infty)$.

b.) $p_{hh}^{(\infty)} = \dfrac{1}{m_{hh}}$

<u>Beweis:</u> a.) h transient $\Longrightarrow \sum_{m=1}^{\infty} p_{ih}^{(m)} < \infty$ nach (7.5)

$$\begin{aligned} &\Longrightarrow p_{ih}^{(m)} \longrightarrow 0\,,\ (m \to \infty) \\ &\Longrightarrow \overline{p}_{ih}^{(n)} \longrightarrow 0 \equiv p_{ih}^{(\infty)}\,,\ (n \to \infty). \end{aligned}$$

h rekurrent $\Longrightarrow \overline{p}_{ih}^{(n)} \equiv \dfrac{1}{n} \sum_{m=1}^{n} p_{ih}^{(m)} = \dfrac{1}{n} \mathrm{E}_i[\,N(n)\,]$ nach (7.3) a.)

$\longrightarrow \dfrac{f_{ih}^*}{m_{hh}}\,,\ (n \to \infty)$ nach (7.6) mit i statt ν.

b.) Ist h rekurrent, so gilt $f_{hh}^* = 1$. Aus $f_{hh}^* < 1$ folgt $m_{hh} = \infty$.

q.e.d.

Nachdem man nun die Cesaro-Konvergenz der n-Schritt-Übergangswahrscheinlichkeiten gezeigt hat, fragt man sich, ob diese an sich konvergieren. Dies ist aber nicht notwendig der Fall, weil sie periodisches Verhalten zeigen können.

<u>Definition 4:</u> Unter der <u>*Periode*</u> d_h von $h \in S$ versteht man die Größe $d_h := \mathrm{ggT}\{\, n \in \mathbb{N} : p_{hh}^{(n)} > 0 \,\}$ mit $\mathrm{ggT}\, \emptyset := \infty$. Ist $d_h = 1$, so heißt h <u>*aperiodisch*</u>.

Der periodische Fall ($d_h > 1$) läßt sich auf den aperiodischen Fall zurückführen:

<u>Lemma 7.10:</u> Für $d \in \mathbb{N}$ ist $(X_{n\cdot d})_{n \in \mathbb{N}_0}$ eine Markoff-Kette mit Übergangsmatrix $\mathbb{P}^d$. Ist $d = d_h$, so ist h aperiodisch bzgl. $\mathbb{P}^d$.

<u>Beweis:</u> Der erste Teil der Aussage folgt sofort mit (5.9). Für den zweiten betrachte man die Fälle
$d_h = 1$: Dann ist $\mathbb{P}^{d_h} \equiv \mathbb{P}$ und $(X_{n\cdot d_h})_{n \in \mathbb{N}} \equiv X$.

$d_h > 1$: $d_h \equiv \mathrm{ggT}\{ n \in \mathbb{N} : p_{hh}^{(n)} > 0 \}$. Sei $c := \mathrm{ggT}\{ n \in \mathbb{N} : p_{hh}^{(n \cdot d_h)} > 0 \}$.

<u>Annahme:</u> $c > 1 \Longrightarrow c \cdot d_h > d_h$ und $c \cdot d_h$ ist Teiler der Menge $\{ n \in \mathbb{N} : p_{hh}^{(n)} > 0 \}$. Dies ist ein Widerspruch zu d_h als ggT.

q.e.d.

Im aperiodischen Fall hat man

Lemma 7.11: Ist h aperiodisch, so gilt $p_{hh}^{(n)} > 0$ für schließlich alle n (d.h. für schließlich alle n kommt man nach einem Besuch in h mit einer positiven Wahrscheinlichkeit in n Schritten wieder nach h zurück).

Beweis: Sei $G := \{ n \in \mathbb{N}_0 : p_{hh}^{(n)} > 0 \}$. Aufgrund der Beziehung $p_{hh}^{(m+n)} \equiv \sum_{i \in S} p_{hi}^{(n)} \cdot p_{ih}^{(m)} \geq p_{hh}^{(n)} \cdot p_{hh}^{(m)}$ bildet G eine Halbgruppe bzgl. der Addition.

Es ist $d_h = \mathrm{ggT}\{G - \{0\}\} = 1$ und daher $G - \{0\}$ nach Definition 4 nichtleer. Setze $\triangle := \min\{ n - m : n > m ;\ m, n \in G \}$ und wähle $m, n \in G$ als minimale Elemente (d.h. $\triangle = n - m$). Ferner sei $n_0 \in G$ so, daß $\frac{1}{n_0} < \frac{1}{m} - \frac{1}{n}$ (z.B. $n_0 = (m+1) \cdot n$).

Man zeigt nun die Eigenschaft

$$(\star) \quad G \cap [n_0, \infty) = n_0 + \triangle \cdot \mathbb{N}_0, \text{ d.h.} \qquad (\star\star) \quad t \in G \cap [n_0, \infty) \Longrightarrow t + \triangle \in G$$

Sei also $t \in G \cap [n_0, \infty) \Longrightarrow \frac{1}{t} \leq \frac{1}{n_0} < \frac{1}{m} - \frac{1}{n}$, d.h. der Abstand von $\frac{1}{m}$ zu $\frac{1}{n}$ ist größer als $\frac{1}{t}$

$\Longrightarrow \exists\, k \in \mathbb{N}$ mit $\frac{1}{n} < \frac{k}{t} < \frac{1}{m} \Longrightarrow k \cdot m < t < k \cdot n \equiv k \cdot m + k \cdot \triangle$.

$m, n \in G \Longrightarrow (k - l) \cdot m + l \cdot n \in G \quad \forall\, 0 \leq l \leq k$, da G eine Halbgruppe ist

$\qquad \Longrightarrow k \cdot m + l \cdot \triangle \in G \quad \forall\, 0 \leq l < k$

Da $t \in G$ und $\triangle$ der minimale Abstand der Elemente aus G ist, folgt:

$\exists\, l < k$ mit $t = k \cdot m + l \cdot \triangle \Longrightarrow t + \triangle = k \cdot m + (l+1) \cdot \triangle \in G$, d.h. $(\star)$ ist gezeigt.

$n_0 \in G \Longrightarrow n_0 + G \subset G \cap [n_0, \infty) = n_0 + \triangle \cdot \mathbb{N}_0$

$\qquad \Longrightarrow G \subset \triangle \cdot \mathbb{N}_0 \Longrightarrow \triangle \leq d_h$, da $\triangle$ alle Elemente aus G teilt

$\qquad \Longrightarrow \triangle = 1$, da $d_h = 1$

Aus $(\star\star)$ folgt damit die Behauptung.

q.e.d.

Es wird nun eine Zerlegung vom Zustandsraum S in verschiedene Klassen vorgenommen, um eine Erreichbarkeitsstruktur zu erhalten:

Definition 5: Für $i, j \in S$ schreibe $i \rightsquigarrow j$ (j ist von i aus *erreichbar*), falls ein $n \in \mathbb{N}$ existiert mit der Eigenschaft $p_{ij}^{(n)} > 0$. $\mathbb{P}$ oder eine Markoff-Kette X mit Übergangsmatrix $\mathbb{P}$ heißen *irreduzibel*, falls $i \rightsquigarrow j \quad \forall\, i, j \in S,\ i \neq j$. Haben ferner alle $i \in S$ eine Eigenschaft $\mathcal{E}$, so sagt man: $\mathbb{P}$ und (oder) die Markoff-Kette X besitzen die Eigenschaft $\mathcal{E}$.

Bemerkungen: 1.) Aufgrund der Beziehung $p_{ij}^{(m+n)} \geq p_{ik}^{(n)} \cdot p_{kj}^{(m)}$ gilt:

(7.12) Die Relation " $\rightsquigarrow$ " ist transitiv.

2.) "Irreduzibel" bedeutet, daß man den Zustandsraum S nicht verkleinern kann.

Der nächste Satz beinhaltet eine erstaunliche Aussage. Es wird sich nämlich zeigen, daß im irreduziblen Fall die Charakteristika eines einzelnen Zustandes den ganzen Zustandsraum charakterisieren. Es genügt also, bei der Untersuchung auf Rekurrenz, Transienz etc. nur einen einfach handhabaren Zustand zu betrachten.

Satz 7.13: Ist $\mathbb{P}$ irreduzibel und existiert ein $i \in S$, so daß

a.) i transient bzw.

b.) i rekurrent bzw.

c.) i positiv-rekurrent bzw.

d.) i null-rekurrent bzw.

e.) i die Periode $d_i \in \mathbb{N}$ hat,

so besitzen $\mathbb{P}$ und X jeweils die entsprechenden Eigenschaften.

Beweis: Zu $i, j \in S,\ i \neq j$, wähle man $m, t \in \mathbb{N}$ mit $p_{ij}^{(m)} = \alpha > 0$ und $p_{ji}^{(t)} = \beta > 0$. Damit gilt:

$$(+) \quad p_{ii}^{(m+n+t)} \geq p_{ij}^{(m)} \cdot p_{jj}^{(n)} \cdot p_{ji}^{(t)} \equiv \alpha \cdot \beta \cdot p_{jj}^{(n)}$$

a.) Sei i transient
$$\Longrightarrow \sum_{n=1}^{\infty} p_{ii}^{(m+n+t)} < \infty \quad \text{nach (7.5)}$$
$$\Longrightarrow \sum_{n=1}^{\infty} p_{jj}^{(n)} < \infty \quad \text{nach } (+)$$
$$\Longrightarrow j \text{ ist transient nach (7.5)}$$

b.) Folgt durch Negation von a.).

d.)
$$\begin{aligned}\frac{1}{m_{ii}} &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{ii}^{(m+n+t)} \quad \text{nach (7.9)}\\ &\geq \alpha \cdot \beta \cdot \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} p_{jj}^{(n)} \quad \text{nach } (+)\\ &= \frac{\alpha \cdot \beta}{m_{jj}} \quad \text{nach (7.9)}\end{aligned}$$

c.) Folgt mit b.) durch Negation von d.).

e.) Sei $p_{jj}^{(n)} > 0 \Longrightarrow p_{jj}^{(2n)} > 0$ per Chapman-Kolmogoroff
$$\begin{aligned}&\Longrightarrow p_{ii}^{(m+2n+t)} > 0 \quad \text{mit } (+)\\ &\Longrightarrow d_i \mid (m+n+t) \text{ und } d_i \mid (m+2n+t)\\ &\Longrightarrow d_i \mid n, \text{ wo } n \text{ beliebig aus der Menge } \{\, m \in \mathbb{N} : p_{jj}^{(m)} > 0 \,\}\\ &\Longrightarrow d_i \leq d_j\end{aligned}$$

Die gleiche Betrachtung mit der Vertauschung von i und j ergibt $d_i = d_j$.

q.e.d.

Irreduzible, rekurrente Markoff-Ketten haben die Eigenschaft, daß bei einem Start in $i \in S$ jeder andere Zustand $j \in S$ fast sicher nach einer endlichen Zeit erreicht wird. Als Vorbereitung

Lemma 7.14: Ist h rekurrent und gilt $h \rightsquigarrow j$, so folgt $g_{jh} = 1$.

Beweis: $h \rightsquigarrow j \Longrightarrow \exists\, n_0 \in \mathbb{N}$ mit $p_{hj}^{(n_0)} > 0$.

$$\begin{aligned}g_{hh} &\equiv \mathrm{P}_h[\, X_m = h \text{ für unendlich viele } m\,]\\ &= \mathrm{P}_h[\, X_m = h \text{ für unendlich viele } m > n_0\,]\\ &\equiv \mathrm{E}_h\big[\mathrm{P}_h[\, X_m = h \text{ für unendlich viele } m > n_0 \mid \mathcal{F}_{n_0}\,]\big]\\ &= \mathrm{E}_h\big[\mathrm{P}_{X_{n_0}}[\, X_m = h \text{ für unendlich viele } m > 0\,]\big] \quad \text{nach (5.7)}\\ &\equiv \mathrm{E}_h[\, g_{X_{n_0}h}\,] = \sum_{i \in S} p_{hi}^{(n_0)} \cdot g_{ih}\end{aligned}$$

$$h \text{ rekurrent} \implies g_{hh} = 1 \quad \text{nach (7.4)}$$

$$\implies 0 = 1 - \sum_{i \in S} p_{hi}^{(n_0)} \cdot g_{ih} \equiv \sum_{i \in S} p_{hi}^{(n_0)} \cdot \underbrace{(1 - g_{ih})}_{\geq 0}$$

Aus $p_{hj}^{(n_0)} > 0$ folgt notwendig $g_{jh} = 1$.

q.e.d.

Nun der schon angesprochene

Satz 7.15: Ist X irreduzibel und rekurrent, so gilt: $f_{ij}^* = g_{ij} = 1 \quad \forall\, i, j \in S$

Beweis: Nach (7.14) ist $g_{ij} = 1 \quad \forall\, i, j \in S$. Aus der Definition von g_{ih} und f_{ih}^* (Definition 2) ersieht man sofort: $f_{ij}^* \geq g_{ij}$. Dies liefert die Behauptung.

q.e.d.

Der nachfolgende Satz beinhaltet eine Methode, mit der man allgemeine Fälle von Zustandsräumen oft auf den irreduziblen Fall reduzieren kann. Für das neugewonnene Modell lassen sich dann aufgrund obiger nützlicher Resultate leicht Ergebnisse herleiten, die man nur noch rückinterpretieren muß.

Satz 7.16: Für rekurrentes h definiere $S(h) := \{\, j \in S : h \sim> j \,\}$. Dann ist $h \in S(h)$ und $\mathbb{P}|_{S(h)} := (p_{ij})_{i,j \in S(h)}$ eine irreduzible Übergangsmatrix.
Insbesondere gilt $p_{ij} = 0 \quad \forall\, i \in S(h)\,,\ j \notin S(h)$.

Beweis: h rekurrent $\implies \sum_{n=1}^{\infty} p_{hh}^{(n)} = \infty$ nach (7.5) $\implies \exists\, n_0 \in \mathbb{N}$ mit $p_{hh}^{(n_0)} > 0$, d.h. $h \sim> h$.
$p_{ij} = 0 \quad \forall\, i \in S(h)\,,\ j \notin S(h)$ folgt mit Widerspruchsannahme und der Transitivität von "$\sim>$".
Die Aussage, daß $\mathbb{P}|_{S(h)}$ irreduzibel ist, folgt mit Hilfe der Eigenschaft $[\mathbb{P}|_{S(h)}]^n = \mathbb{P}^n|_{S(h)}$.

□

Es wird nun der Frage nachgegangen, wann eine Markoff-Kette X eine Folge identisch verteilter Zufallsvariablen darstellt. Zunächst gilt im Bezug auf die Verteilungsgleichheit von X_0 und X_1 mit Hilfe von (5.8′) die Äquivalenz

$$\begin{aligned} \mathrm{P}_\nu X_1^{-1} = \mathrm{P}_\nu X_0^{-1} =: \nu \iff \nu(\{\, j \,\}) &= \mathrm{P}_\nu[\, X_1 = j \,] = \sum_{k \in S} \mathrm{P}_\nu[\, X_0 = k\,,\ X_1 = j \,] \\ &= \sum_{k \in S} \nu(\{\, k \,\}) \cdot p_{kj} \quad \forall\, j \in S. \end{aligned}$$

Dies motiviert

Definition 6: Ein Wahrscheinlichkeitsmaß ν auf S heißt <u>*stationäre Verteilung*</u> (<u>*bzgl. $\mathbb{P}$*</u>), falls gilt:

$$\textbf{(7.17)} \quad \nu(\{\, j \,\}) = \sum_{i \in S} \nu(\{\, i \,\}) \cdot p_{ij} \quad \forall\, j \in S$$

Eine Charakterisierung von stationären Verteilungen liefert

Proposition 7.17′: (Balance-Gleichungen)
Ein Wahrscheinlichkeitsmaß ν auf S ist eine stationäre Verteilung

$$\iff \sum_{j \in B} \sum_{i \notin B} \nu(\{\, j \,\}) \cdot p_{ji} = \sum_{i \notin B} \sum_{j \in B} \nu(\{\, i \,\}) \cdot p_{ij} \quad \text{für } B \subset S$$

(d.h. $\mathrm{P}_\nu[\, X_0 \in B, X_1 \in B^C \,] = \mathrm{P}_\nu[\, X_0 \in B^C, X_1 \in B \,]$, der Fluß befindet sich in den Zeitpunkten 0,1 im Gleichgewicht)

Beweis: ν erfülle die Balance-Gleichungen. Mit $B := \{j\}$ gilt:

$$\nu(\{j\})\cdot(1-p_{jj}) = \sum_{i\neq j}\nu(\{i\})\cdot p_{ij} \iff \nu(\{j\}) = \sum_{i\in S}\nu(\{i\})\cdot p_{ij},$$ d.h. ν ist stationär.

Sei umgekehrt ν stationär. Für $A, B \subset S$ setze man $F(A,B) := \sum_{j\in B}\sum_{i\in A}\nu(\{j\})\cdot(p_{ij}-\delta_{ji})$.

Dann ist die linke Seite der Balance-Gleichung gleich $F(B^C, B)$ sowie die rechte gleich $F(B, B^C)$.

Weiter gelten: $\quad F(B^C,B) = F(S,B) - F(B,B)$

und $\quad F(B^C,B) = F(B,S) - F(B,B)$

Dabei ist $F(S,B) = 0$, da $\sum_{i\in S}(p_{ij}-\delta_{ji}) = 0 \quad \forall j \in S$, und $F(B,S) = 0$, da ν stationär ist.

q.e.d.

Nun die Antwort auf die oben aufgeworfene Frage nach der Verteilungsgleichheit:

Satz 7.18: Ist die Startverteilung ν der Markoff-Kette X eine stationäre Verteilung, so ist X ein stationärer Prozeß, d.h. $\mathrm{P}_\nu(X_n,\ldots,X_{n+k})^{-1}$ ist unabhängig von $n \in \mathbb{N}_0 \quad \forall k \in \mathbb{N}_0$. (Man beachte den Spezialfall $k = 0$!)

Beweis: Mit Induktion und Chapman-Kolmogoroff folgt

$$(+) \qquad \nu(\{j\}) = \sum_{i\in S}\nu(\{i\})\cdot p_{ij}^{(n)} \quad \forall j\in S,\ n\in\mathbb{N}.$$

Beweis: $$\nu(\{j\}) = \sum_{i\in S}\nu(\{i\})\cdot p_{ij} = \sum_{i\in S}\Big(\sum_{k\in S}\nu(\{k\})\cdot p_{ki}\Big)\cdot p_{ij}$$
$$= \sum_{k\in S}\nu(\{k\})\cdot\sum_{i\in S}p_{ki}\cdot p_{ij} \equiv \sum_{k\in S}\nu(\{k\})\cdot p_{kj}^{(2)} \quad \text{etc.}$$

$$\Longrightarrow \mathrm{P}_\nu[X_n = i_0,\ldots,X_{n+k} = i_k] = \mathrm{P}_\nu[X_n = i_0]\cdot p_{i_0i_1}\cdot\ldots\cdot p_{i_{k-1}i_k} \quad \text{nach (5.6)}$$
$$= \Big(\sum_{i\in S}\nu(\{i\})\cdot \mathrm{P}_i[X_n = i_0]\Big)\cdot p_{i_0i_1}\cdot\ldots\cdot p_{i_{k-1}i_k} \quad \text{nach } (5.8')$$
$$\equiv \Big(\sum_{i\in S}\nu(\{i\})\cdot p_{ii_0}^{(n)}\Big)\cdot p_{i_0i_1}\cdot\ldots\cdot p_{i_{k-1}i_k}$$
$$= \nu(\{i_0\})\cdot p_{i_0i_1}\cdot\ldots\cdot p_{i_{k-1}i_k} \quad \text{nach } (+)$$

q.e.d.

Einige spezielle Kandidaten für stationäre Verteilungen werden im nächsten Satz vorgestellt. Dazu auch eine Erweiterung der Chapman-Kolmogoroff-Gleichungen auf $\overline{\mathbb{N}}_0$:

Satz 7.19: a.) $\sum_{j\in S}p_{ij}^{(\infty)} \leq 1,\ i\in S.$

b.) Ist ν eine stationäre Verteilung, so gilt: $\nu(\{j\}) = \sum_{i\in S}\nu(\{i\})\cdot p_{ij}^{(\infty)}$

$\Big($ d.h. $\big(\nu(\{i\})\big)_{i\in S}^{\top}\cdot\mathbb{P}^\infty = \big(\nu(\{i\})\big)_{i\in S}^{\top}\Big)$

c.) $p_{ij}^{(\infty)} = \sum_{k\in S}p_{ik}^{(\infty)}\cdot p_{kj} \quad$ (d.h. $\mathbb{P}^\infty = \mathbb{P}^\infty\cdot\mathbb{P}$)

d.) $p_{ij}^{(\infty)} = \sum_{k\in S}p_{ik}\cdot p_{kj}^{(\infty)} \quad$ (d.h. $\mathbb{P}^\infty = \mathbb{P}\cdot\mathbb{P}^\infty$)

e.) $p_{ij}^{(\infty)} = \sum_{k\in S}p_{ik}^{(\infty)}\cdot p_{kj}^{(\infty)} \quad$ (d.h. $\mathbb{P}^\infty = \mathbb{P}^\infty\cdot\mathbb{P}^\infty$)

f.) Gilt für ein $i \in S$ $\sum\limits_{j \in S} p_{ij}^{(\infty)} > 0$, so ist ν_i, wo ν_i definiert durch $\nu_i(\{k\}) := \dfrac{p_{ik}{}^{(\infty)}}{\sum\limits_{j \in S} p_{ij}^{(\infty)}}$, eine stationäre Verteilung.

Ist speziell $|S| < \infty$, so ist $\nu_i(\{k\}) = p_{ik}^{(\infty)}$ und die Stationarität gilt für alle $i \in S$.

Beweis: a.) $$\sum_{j \in S} p_{ij}^{(\infty)} \overset{(7.9)}{\equiv} \sum_{j \in S} \lim_{n\to\infty} \overline{p}_{ij}^{(n)} \le \varliminf_{n\to\infty} \sum_{j \in S} \overline{p}_{ij}^{(n)} \quad \text{nach Fatou}$$
$$\equiv \varliminf_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \underbrace{\sum_{j \in S} p_{ij}^{(n)}}_{=1} = 1$$

b.) Aus (+) im Beweis von (7.18) folgt leicht: $\nu(\{j\}) = \sum\limits_{i \in S} \nu(\{i\}) \cdot \overline{p}_{ij}^{(n)} \quad \forall n \in \mathbb{N}$

$$\Longrightarrow \nu(\{j\}) = \lim_{n\to\infty} \sum_{i \in S} \nu(\{i\}) \cdot \overline{p}_{ij}^{(n)}$$
$$= \sum_{i \in S} \nu(\{i\}) \cdot \lim_{n\to\infty} \overline{p}_{ij}^{(n)} \quad \text{mit majorisierter Konvergenz}$$
$$= \sum_{i \in S} \nu(\{i\}) \cdot \overline{p}_{ij}^{(\infty)}$$

c.) $$p_{ij}^{(\infty)} \overset{(7.9)}{=} \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ij}^{(m+1)} = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sum_{k \in S} p_{ik}^{(m)} \cdot p_{kj}$$
$$\equiv \lim_{n\to\infty} \sum_{k \in S} \overline{p}_{ik}^{(n)} \cdot p_{kj} \ge \sum_{k \in S} \overline{p}_{ik}^{(\infty)} \cdot p_{kj} \quad \text{mit Fatou}$$

Hier ist eine Anwendung des Satzes über majorisierte Konvergenz nicht möglich, da $\sum\limits_{k \in S} p_{kj} = \infty$ sein kann. Mache stattdessen die Betrachtung

$$\sum_{j \in S} \underbrace{\{ p_{ij}^{(\infty)} - \sum_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj} \}}_{\ge 0 \text{ nach vorigem}} = \underbrace{\sum_{j \in S} p_{ij}^{(\infty)}}_{\substack{\le 1 \text{ nach a.), also} \\ \text{Summe zerlegbar}}} - \sum_{k \in S} p_{ik}^{(\infty)} \cdot \underbrace{\sum_{j \in S} p_{kj}}_{=1} = 0.$$

$$\Longrightarrow p_{ij}^{(\infty)} - \sum_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj} = 0 \quad \forall j \in S.$$

d.) Analog zu c.) hat man:
$$p_{ij}^{(\infty)} = \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} p_{ij}^{(m+1)} \equiv \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sum_{k \in S} p_{ik} \cdot p_{kj}^{(m)} = \lim_{n\to\infty} \sum_{k \in S} p_{ik} \cdot \overline{p}_{kj}^{(n)}$$
$$= \sum_{k \in S} p_{ik} \cdot \overline{p}_{kj}^{(\infty)} \quad \text{mit majorisierter Konvergenz.}$$

e.) Nach c.) gilt $p_{ij}^{(\infty)} = \sum\limits_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj}$

$$\Longrightarrow p_{ij}^{(\infty)} = \sum_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj}^{(n)} \quad \forall n \in \mathbb{N} \quad \text{analog zu (+) in (7.18)}$$
$$\Longrightarrow p_{ij}^{(\infty)} = \lim_{n\to\infty} \sum_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj}^{(n)} = \sum_{k \in S} p_{ik}^{(\infty)} \cdot p_{kj}^{(\infty)} \quad \text{mit majorisierter Konvergenz, da}$$

nach a.) $\sum\limits_{k \in S} p_{ik}^{(\infty)} \le 1 \quad \forall i \in S$.

f.) Mit dieser Normierung ist ν_i ein Wahrscheinlichkeitsmaß auf S. c.) besagt dann gerade, daß ν_i eine stationäre Verteilung ist. Ist $|\, S \,| < \infty$, so gilt in a.) aufgrund der Grenzwertsätze für endliche Summen $\sum_{j \in S} p_{ij}^{(\infty)} = 1 > 0 \quad \forall\, i \in S$ und damit die Behauptung.

q.e.d.

Im irreduziblen Fall tritt wieder eine Besonderheit auf, nämlich die Unabhängigkeit der $p_{ij}^{(\infty)}$ vom Startpunkt:

Lemma 7.20: Ist X irreduzibel, so gilt für $i, j \in S$: $\quad p_{ij}^{(\infty)} = p_{jj}^{(\infty)} \equiv \dfrac{1}{m_{jj}}$

Beweis: j transient $\Longrightarrow i$ transient $\quad \forall\, i \in S$ nach (7.13)

$$\Longrightarrow p_{ij}^{(\infty)} \equiv \frac{f_{ij}^*}{m_{jj}} = 0 \quad \forall\, i \in S$$

j rekurrent $\Longrightarrow i$ rekurrent $\quad \forall\, i \in S$ nach (7.13)

$$\Longrightarrow f_{ij}^* = 1 \quad \forall\, i \in S \text{ nach (7.15)}$$

q.e.d.

Folgender Satz enthält für den irreduziblen Fall ein Kriterium für die Existenz einer stationären Verteilung und charakterisiert diese:

Satz 7.21: Ist X irreduzibel, so sind äquivalent:

(i) X ist positiv rekurrent.

(ii) Es existiert eine stationäre Verteilung.

(iii) $\nu(\{\, j \,\}) := \dfrac{1}{m_{jj}}$, $j \in S$, ist die einzige stationäre Verteilung.

Beweis: Der Beweis wird geführt in den Schritten (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i).

$$\text{(iii)} \Longrightarrow \text{(i)}: \quad \sum_{j \in S} \frac{1}{m_{jj}} = 1 \Longrightarrow \exists\, i \in S \text{ mit } m_{ii} < \infty$$

$$\Longrightarrow X \text{ ist positiv rekurrent nach (7.13), d.h. (i).}$$

$$\text{(i)} \Longrightarrow \text{(ii)}: \quad X \text{ positiv rekurrent } \Longrightarrow m_{hh} < \infty \quad \forall\, h \in S$$

$$\Longrightarrow 0 < c := \sum_{j \in S} \frac{1}{m_{jj}} \leq 1 \text{ nach (7.20) und (7.19) a.)}$$

Nach (7.19) f.) existiert demnach eine stationäre Verteilung.

$$\begin{aligned}
\text{(ii)} \Longrightarrow \text{(iii)}: \quad \nu(\{\, j \,\}) &= \sum_{i \in S} \nu(\{\, i \,\}) \cdot p_{ij}^{(\infty)} \text{ nach (7.19) b.) für } \nu \text{ stationär} \\
&= \Big(\sum_{i \in S} \nu(\{\, i \,\})\Big) \cdot \frac{1}{m_{jj}} \text{ gemäß (7.20)} \\
&\equiv \frac{1}{m_{jj}}
\end{aligned}$$

q.e.d.

Nun werden die Irrfahrt auf $\mathbb{Z}$ und das M/G/1-Warteschlangenmodell (beide schon bekannt) auf ihre Eigenschaften bzgl. der in diesem Abschnitt eingeführten Begriffe hin untersucht.

Satz 7.22: Die Markoff-Kette X der Irrfahrt auf $\mathbb{Z}$ ist irreduzibel im Falle $0 < p < 1$ und hat dann die Periode 2. Im Falle $p = \dfrac{1}{2}$ ist X null-rekurrent, bei $p \neq \dfrac{1}{2}$ transient.

Beweis: 1.) $p_{i,i+1} = p > 0\,,\ p_{i,i-1} = 1-p > 0 \Longrightarrow i \sim> i \pm k \quad \forall\, k \in \mathbb{Z}$ nach (7.12).

2.) Bei einem Start in $a \in \mathbb{Z}$ gilt:

$$\left.\begin{array}{l} \mathrm{P}[\,X_{2n+1} \text{ ungerade}\,] = \mathrm{P}[\,X_{2n} \text{ gerade}\,] = 1 \quad \forall\, n \in \mathbb{N}_0\,, \text{ falls } a \text{ gerade} \\ \mathrm{P}[\,X_{2n+1} \text{ gerade}\,] = \mathrm{P}[\,X_{2n} \text{ ungerade}\,] = 1 \quad \forall\, n \in \mathbb{N}_0\,, \text{ falls } a \text{ ungerade} \end{array}\right\} \Longrightarrow d_a = 2$$

3.) Bei der Irrfahrt auf $\mathbb{Z}$ kann man nur nach einer geraden Anzahl $2n$ von Schritten zum Ausgangspunkt zurückkehren, und dies nach genau n Aufwärts- und n Abwärtsschritten. (vgl. Ruinproblem)

$$\Longrightarrow p_{hh}^{(2n-1)} = 0 \text{ und } p_{hh}^{(2n)} = \binom{2n}{n} \cdot p^n \cdot (1-p)^n \quad \forall\, n \in \mathbb{N} \quad \forall\, h \in \mathbb{Z} \quad (\text{ per Induktion })$$

$$p \neq \frac{1}{2} \Longrightarrow \sum_{n=1}^{\infty} p_{hh}^{(n)} = \sum_{n=1}^{\infty} p_{hh}^{(2n)} = \big(1 - 4p\cdot(1-p)\big)^{-\frac{1}{2}} < \infty$$

$$\Longrightarrow X \text{ ist transient nach (7.5)}$$

$$p = \frac{1}{2} \Longrightarrow \sum_{n=1}^{\infty} p_{hh}^{(n)} = \sum_{n=1}^{\infty} p_{hh}^{(2n)} = \lim_{q\uparrow\frac{1}{2}} \sum_{n\to\infty}^{\infty} \binom{2n}{n} \cdot q^n \cdot (1-q)^n$$

$$= \lim_{q\uparrow\frac{1}{2}} \big(1 - 4q\cdot(1-q)\big)^{-\frac{1}{2}} = \infty$$

$$\Longrightarrow X \text{ ist für } p = \frac{1}{2} \text{ rekurrent.}$$

4.) Es bleibt noch z.z.: Es existiert keine stationäre Verteilung.

Annahme: ν sei stationäre Verteilung.

Setze $B := \{\,\ldots, j-2, j-1, j\,\}$. Dann erhält man mit (7.17′) die Gleichung $\nu(\{\,j\,\})\cdot p = \nu(\{\,j+1\,\})\cdot(1-p)$, da $p_{hh} = 0\,,\ p_{kl} = 0 \quad \forall\, k,l \in \mathbb{Z}$ mit $|\,k-l\,| > 1$. Für $p = \frac{1}{2}$ hat man also $\nu(\{\,j\,\}) = \nu(\{\,j+1\,\})$ und damit einen Widerspruch zu $\sum_{j\in\mathbb{Z}} \nu(\{\,j\,\}) = 1$.

q.e.d.

Satz 7.23: Die Markoff-Kette X des Warteschlangenmodells M/G/1 mit Verkehrsrate ϱ ist irreduzibel und aperiodisch. Ferner gelten:

a.) Ist $\varrho > 1$, so ist X transient.

b.) Ist $\varrho = 1$, so ist X null-rekurrent.

c.) Ist $\varrho < 1$, so ist X positiv rekurrent.

Für $\varrho \geq 1$ existiert keine stationäre Verteilung und es gilt:

$$\mathrm{P}[\,X_n \geq k\,] \longrightarrow 1\,,\ (n \to \infty)\,, \quad \forall\, k \in \mathbb{N}.$$

Im Fall $\varrho < 1$ existiert genau eine stationäre Verteilung ν und es ergibt sich:

$$\mathrm{P}[\,X_n = k\,] \longrightarrow \nu(\{\,k\,\}) = \frac{1}{m_{kk}}\,,\ (n \to \infty)\,, \quad \forall\, k \in \mathbb{N}.$$

Dabei ist $\nu(\{\,0\,\}) = 1 - \varrho$ und die $\nu(\{\,k\,\})\,,\ k \in \mathbb{N}$, lassen sich mit Hilfe von (7.17) (bzw. (7.17′)) rekursiv berechnen.

Beweis: Es gilt $p_{i,i-1} \equiv b_{i-1-(i-1)^+} = b_0 > 0$ für $i > 0 \Longrightarrow i \sim> j \quad \forall\, j < i$ gemäß (7.12).

Ferner existiert wegen $b_0 + b_1 < 1$ ein $k \in \mathbb{N}$ mit $b_{k+1} > 0$

$\Longrightarrow p_{i,i+k} = p_{0,k+1} = b_{k+1} > 0 \quad \forall\, i > 0 \Longrightarrow 0 \sim> 1 + n\cdot k \quad \forall\, n \in \mathbb{N}$

$\Longrightarrow X$ ist irreduzibel.

$p_{00} \equiv p_{00}^{(1)} = b_0 > 0 \Longrightarrow d_0 = 1 \Longrightarrow X$ ist aperiodisch nach (7.13).

Die Aussagen a.), b.) und c.) folgen direkt aus (5.22) in Verbindung mit (7.13).

Daher gilt mit (7.21): Es existiert eine stationäre Verteilung ν genau dann, wenn $\varrho < 1$ ist. Diese ist dann eindeutig.

Wegen (7.29) (siehe unten) konvergiert $p_{ik}^{(n)}$ gegen $\dfrac{1}{m_{kk}}$, $(n \to \infty)$. Somit hat man mit majorisierter Konvergenz:

$$\begin{aligned} \mathrm{P}[\, X_n = k\,] &= \sum_{i \in \mathbb{N}} \mathrm{P}[\, X_0 = i\,] \cdot p_{ik}^{(n)} \quad \text{nach (5.9)} \\ &\longrightarrow \frac{1}{m_{kk}} \ \big(\, \equiv \nu(\{\, k\,\}) \text{ im Fall } \varrho < 1 \,\big)\,,\ (n \to \infty). \end{aligned}$$

Im Fall $\varrho \geq 1$ ist X transient bzw. null-rekurrent, d.h. $m_{kk} = \infty \quad \forall\, k \in \mathbb{N}$

$$\Longrightarrow \mathrm{P}[\, X_n \geq k\,] = 1 - \sum_{j<k} \mathrm{P}[\, X_n = j\,] \longrightarrow 1 - \sum_{j<k} \frac{1}{m_{jj}}\,,\ (n \to \infty)$$

$$(\ = 1 \text{ im Fall } \varrho \geq 1\ ).$$

Es ist $\dfrac{1}{m_{00}} = 1 - \varrho$ $\big($ vgl. (5.22) $\big)$, daher ist bei $\varrho < 1 \qquad \nu(\{\, 0\,\}) \equiv \dfrac{1}{m_{00}} = 1 - \varrho.$

Hat man $\nu(\{\, i\,\})$ berechnet für $i \leq j \in \mathbb{N}$, so erhält man $\nu(\{\, j+1\,\})$ durch Auflösen der Gleichung

$$\begin{aligned} \nu(\{\, j\,\}) &= \sum_{i \in \mathbb{N}} \nu(\{\, i\,\}) \cdot p_{ij} \quad \text{nach (7.17)} \\ &= \sum_{i \leq j} \nu(\{\, i\,\}) \cdot b_{j-(i-1)^+} + \nu(\{\, j+1\,\}) \cdot b_0 \\ &= \sum_{0<i\leq j} \nu(\{\, i\,\}) \cdot b_{j-1} + \nu(\{\, 0\,\}) \cdot b_j + \nu(\{\, j+1\,\}) \cdot b_0 \end{aligned}$$

oder direkt durch Anwendung von (7.17′).

q.e.d.

Zur Untersuchung der Konvergenz der $p_{ik}^{(n)}$ benötigt man den Begriff der Kopplung zweier Markoff-Ketten. Dabei werden zwei Markoff-Ketten X, X' mit gleicher Übergangsmatrix $\mathbb{P}$ und gleichem Zustandsraum S durch eine einzige Markoff-Kette ausgedrückt. Dazu das

Lemma 7.24: (Kopplung von Markoff-Ketten)

Sei $\mathbb{P}$ eine Übergangsmatrix zum Raum S. Dann ist $\widehat{\mathbb{P}} = (\widehat{p}_{(i,j),(k,l)})_{i,j,k,l \in S}$ definiert durch $\widehat{p}_{(i,j),(k,l)} := p_{ik} \cdot p_{jl}$ eine Übergangsmatrix zum Raum $S \times S$ mit der Eigenschaft $\widehat{p}^{(n)}_{(i,j),(k,l)} = p_{ik}^{(n)} \cdot p_{jl}^{(n)}$ für $i, j, k, l \in S\,,\ n \in \mathbb{N}$.

Beweis: Für $i, j \in S$ gilt: $\displaystyle\sum_{k,l \in S} \widehat{p}_{(i,j),(k,l)} \equiv \sum_{k,l \in S} p_{ik} \cdot p_{jl} = \sum_{k \in S} p_{ik} \cdot \sum_{l \in S} p_{jl} = 1.$

Ferner hat man (induktiv):

$$\begin{aligned} \widehat{p}^{(n+1)}_{(i,j),(k,l)} &\equiv \sum_{(i',j') \in S \times S} \widehat{p}^{(n)}_{(i,j),(i',j')} \cdot \widehat{p}_{(i',j'),(k,l)} = \sum_{(i',j') \in S \times S} p_{ii'}^{(n)} \cdot p_{jj'}^{(n)} \cdot p_{i'k} \cdot p_{j'l} \\ &= \sum_{i' \in S} p_{ii'}^{(n)} \cdot p_{i'k} \cdot \sum_{j' \in S} p_{jj'}^{(n)} \cdot p_{j'l} \equiv p_{ik}^{(n+1)} \cdot p_{jl}^{(n+1)} \end{aligned}$$

q.e.d.

Nun wird zunächst im irreduziblen, aperiodischen Fall gezeigt, daß die Abhängigkeit der Übergangswahrscheinlichkeiten vom Startpunkt bei fortschreitender Zeit verschwindet. Dazu vergleicht man zwei gekoppelte Markoff-Ketten, die in verschiedenen Punkten starten.

Proposition 7.25: Ist $\mathbb{P}$ irreduzibel und aperiodisch, so gilt:

$$p_{ik}^{(n)} - p_{jk}^{(n)} \longrightarrow 0\,,\ (n \to \infty) \quad \forall\, i, j, k \in S.$$

Beweis: 1.) Zunächst erweitert man (7.11) zu

(7.26) $\forall\, j, k \in S \quad \exists\, n_0 = n_0(j,k)$ mit $p_{jk}^{(n)} > 0 \quad \forall\, n \geq n_0$.

Denn nach (7.11) existiert zunächst ein $n_1 = n_1(k)$ mit $p_{kk}^{(n)} > 0 \quad \forall\, n \geq n_1$. Da $j \sim> k$ existiert ein $n_2 \in \mathbb{N}$ mit $p_{jk}^{(n_2)} > 0$. Damit hat man $p_{jk}^{(n_2+n)} \geq p_{jk}^{(n_2)} \cdot p_{kk}^{(n)} > 0 \quad \forall\, n \geq n_1$.

2.) Sei $\widehat{X} = (X, X')$ eine Markoff-Kette mit Übergangsmatrix $\widehat{\mathbb{P}}$ zum Raum $S \times S$ wie in (7.24), welche nach Ionescu-Tulcea existiert. Dann ist $\widehat{X}$ irreduzibel, denn für $i, j, k, l \in S$ hat man $\widehat{p}_{(i,j),(k,l)}^{(n)} \equiv p_{ik}^{(n)} \cdot p_{jl}^{(n)} > 0 \quad \forall\, n \geq n_0(i,k) \vee n_0(j,l)$.

3.) Ist $\widehat{X}$ transient, so hat man mit (7.5): $\widehat{p}_{(i,i),(k,k)}^{(n)} \equiv (p_{ik}^{(n)})^2 \longrightarrow 0\,,\ (n \to \infty)$

$\Longrightarrow p_{ik}^{(n)} \longrightarrow 0\,,\ (n \to \infty)$, d.h. die Behauptung.

4.) Ist $\widehat{X}$ rekurrent, dann ist nach (7.15) $\quad f_{(i,j),(k,k)}^{*} = 1 \quad \forall\, i, j, k, \in S$.

$T := \inf\{\, n \in \mathbb{N} : \widehat{X}_n = (k,k)\,\}$ werde Kopplungszeitpunkt von X, X' genannt, d.h. der Zeitpunkt, in dem X, X' in k zum ersten Mal zusammentreffen. Da $f_{(i,j),(k,k)}^{*} = 1$ (d.h. der besagte Zeitpunkt ist nach einem Start in i bzw. in j fast sicher endlich) hat man

(7.27) $\mathrm{P}_{(i,j)}[\,T > n\,] \longrightarrow 0\,,\ (n \to \infty)$.

Außerdem gelten die Gleichungen

(7.28) $p_{ik}^{(n)} = \mathrm{P}_{(i,j)}[\,X_n = k\,]\,, \quad p_{jk}^{(n)} = \mathrm{P}_{(i,j)}[\,X'_n = k\,]$.

Denn es ist $\{\,X_n = k\,\} = \bigcup\limits_{l \in S} \{\,\widehat{X}'_n = (k,l)\,\}$ und damit

$$\begin{aligned}\mathrm{P}_{(i,j)}[\,X_n = k\,] &= \sum_{l \in S} \mathrm{P}_{(i,j)}[\,\widehat{X}_n = (k,l)\,] = \sum_{l \in S} \widehat{p}_{(i,j),(k,l)}^{(n)} \equiv \sum_{l \in S} p_{ik}^{(n)} \cdot p_{jl}^{(n)} \quad \text{nach (7.24)}\\ &= p_{ik}^{(n)}\,.\end{aligned}$$

Die zweite Gleichung beweist man analog.

Mit $(i,j) := (k,k)$ erhält man daher $\quad \mathrm{P}_{(k,k)}[\,X_m = k\,] = \mathrm{P}_{(k,k)}[\,X'_m = k\,]$
und unter Herbeiziehung von (7.8) (angewandt auf $B := \{\,k\,\} \times S$):

$$\mathrm{P}_{(i,j)}[\,X_n = k, T \leq n\,] = \sum_{t=1}^{n} \mathrm{P}_{(i,j)}[\,T = t\,] \cdot \mathrm{P}_{(k,k)}[\,X_{n-t}^{(\prime)} = k\,] = \mathrm{P}_{(i,j)}[\,X'_n = k, T \leq n\,]\,.$$

Die Behauptung folgt wegen $\mathrm{P}_{(i,j)}[\,X_n^{(\prime)} = k\,] - \mathrm{P}_{(i,j)}[\,X_n^{(\prime)} = k, T \leq n\,] \overset{(7.27)}{\longrightarrow} 0\,,\ (n \to \infty)$
$\Longrightarrow \mathrm{P}_{(i,j)}[\,X'_n = k\,] - \mathrm{P}_{(i,j)}[\,X_n = k\,] \longrightarrow 0\,,\ (n \to \infty)$, nach obigem, d.h. mit (7.28):
$p_{jk}^{(n)} - p_{ik}^{(n)} \longrightarrow 0\,,\ (n \to \infty)$.

q.e.d.

Dies war die Vorstufe zum Hauptgrenzwertsatz:

Satz 7.29: Ist $\mathbb{P}$ irreduzibel und aperiodisch, so gilt: $\quad p_{ik}^{(n)} \longrightarrow \dfrac{1}{m_{kk}}\,,\ (n \to \infty), \quad \forall\, i, k \in S.$

Beweis: Da nach (7.25) die Abhängigkeit der $p_{ik}^{(n)}$ vom Startpunkt bei wachsendem n verschwindet, reicht es also o.E. aus, den Fall $i = k$ zu betrachten.

Fall 1: Ist $\mathbb{P}$ transient, so gilt nach (7.5) $\displaystyle\sum_{n=1}^{\infty} p_{kk}^{(n)} < \infty$

$\Longrightarrow p_{kk}^{(n)} \longrightarrow 0 \equiv \frac{1}{\infty}$, $(n \to \infty)$. Man beachte, daß bei Transienz $m_{kk} = \infty$ ist.

Fall 2: Ist $\mathbb{P}$ positiv rekurrent, so ist nach (7.21) $\nu(\{\,j\,\}) := \dfrac{1}{m_{jj}}$ stationäre Verteilung. Nach (+) im Beweis von (7.18) gilt ferner $\nu(\{\,k\,\}) = \displaystyle\sum_{j\in S} \nu(\{\,j\,\}) \cdot p_{jk}^{(n)}$

$\Longrightarrow p_{kk}^{(n)} - \nu(\{\,k\,\}) = \displaystyle\sum_{j\in S} \nu(\{\,j\,\}) \cdot (p_{kk}^{(n)} - p_{jk}^{(n)}) \longrightarrow 0$, $(n \to \infty)$, mit (7.25) und majorisierter Konvergenz.

Fall 3: $\mathbb{P}$ sei null-rekurrent ($\Longrightarrow m_{kk} = \infty \quad \forall\, k \in S$)

Annahme: Die Behauptung sei falsch, d.h. $\exists\, k \in S$ mit $\alpha_k := \overline{\lim\limits_{n\to\infty}}\, p_{kk}^{(n)} > 0$.

Indem man zeigt, daß eine stationäre Verteilung existiert, hat man mit (7.21) einen Widerspruch. Mit einem Diagonalverfahren ausgehend von k erhält man eine Teilfolge $(n')_{n\in\mathbb{N}}$ mit der Eigenschaft $\forall\, j \in S\ \exists\, \alpha_j$ mit $\lim\limits_{n'\to\infty} p_{kj}^{(n')} = \alpha_j$.

Wegen $\alpha_k > 0$ gilt $0 < \displaystyle\sum_{j\in S} \alpha_j$ und analog zu (7.19) a.) mit Fatou: $\displaystyle\sum_{j\in S} \alpha_j \leq 1$.

Nach (7.25) gilt sogar $\lim\limits_{n'\to\infty} p_{ij}^{(n')} = \alpha_j \quad \forall\, i, j \in S$ und

$$\begin{aligned}
\lim_{n'\to\infty} p_{kj}^{(n'+1)} &\equiv \lim_{n'\to\infty} \sum_{i\in S} p_{ki} \cdot p_{ij}^{(n')} \quad \text{(Chapman-Kolmogoroff)}\\
&= \sum_{i\in S} p_{ki} \cdot \lim_{n'\to\infty} p_{ij}^{(n')} \quad \text{(majorisierte Konvergenz)}\\
&\equiv \sum_{i\in S} p_{ki} \cdot \alpha_j \equiv \alpha_j\,.
\end{aligned}$$

Analog zu (7.19) c.) folgt aus $p_{kj}^{(n'+1)} = \displaystyle\sum_{i\in S} p_{ki}^{(n')} \cdot p_{ij}$:

$$\alpha_j \equiv \lim_{n'\to\infty} p_{kj}^{(n'+1)} = \sum_{i\in S} \lim_{n'\to\infty} p_{ki}^{(n')} \cdot p_{ij} \equiv \sum_{i\in S} \alpha_i \cdot p_{ij}\,.$$

Somit ist $\nu(\{\,j\,\}) := \dfrac{\alpha_j}{\sum\limits_{i\in S} \alpha_i}$ stationäre Verteilung.

q.e.d.

Aus diesem allgemeinen Konvergenzsatz lassen sich einige Spezialfälle über Konvergenz ableiten, so z.B.

Korollar 7.30: Ist h aperiodisch, so gilt $\mathrm{P}_\nu[\,X_n = h\,] \longrightarrow \dfrac{\mathrm{P}_\nu[\,T < \infty\,]}{m_{hh}}$, $(n \to \infty)$.

Insbesondere: $p_{ih}^{(n)} \longrightarrow p_{ih}^{(\infty)}$, $(n \to \infty)$.

Beweis: Die zweite Behauptung ist die erste mit P_i statt P_ν, die erste folgt so:

Nach (7.7) b.) gilt $\mathrm{P}_\nu[\,X_n = h\,] = \displaystyle\sum_{t=1}^{\infty} \mathrm{P}_\nu[\,T = t\,] \cdot p_{hh}^{(n-t)}$ mit $p_{hh}^{(m)} := 0$, $m < 0$.

Es ist also z.z.: $p_{hh}^{(n)} \longrightarrow \dfrac{1}{m_{hh}}$, $(n \to \infty)$.

Fall 1: Ist h transient, so verwendet man analog zum Beweis von (7.29) die Aussage (7.5) zum Beweis von $p_{hh}^{(n)} \longrightarrow \dfrac{1}{m_{hh}}$.

Fall 2: Ist h rekurrent, so betrachte man $S(h)$ $\big($ vgl. (7.16) $\big)$. Denn nach (7.16) ist $\mathbb{P}|_{S(h)}$ irreduzibel und damit liefert (7.29) die Behauptung.

q.e.d.

Bemerkung: Die Existenz von $\lim\limits_{n\to\infty} \mathrm{P}_\nu[\,X_n = h\,]$ bedeutet, daß $\mathrm{P}_\nu[\,X_n = h\,]$ mit wachsendem n nahezu unabhängig von n ist. Das stochastische System ist dann nahezu im stationären Zustand, d.h. die Verteilungen ändern sich nahezu nicht mehr.

Kapitel 2: Zeitstetige stochastische Prozesse

Bisher wurden nur stochastische Prozesse mit diskreter Zeitparametermenge betrachtet (z.B. $\mathbb{N}, \mathbb{Z}$ etc.). Diese Art von Modellen reicht jedoch nicht aus, zufällige Geschehen mit kontinuierlicher zeitlicher Entwicklung exakt beschreiben zu können. Diese Problematik führt zum Begriff des zeitstetigen stochastischen Prozesses, welcher in den nächsten Abschnitten an einigen wichtigen, praxisnahen Beispielen erläutert wird. Dabei wird aber auch immer auf die aus dem zeitdiskreten Fall her bekannten Ergebnisse zurückgegriffen werden, indem man den zeitstetigen Prozeß durch zeitdiskrete annähert und auf diese die zeitdiskreten Methoden anwendet.

Nach diesen Vorbemerkungen zu einem ersten Beispiel:

Abschnitt 8: Die Brown'sche Bewegung

Die Brown'sche Bewegung ist ein stochastisches Modell für die Bewegung eines Teilchens (Staubkorn, markiertes Molekül etc.) unter dem Einfluß chaotischer Stöße (z.B. Zusammenstößen mit anderen Molekülen eines Gases bzw. einer Flüssigkeit).

Historisch wurde dieses Phänomen 1828 durch den Botaniker Brown beschrieben, später von Bachelier (1900) im Zusammenhang mit der Fluktuation von Aktienkursen weiter untersucht. Einstein beschrieb 1905 diese Art von Bewegungen bei der Untersuchung der Bewegung eines Teilchens unter dem Einfluß von Kollisionen mit einer großen Zahl anderer Teilchen (dies führte zur Anerkennung der Realität der Atome). Das exakte mathematische Modell wurde 1923 von Wiener entwickelt.

Hier wird nun die Bewegung in einer Richtung (Dimension) untersucht, also eine eindimensionale Betrachtung durchgeführt:

Ein Prozeß $X = (X_t)_{t \geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ beschreibe eine Komponente der Bewegung eines Objekts bei fortschreitender Zeit, d.h. X_t gebe eine Komponente des Orts des Objekts im Zeitpunkt t wieder. Man diskretisiert das Problem, indem man den Prozeß in äquidistanten Zeitpunkten $0, \Delta t, 2\Delta t, \ldots$ untersucht.
Bei einem Start $X_0 = 0$ (o.E. $X_0 = 0$ durch Skalenänderung) hat man $X_{(k+1)\Delta t} = X_{k\Delta t} + Z_{k+1}^{\Delta t}$, $k \in \mathbb{N}_0$, wo $Z_{k+1}^{\Delta t}$ den Zuwachs bzw. die Änderung im Zeitintervall $\big(k\Delta t, (k+1)\Delta t\big]$ darstellt. Der Übergang $\Delta t \to 0$ liefert eine Annäherung (Approximation) an einen zeitstetigen Prozeß.

Als Modellansatz seien die $(Z_k^{\Delta t})_{k\in\mathbb{N}}$ identisch verteilt und unabhängig $\big($ was zeitliche und räumliche Homogenität ausdrücken soll, d.h. die Bedingungen (Temperatur, Druck, Konzentration etc.) sind in jedem Ort gleich und ändern sich nicht mit der Zeit $\big)$. Des weiteren sei $\mathrm{E}[\, Z_{k+1}^{\Delta t}\,] = 0 \quad \forall\, k \in \mathbb{N}_0$.

Es ist vorteilhaft, als äquidistante Unterteilung $\Delta t = \dfrac{1}{n}$ zu wählen, dies erleichtert den späteren Grenzübergang $\Delta t \to 0$. Mit dieser Wahl hat man z.B. im Zeitpunkt $t = 1$:

$\mathrm{E}[\, X_1\,] \equiv \mathrm{E}[\sum_{k=1}^{n} Z_k^{\frac{1}{n}}\,] = 0$ sowie $\mathrm{Var}\,(X_1) = n \cdot \mathrm{Var}\,(Z_1^{\frac{1}{n}}) \equiv \mathrm{Var}\,(\sqrt{n} \cdot Z_1^{\frac{1}{n}})$

Man setzt an: $Z_j^{\frac{1}{n}} =: \dfrac{1}{\sqrt{n}} \cdot Z_j$ (oder $:= Z_j^{(n)}$, wo $Z_j^{(n)}$ identisch verteilt $\forall\, j, n \in \mathbb{N}$ und bei festem n unabhängig) und erhält damit eine eindimensionale Irrfahrt $(X_{k\Delta t})_{n\geq 0}$ (ein schon aus Kapitel 1 her bekannter Begriff).

Bei der bisherigen Modellierung der Bewegung wurden nur die Zeitpunkte $(k\Delta t)_{k\in\mathbb{N}_0}$ betrachtet und die Zwischenräume vernachlässigt. Es wird in erster Näherung angenommen, daß die Bewegung zwischen $k\Delta t$ und $(k+1)\Delta t$ linear verläuft, d.h. $X_t^n := \frac{1}{\sqrt{n}}\sum_{j=1}^{k} Z_j + \frac{nt-k}{\sqrt{n}} \cdot Z_{k+1}$ für $\frac{k}{n} \le t < \frac{k+1}{n}$.

Beachte: X^n stimmt an den Stellen $t = \frac{k}{n}$, $k \in \mathbb{N}_0$, mit dem ursprünglichen Prozeß X überein.
X^n ist ein Kontrollpolygonzug zu X.

Im folgenden wird die Problematik auf den kanonischen Raum $C \equiv C[0,\infty)$ verlagert, mache dazu die

Definition 1: $C := C[0,\infty)$ sei die Menge der stetigen Funktionen von $[0,\infty)$ nach $\mathbb{R}$.
Die Abbildungen $\Pi_t : C \to \mathbb{R}$, $t \ge 0$, seien definiert durch $\Pi_t(x) := x(t)$, $x \in C$.
(Mithin wertet Π_t eine Funktion an der Stelle t aus.)

Der Übergang zum Raum C erfolgt dadurch, daß man $X^n(\omega)$ als die Abbildung $[t \mapsto X_t^n(\omega)\,,\ t \ge 0]$ auffaßt ($X^n(\omega)$ heißt auch *ω-Pfad*), d.h. $X^n : \Omega \to C$. Dabei sind jedoch folgende Probleme zu lösen:

1.) Man muß auf C eine σ-Algebra $\mathcal{L}$ finden derart, daß X^n zu einer Zufallsvariable auf $(\Omega, \mathcal{F}, \mathrm{P})$ wird.
2.) Es ist die Frage zu klären, ob X^n in Verteilung gegen ein Maß μ aus $\mathbb{P}(C)$ konvergiert (schwache Konvergenz von $\mathbb{P}(X^n)^{-1} =: \mu_n \in \mathbb{P}(C)$ gegen $\mu \in \mathbb{P}(C)$) um ein Modell für den Ursprungsprozeß zu gewinnen.

Im Bezug auf 2.) ist es wegen der Beziehung $X_t^n = \Pi_t \circ X^n$ möglich, bei der Untersuchung von $\mathrm{P}(X^n)^{-1}$ auf den Projektionsprozeß $(\Pi_t)_{t\ge 0}$ auszuweichen (analog zum kanonischen Prozeß beim Markoff-Prozeß). Es gilt nämlich $\mathrm{P}(X_{t_1}^n, \ldots, X_{t_m}^n)^{-1} = \mathrm{P}(X^n)^{-1}(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}$, $t_i \ge 0$, $m \in \mathbb{N}$.

Nach diesen Vorüberlegungen nun die Einführung einer topologischen Struktur auf dem Raum C:

Definition 2: Zu $0 < t < \infty$ und $x \in C$ sei $\| x \|_{\infty,t} := \max_{0\le s\le t} |x(s)|$ die Maximumsnorm auf $[0,t]$.
Mit deren Hilfe definiert man für $x, y \in C$: $d(x,y) := \sum_{n=1}^{\infty} \frac{1}{2^n} \wedge \| x-y \|_{\infty,n}$.
Des weiteren werde für $x \in C$ und $\epsilon > 0$ die *ε-Umgebung von x bzgl. d* gegeben durch $B(x,\epsilon) := \{ y \in C : d(x,y) < \epsilon \}$.

Bemerkung: Für eine Folge $(x_n)_{n\in\mathbb{N}_0} \subset C$ gelten die Äquivalenzen

(8.1)
$$\begin{aligned} & d(x_n, x_0) \longrightarrow 0\,,\ (n\to\infty) \\ \overset{(1)}{\Longleftrightarrow}\ & \| x_n - x_0 \|_{\infty,m} \longrightarrow 0\,,\ (n\to\infty) \quad \forall\, m \in \mathbb{N} \\ \overset{(2)}{\Longleftrightarrow}\ & x_n \longrightarrow x_0 \quad \text{gleichmäßig auf Kompakta} \end{aligned}$$

Beweis: (2) sowie (1) "$\Rightarrow$" sind klar.
Für (1) "$\Leftarrow$" verwende man ein $\frac{\epsilon}{2}$-Argument für den Reihenrest von $\sum_{j=1}^{\infty} \frac{1}{2^j}$ und $\| x-y \|_{\infty,m}$.

□

d verhilft C zu folgenden Eigenschaften:

Lemma 8.2: d ist eine Metrik auf C und C ist unter d ein polnischer Raum.
(Zur Definition eines polnischen Raumes vgl. (1.5))

Beweis: Daß d eine Metrik auf C ist, beweist man elementar. Die Vollständigkeit von C unter d folgt aus der Tatsache, daß der gleichmäßige Limes stetiger Funktionen wieder stetig ist. Die Separabilität erhält man dadurch, daß stetige Funktionen durch Polynome mit rationalen Koeffizienten (z.B. Bernstein-Polynome) auf jedem kompakten Intervall approximiert werden können.

□

Nach der topologischen Struktur nun, wie oben schon angesprochen, die Einführung einer Meßbarkeitsstruktur auf C:

Definition 3: Setze $\mathcal{L} := \sigma(\Pi_t, t \geq 0) \equiv \sigma\big(\bigcup_{t\geq 0} \Pi_t^{-1}(\mathcal{B}_1)\big)$. Mit $\mathcal{B}(C)$ werde die Borel-σ-Algebra auf C bzgl. d bezeichnet. ($\mathcal{B}_1$ ist die Borel-σ-Algebra auf $\mathbb{R}$ bzgl. euklid'scher Metrik)

Zwischen der topologischen und der Meßbarkeitsstruktur bestehen die folgenden Zusammenhänge:

Lemma 8.3: a.) Π_t ist stetig.

b.) Die Abbildung $x \mapsto d(x,y)$ ist $\mathcal{L} - \mathcal{B}_1$-meßbar $\forall\, y \in C$.

Beweis: a.) Folgt sofort aus (8.1).

b.) Die Abbildung $x \mapsto d(x,y)$ ist eine Komposition von meßbaren Abbildungen.

$\square$

Mit Hilfe dieses Lemmas hat man

Proposition 8.4: Es gilt: $\mathcal{L} = \mathcal{B}(C)$

Beweis: "$\subseteq$": Nach (8.3) a.) ist Π_t $\mathcal{B}(C) - \mathcal{B}_1$-meßbar $\forall\, t \geq 0$.
$\Longrightarrow \mathcal{L} \subset \mathcal{B}(C)$, da $\mathcal{L}$ die kleinste σ-Algebra ist, bzgl. der die Π_t meßbar sind.

"$\supseteq$": Nach (8.3) b.) ist $B(x,\epsilon) \in \mathcal{L}$ $\forall\, x \in C$, $\epsilon > 0$. Da nach (8.2) C separabel ist, ist jede offene Menge in C als abzählbare Vereinigung derartiger Kugeln darstellbar und daher aus $\mathcal{L}$.

$\square$

Im Hinblick auf obige Einführung definiert man:

Definition 4: Für $\mathrm{P} \in \mathbb{P}(C)$ und $t_1, \ldots, t_m \in \mathbb{R}_0^+$, $m \in \mathbb{N}$, wird $\mathrm{P}(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}$ als die <u>*m-dimensionale Randverteilung von P*</u> bezeichnet.

Mit diesen Randverteilungen ist es möglich, Wahrscheinlichkeitsmaße auf C zu vergleichen/charakterisieren:

Lemma 8.5: (Eindeutigkeit) Für $\mathrm{P}_1, \mathrm{P}_2 \in \mathbb{P}(C)$ sind äquivalent:

(i) $\mathrm{P}_1 = \mathrm{P}_2$

(ii) $\mathrm{P}_1(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} = \mathrm{P}_2(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}$ $\forall\, 0 \leq t_1 < t_2 < \ldots < t_m < \infty$, $m \in \mathbb{N}$.

Beweis: (i) $\Rightarrow$ (ii) ist klar.

Für (ii) $\Rightarrow$ (i) beachte man, daß $\bigcup_{\substack{0 \leq t_1 < \ldots < t_m < \infty \\ m \in \mathbb{N}}} \sigma\big((\Pi_{t_1}, \ldots, \Pi_{t_m})\big)$ ein durchschnittstabiles Erzeugendensystem von $\mathcal{L}$ ist, auf dem P_1 und P_2 übereinstimmen. Die Gleichheit auf ganz $\mathcal{L}$ folgt dann mit dem Eindeutigkeitssatz für Wahrscheinlichkeitsmaße.

$\square$

Nach diesen maßtheoretischen und topologischen Vorbereitungen ist es nun möglich, die oben motivierte eindimensionale Irrfahrt zu untersuchen:

Die Polygonprozesse

Definition 5: Sei $(Z_j)_{j \in \mathbb{N}}$ eine Folge reeller, identisch verteilter, unabhängiger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Erwartungswert 0 und Varianz 1.

Definiere für $k \in \mathbb{N}$ $\quad S_k := \sum_{j=1}^{k} Z_j$, d.h. $(S_k)_{k\in\mathbb{N}}$ stellt eine eindimensionale Irrfahrt auf $\mathbb{R}$ mit Start in $S_0 := 0$ dar. Des weiteren setze man

$$X_t^n(\omega) := \frac{1}{\sqrt{n}} \cdot [\, S_k(\omega) + (nt-k) \cdot Z_{k+1}(\omega)] \quad \text{für} \quad \frac{k}{n} \leq t < \frac{k+1}{n} \quad (\iff k = [nt]\ ).$$

Mit $X^n(\omega)$ wird der Pfad $t \mapsto X_t^n(\omega)$ bezeichnet.

Die Meßbarkeit von X^n zeigt

Lemma 8.6: a.) $X : \Omega \to C$ ist $\mathcal{F}-\mathcal{L}$-meßbar $\iff \Pi_t \circ X$ ist $\mathcal{F}-\mathcal{B}_1$-meßbar $\quad \forall\, t \geq 0$.

b.) $X^n : \Omega \to C$ ist $\mathcal{F}-\mathcal{L}$-meßbar $\quad \forall\, n \in \mathbb{N}$ $\big(\implies \mathrm{P}(X^n)^{-1} \in \mathbb{P}(C) \quad \forall\, n \in \mathbb{N} \big)$.

Beweis: a.) "$\Rightarrow$": Π_t ist $\mathcal{L}-\mathcal{B}_1$-meßbar $\quad \forall\, t \geq 0 \implies \Pi_t \circ X$ ist $\mathcal{F}-\mathcal{B}_1$-meßbar $\quad \forall\, t \geq 0$.

"$\Leftarrow$": $X^{-1}\Pi_t^{-1}(\mathcal{B}_1) \subset \mathcal{F} \quad \forall\, t \geq 0 \implies X^{-1}\big(\bigcup_{t\geq 0} \Pi_t^{-1}(\mathcal{B}_1)\big) \subset \mathcal{F}$.

b.) $\Pi_t \circ X^n = X_t^n$, somit ist X^n nach a.) $\mathcal{F}-\mathcal{L}$-meßbar.

q.e.d.

Als Vorbereitung für den Nachweis der oben in Punkt 2.) angesprochenen schwachen Konvergenz von $\mathrm{P}(X^n)^{-1}$ nun zunächst

Konvergenz der endlich dimensionalen Verteilung

Nach (8.3) war Π_t stetig $\quad \forall\, t \geq 0$, folglich $(\Pi_{t_1}, \ldots, \Pi_{t_m})$ stetig $\quad \forall\, t_i \geq 0$, $m \in \mathbb{N}$. Daher ist es nach W-Theorie 17.8 für die schwache Konvergenz von $\mathrm{P}(X^n)^{-1}$ gegen ein $\mu \in \mathbb{P}(C)$ notwendig, daß gilt:

(+) $\quad \mathrm{P}(X^n)^{-1}(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} \equiv \mathrm{P}(X_{t_1}^n, \ldots, X_{t_m}^{-1})^{-1} \xrightarrow{w} \mu(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}$, $(n \to \infty)$

Es folgen nun zwei Lemmata, die für den Beweis des Grenzwertsatzes für die endlich dimensionale Randverteilung unentbehrlich sind:

Lemma 8.7: Seien $(\mu_n^{(j)})_{n\in\overline{\mathbb{N}}}$, $j = 1, \ldots, m \in \mathbb{N}$, Folgen aus $\mathbb{P}(\mathbb{R})$. Dann sind äquivalent:

(i) $\mu_n^{(j)} \xrightarrow{w} \mu_\infty^{(j)}$, $(n \to \infty)$, $\forall\, j \in \{1, \ldots, m\}$

(ii) $\bigtimes_{j=1}^{m} \mu_n^{(j)} \xrightarrow{w} \bigtimes_{j=1}^{m} \mu_\infty^{(j)}$, $(n \to \infty)$.

Beweis: (i) $\Rightarrow$ (ii): Nach dem Stetigkeitssatz von Lévy-Cramèr gelten für die charakteristischen Funktionen $\varphi_n^{(j)}$ von $\mu_n^{(j)}$: $\quad \varphi_n^{(j)}(t) \longrightarrow \varphi_\infty^{(j)}(t)$, $(n \to \infty)$, $\forall\, t$, $1 \leq j \leq m$ (vergleiche W-Theorie 18.11)

$$\iff \prod_{j=1}^{m} \varphi_n^{(j)}(t_j) \longrightarrow \prod_{j=1}^{m} \varphi_\infty^{(j)}(t_j)\,,\ (n \to \infty)\,,\ \forall\, t_j\,,\ 1 \leq j \leq m$$

d.h. die charakteristischen Funktionen von $\bigtimes_{j=1}^{m} \mu_n^{(j)}$ konvergieren gegen die von $\bigtimes_{j=1}^{m} \mu_\infty^{(j)}$ $\big($ dabei gilt "$\Leftarrow$" wegen $\varphi_n^{(j)}(0) = 1$ $\big)$.

Dies liefert nach Lévy-Cramèr gerade die Behauptung.

Die Richtung (ii) $\Rightarrow$ (i) ist klar.

$\square$

Eine wichtige Eigenschaft der mehrdimensionalen Normalverteilung zeigt

Lemma 8.8: Seien $t_0 = 0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$, und $(W_i)_{0\leq i\leq m}$ reelle Zufallsvariablen. Dann sind äquivalent:

(i) $W_i - W_{i-1}$ sind unabhängig, $W_i - W_{i-1} \sim \mathrm{N}(0, t_i - t_{i-1})$, $i \in \{1, \ldots, m\}$, $W_0 := 0$

(ii) $(W_1, \ldots, W_m) \sim \mathrm{N}(0, \Gamma)$ mit $\Gamma := (t_j \wedge t_i)_{1\leq i,j\leq m}$

Beweis: Der Beweis ist relativ einfach zu führen, man verwendet dabei nur den die mehrdimensionale Normalverteilung charakterisierenden Satz 16.6 der W-Theorie sowie deren Definition.

□

Nun die schon mehrfach angesprochene

Proposition 8.9: Für $\_0 = 0 \leq t_1 < t_2 < \ldots < t_m < \infty$, $m \in \mathbb{N}$, gilt mit $\Gamma := (t_i \wedge t_j)_{1 \leq i,j \leq m}$:

$$P(X^n_{t_1}, \ldots, X^n_{t_m})^{-1} \xrightarrow{w} N(0,\Gamma)\,,\ (n \to \infty).$$

Bemerkung: Im Fall $m = 1$ hat man insbesondere: X^n_t konvergiert in Verteilung gegen $N(0,t)$, d.h. für $t = 1$: $X^n_1 \equiv \frac{S_n}{\sqrt{n}} \longrightarrow N(0,1)$ in Verteilung, $(n \to \infty)$.

Somit enthält (8.9) den zentralen Grenzwertsatz und verallgemeinert diesen dadurch.

Beweis: O.E. sei $t_1 > 0$. Ist nämlich $t_1 = 0$, so ist $X^n_{t_1} \equiv X^n_0 = 0$. Andererseits ist in (8.8) mit $t_1 = 0$ auch $W_1 = \underline{0}$ und folgende Betrachtungen können mit $(t_2, \ldots, t_m)$ getätigt werden.

Zur Vereinfachung zeigt man die Behauptung zunächst für den stückweise konstanten Prozeß $\frac{1}{\sqrt{n}} \cdot S_{[n \cdot t]}$ anstelle von X^n_t $\quad (X^n_{\frac{[n \cdot t]}{n}} = \frac{1}{\sqrt{n}} \cdot S_{[n \cdot t]}$).

Um dabei die Form von (8.8) zu erreichen, ist es günstiger, die Zuwächse

$$Y^n_i := \frac{1}{\sqrt{n}} \cdot (S_{[n \cdot t_i]} - S_{[n \cdot t_{i-1}]}) \equiv \frac{1}{\sqrt{n}} \sum_{j=[n \cdot t_{i-1}]}^{[n \cdot t_i]+1} Z_j\ ,\ i \in \{1, \ldots, m\}\text{, zu untersuchen.}$$

Für festes n sind diese aufgrund der Unabhängigkeit der $(Z_k)_{k \in \mathbb{N}}$ ebenfalls unabhängig.
Im Hinblick auf (8.7) sei jetzt i fest gewählt. Da die Z_k identisch verteilt sind, ist Y^n_i verteilungsgleich mit $\frac{S_{k(n)}}{\sqrt{n}}$, wo $k(n)$ definiert durch $k(n) := [n \cdot t_i] - [n \cdot t_{i-1}]$.

Wegen $k(n) \longrightarrow \infty$ für $n \to \infty$ folgt aus dem zentralen Grenzwertsatz $\frac{S_{k(n)}}{\sqrt{k(n)}} \longrightarrow N(0,1)$ in Verteilung . Mit $\frac{k(n)}{n} \longrightarrow t_i - t_{i-1}$, $(n \to \infty)$, und W-Theorie 17.21/17.22 (Cramèr) ergibt sich daher:

$$\frac{S_{k(n)}}{\sqrt{n}} \equiv \sqrt{\frac{k(n)}{n}} \cdot \frac{S_{k(n)}}{\sqrt{k(n)}} \longrightarrow N(0, t_i - t_{i-1}) \text{ in Verteilung, } (n \to \infty).$$

Somit folgt aus (8.7) $\quad (Y^n_1, \ldots, Y^n_m) \longrightarrow \bigtimes_{j=1}^{m} N(0, t_i - t_{i-1}) =: \nu$ in Verteilung, $(n \to \infty)$, da $P(Y^n_1, \ldots, Y^n_m)^{-1} = \bigtimes_{j=1}^{m} P(Y^n_j)^{-1}$ aufgrund der Unabhängigkeit.

Definiere nun $h : \mathbb{R}^m \to \mathbb{R}^m$ durch $h(y_1, \ldots, y_m) := (y_1, y_1 + y_2, \ldots, y_1 + \ldots + y_m)$.
Da h stetig und $h(Y^n_1, \ldots, Y^n_m) = \frac{1}{\sqrt{n}} \cdot (S_{[n \cdot t_1]}, \ldots, S_{[n \cdot t_m]})$ gilt nach W-Theorie 17.8:

$$\frac{1}{\sqrt{n}} \cdot (S_{[n \cdot t_1]}, \ldots, S_{[n \cdot t_m]}) \longrightarrow \nu h^{-1} \text{ in Verteilung, } (n \to \infty).$$

Da aber ferner unter Verwendung von (8.8) $\quad h(Y^n_1, \ldots, Y^n_m) \longrightarrow N(0,\Gamma)$ in Verteilung, $(n \to \infty)$, ergibt sich notwendigerweise $\quad \nu h^{-1} = N(0,\Gamma)$. $\big($ Wende dabei (8.8) auf eine Folge $(V_i)_{0 \leq i \leq m}$ unabhängiger Zufallsvariablen mit $V_i \sim N(0, t_i - t_{i-1})$ an $\big)$.

Betrachte nun wieder den ursprünglichen Prozeß X^n_t:

Es ist $\left|\frac{S_{[n \cdot t]}}{\sqrt{n}} - X^n_t\right| \equiv |(n \cdot t - [n \cdot t]) \cdot \frac{Z_{[n \cdot t]+1}}{\sqrt{n}}| \leq \frac{|Z_{[n \cdot t]+1}|}{\sqrt{n}}$ und somit für $\epsilon > 0$:

$P[\,|\frac{S_{[n \cdot t]}}{\sqrt{n}} - X^n_t| > \epsilon\,] \leq P[\,\frac{|Z_1|}{\sqrt{n}} > \epsilon\,] \downarrow P[\,|Z_1| = \infty\,] = 0$, $(n \to \infty)$, d.h. $\frac{S_{[n \cdot t]}}{\sqrt{n}} - X^n_t \longrightarrow 0$ nach Wahrscheinlichkeit, $(n \to \infty)$. (und damit auch in Verteilung, vgl. W-Theorie 17.18)

$\Longrightarrow (\frac{S_{[n \cdot t_i]}}{\sqrt{n}} - X^n_{t_i})_{1 \leq i \leq m} \longrightarrow 0$ nach Wahrscheinlichkeit, $(n \to \infty)$.

(dies folgt leicht mit einem $\frac{\varepsilon}{m}$-Argument)

$\Longrightarrow (X^n_{t_1}, \ldots, X^n_{t_m}) \longrightarrow \mathrm{N}(0, \Gamma)$ in Verteilung, $(n \to \infty)$, nach W-Theorie 17.20

q.e.d.

Nun wird für die Konvergenz von $\mathrm{P}(X^n)^{-1}$ ein notwendiges und hinreichendes Kriterium für die Existenz eines schwachen Grenzwertes einer Folge von Wahrscheinlichkeitsmaßen auf C erarbeitet:

Proposition 8.10: Sei $(\mu_n)_{n \in \mathbb{N}}$ eine Folge von Elementen aus $\mathbb{P}(C)$ und $\mu_{t_1,\ldots,t_m} \in \mathbb{P}(\mathbb{R}^m)$ für $0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$. Dann sind äquivalent:

(i) Es existiert ein $\mu_\infty \in \mathbb{P}(C)$ mit den beiden Eigenschaften

a.) $\mu_n \xrightarrow{w} \mu_\infty$, $(n \to \infty)$

b.) $\mu_\infty(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} = \mu_{t_1,\ldots,t_m} \quad \forall\, 0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$

(ii) a.) $\mu_n(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} \xrightarrow{w} \mu_{t_1,\ldots,t_m} \quad \forall\, 0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$

b.) $\{\mu_n\}_{n \in \mathbb{N}}$ ist relativ folgenkompakt, d.h. für alle Teilfolgen $\{\mu_{n'}\}_{n \in \mathbb{N}}$ existiert eine Teilfolge $(n'')_{n \in \mathbb{N}} \subset (n')_{n \in \mathbb{N}}$ und ein $\mu \in \mathbb{P}(C)$ mit der Eigenschaft $\mu_{n''} \xrightarrow{w} \mu$, $(n'' \to \infty)$.

Beweis: (i) $\Rightarrow$ (ii): (ii) a.) folgt sofort aus der Bemerkung (+) vor (8.7). Aus (i) a.) folgt direkt (ii) b.), da mit μ_n auch alle Teilfolgen $\mu_{n''}$ gegen μ_∞ in Verteilung konvergieren.

(ii) $\Rightarrow$ (i): (vgl. dazu auch den Beweis des Stetigkeitssatzes von Cramèr)
Nach (ii) b.) existiert eine Teilfolge $(n')_{n \in \mathbb{N}}$ und ein $\mu_\infty \in \mathbb{P}(C)$ mit der Eigenschaft $\mu_{n'} \xrightarrow{w} \mu_\infty$, $(n' \to \infty)$. Mit (+) gilt daher $\forall\, 0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$:

$$\begin{aligned} \mu_{n'}(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} &\xrightarrow{w} \mu_\infty(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}, \ (n' \to \infty) \\ &\equiv \mu_{t_1,\ldots,t_m} \quad \text{nach (ii) a.)} \end{aligned}$$

d.h. man erhält (i) b.).

Für (i) a.) ist nun noch die Unabhängigkeit von μ_∞ von der Wahl der Teilfolge n' zu zeigen. Sei dazu $(m')_{m \in \mathbb{N}} \subset \mathbb{N}$ beliebig. Nach W-Theorie 18.10 ist nur die Existenz einer Teilfolge $(m'')_{m \in \mathbb{N}} \subset (m')_{m \in \mathbb{N}}$ nachzuweisen derart, daß $\mu_{m''} \xrightarrow{w} \mu_\infty$ für $m'' \to \infty$. (ii) b.) liefert $(m'')_{m \in \mathbb{N}} \subset (m')_{m \in \mathbb{N}}$ und ein $\mu \in \mathbb{P}(C)$ mit der Eigenschaft $\mu_{m''} \xrightarrow{w} \mu$, $(m'' \to \infty)$.

Daraus folgt wieder mit (+) für alle $0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$:

$$\begin{aligned} \mu_{m''}(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} &\xrightarrow{w} \mu(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1}, \ (m'' \to \infty) \\ &= \mu_{t_1,\ldots,t_m} \quad \text{nach (ii) a.)} \\ &= \mu_\infty(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} \quad \text{nach obiger Rechnung} \end{aligned}$$

Mit (8.5) folgt daher $\mu = \mu_\infty$.

q.e.d.

Man nähert sich nun schnell der Lösung der Konvergenzfrage der $\mathrm{P}(X^n)^{-1}$, denn es zeigt sich, daß $\{\mathrm{P}(X^n)^{-1} : n \in \mathbb{N}\}$ die Voraussetzung (ii) b.) aus (8.10) erfüllt:

Proposition 8.11: Die Menge $\{\mathrm{P}(X^n)^{-1} : n \in \mathbb{N}\}$ ist relativ folgenkompakt in $\mathbb{P}(C)$.

Beweis: Die Folgenkompaktheit wird sich in Abschnitt 9 als äquivalent zur Straffheit erweisen, der Beweis wird dort geführt.

□

Nun zu den beiden Hauptsätzen dieses Abschnitts, die die Konvergenzfrage endgültig klären:

Satz 8.12: Es existiert genau ein Wahrscheinlichkeitsmaß μ auf $(C, \mathcal{L})$ mit $\mu(\Pi_{t_1}, \ldots, \Pi_{t_m})^{-1} = \mathrm{N}(0, \Gamma)$, wo $\Gamma = (t_i \wedge t_j)_{1 \leq i,j \leq m}$, $0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$.

Beweis: Zum Existenznachweis setze man in (8.10) für $n \in \mathbb{N}$ $\mu_n := \mathrm{P}(X^n)^{-1}$ sowie für $0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$, $\mu_{t_1,\ldots,t_m} := \mathrm{N}(0,\Gamma)$, wobei Γ wie oben definiert sei. Wegen $\mathrm{P}(X^n_{t_1},\ldots,X^n_{t_m})^{-1} = \mathrm{P}(X^n)^{-1}(\Pi_{t_1},\ldots,\Pi_{t_m})^{-1}$ hat man, da nach (8.9) $\mathrm{P}(X^n_{t_1},\ldots,X^n_{t_m})^{-1} \xrightarrow{w} \mathrm{N}(0,\Gamma)$, die Eigenschaft $\mu_n(\Pi_{t_1},\ldots,\Pi_{t_m})^{-1} \xrightarrow{w} \mu_{t_1,\ldots,t_m}$, $(n \to \infty)$. Aufgrund von (8.11) ist $\{\, \mathrm{P}(X^n)^{-1} : n \in \mathbb{N} \,\}$ folgenkompakt, d.h. die Voraussetzungen von (ii) in (8.10) sind erfüllt. Daher existiert ein $\mu \in \mathbb{P}(C)$ mit der gewünschten Eigenschaft, welches gemäß (8.5) eindeutig ist.

q.e.d.

Dieses Maß hat eine fundamentale Bedeutung:

Definition 6: Das Maß aus (8.12) heißt *Wiener-Maß* auf $C[\,0,\infty)$. μ ist neben dem Lebesgue-Maß das wichtigste Maß.

Betrachtet man noch einmal den Beweis von (8.12), so stellt man fest, daß man das Ziel "Konvergenz der X^n in Verteilung" schon erreicht hat, es gilt nämlich

Satz 8.13: (Donsker) X^n konvergiert in Verteilung gegen das Wienermaß μ, $(n \to \infty)$.

Beweis: Im Beweis von (8.12) wurde schon gezeigt, daß die Menge $\{\, \mathrm{P}(X^n)^{-1} : n \in \mathbb{N} \,\}$ die Voraussetzungen von (8.10) (ii) erfüllt, somit liefert (8.10) (i) die Behauptung zusammen mit dem Eindeutigkeitssatz (8.5).

q.e.d.

Es wird nun eine exakte Definition des Begriffs "Brown'sche Bewegung" angegeben, welche im nachhinein die bisherige Vorgehensweise rechtfertigt. Zuvor jedoch noch

Definition 7: Ein stochastischer Prozeß $(X_t)_{t\geq 0}$ mit Zustandsraum $\mathbb{R}^d$ hat

(i) *unabhängige Zuwächse*, falls die Zufallsvariablen $X_0, X_{t_1} - X_0, \ldots, X_{t_m} - X_{t_{m-1}}$ unabhängig sind für alle $0 \leq t_1 < \ldots < t_m < \infty$, $m \in \mathbb{N}$.

(ii) *stationäre Zuwächse*, falls $X_t - X_s$ verteilungsgleich ist mit $X_{t-s} - X_0$ für alle $0 \leq s \leq t < \infty$.

Nun die angekündigte

Definition 8: Ein stochastischer Prozeß $(W_t)_{t\geq 0}$ mit Zustandsraum $\mathbb{R}^d$ heißt *d-dimensionale standardisierte Brown'sche Bewegung* (im Fall $d = 1$ auch *Wienerprozeß*), falls gilt:

(0) $(W_t)_{t\geq 0}$ hat stetige Pfade

(i) $W_0 = 0$ fast sicher

(ii) $(W_t)_{t\geq 0}$ hat stationäre unabhängige Zuwächse

(iii) $W_t - W_s \sim \bigtimes_{i=1}^{d} \mathrm{N}(0, t-s)$, $0 \leq s \leq t < \infty$

Bemerkung: Insbesondere gilt für $W_t \equiv (W_t^{(1)},\ldots,W_t^{(d)})$: $W_t^{(1)},\ldots,W_t^{(d)}$ sind unabhängig (W-Theorie 13.1 a'.)) sowie $W_t^{(i)} \sim \mathrm{N}(0,t)$.

Mit Hilfe des Projektionsprozesses und des Wienermaßes hat man einen kanonischen Wienerprozeß gefunden:

Korollar 8.14: Es sind äquivalent: (i) μ ist Wienermaß auf $(C,\mathcal{L})$

(ii) $(\Pi_t)_{t\geq 0}$ ist ein Wienerprozeß auf $(C,\mathcal{L},\mu)$

Beweis: leicht mit (8.12) sowie (8.8).

□

Bemerkung: X^n war als Annäherung an den ursprünglichen Bewegungsprozeß X konzipiert worden. Nun zeigen (8.13) und (8.14), daß die Verteilung des Pfades $t \mapsto X_t^n(\omega)$ von X^n beliebig gut an die Verteilung eines Wienerprozesses $\big($ hier $(\Pi_t)_{t\geq 0}$ $\big)$ angenähert werden kann. Man erhält damit auch ein Modell für den ursprünglichen Bewegungsprozesses X.

Mit (8.13) erhält man eine *funktionale Variante* des zentralen Grenzwertsatzes:

Man hat nicht nur, wie in (8.9) angesprochen, $\dfrac{S_n}{\sqrt{n}} \equiv X_1^n \longrightarrow \mathrm{N}(0,1)$ in Verteilung, $(n \to \infty)$, sondern auch

die Aussage: Ist $h : C \to \mathbb{R}$ (μ-fast überall) stetig, so gilt $h(X^n) \longrightarrow \mu h^{-1}$ in Verteilung, $(n \to \infty)$, wo μ das Wienermaß ist (W-Theorie 17.8).

Dazu ein

Beispiel: Definiere $h : C \to \mathbb{R}$ durch $h(x) := \max\limits_{0\leq t\leq 1} x(t)$. Dann ist h stetig.

Beweis: $(x_n)_{n\in\mathbb{N}} \subset C$ mit $d(x_n, x) \longrightarrow 0$, $(n \to \infty)$ für ein $x \in C$

$\Longrightarrow \| x_n - x \|_{\infty,1} \longrightarrow 0$, $(n \to \infty)$

Ferner gilt $h(X^n) = \max\limits_{0\leq k\leq n} \dfrac{S_k}{\sqrt{n}}$, da X^n ein Polygonprozeß ist. Für $A \in \mathcal{B}_1$ hat man

$\mu h^{-1}[A] = \mu[x : \max\limits_{0\leq t\leq 1} x(t) \in A] = \mu(\max\limits_{0\leq t\leq 1} \Pi_t)^{-1}[A]$. Da $\mathrm{P}(X^n)^{-1} \xrightarrow{w} \mu$, $(n \to \infty)$,

$\big($ und somit $\mathrm{P}(X^n)^{-1}h^{-1} \xrightarrow{w} \mu h^{-1}$ $\big)$ gilt $\max\limits_{0\leq k\leq n} \dfrac{S_k}{\sqrt{n}} \longrightarrow \max\limits_{0\leq t\leq 1} W_t$ in Verteilung, $(n \to \infty)$,

wobei W ein beliebiger Wienerprozeß ist $\big($ da unter μ $(\Pi_t)_{t\geq 0}$ Wienerprozeß $\big)$.

Es wird nun mit Hilfe eines *Spiegelungsprinzips* gezeigt: $\max\limits_{0\leq t\leq 1} W_t$ ist verteilungsgleich mit $|W_1|$, wobei bekanntlich $W_1 \sim \mathrm{N}(0,1)$.

Das Spiegelungsprinzip kann schon auf der Ebene der Irrfahrten eingesetzt werden, indem man das (8.13) innewohnende *Invarianzprinzip* ausnutzt. Man erinnere sich, daß der schwache Limes der $\mathrm{P}(X^n)^{-1}$ nicht von der Verteilung der Z_n abhängt, sondern nur von den beiden Voraussetzungen $\mathrm{E}[Z_n] = 0$ sowie $\mathrm{Var}(Z_n) = 1$. Daher wähle man der Einfachheit halber die Z_n so, daß $\mathrm{P}[Z_n = \pm 1] = \dfrac{1}{2}$ ist, mithin $(S_n)_{n\in\mathbb{N}_0}$ eine symmetrische Irrfahrt auf $\mathbb{Z}$ darstellt.

Mit dem Spiegelungsprinzip kann man mit etwas Aufwand zeigen (Übung !):

$\mathrm{P}[\max\limits_{0\leq k\leq n} S_k < b] = \mathrm{P}[-b \leq S_n < b]$ für $b \in \mathbb{N}$

$$
\begin{aligned}
\Longrightarrow \mathrm{P}[\max_{0\leq k\leq n} S_k \leq t\cdot\sqrt{n}] &= \mathrm{P}\big[\max_{0\leq k\leq n} S_k < [t\cdot\sqrt{n}]+1\big] \quad \text{da } S_k \in \mathbb{Z} \\
&\equiv \mathrm{P}\big[-[t\cdot\sqrt{n}]-1 \leq S_n < [t\cdot\sqrt{n}]+1\big] \quad \text{nach obigem} \\
&= \mathrm{P}[S_n \leq t\cdot\sqrt{n}] - \mathrm{P}\big[S_n < -[t\cdot\sqrt{n}]-1\big] \\
&= \mathrm{P}[\frac{S_n}{\sqrt{n}} \leq t] - \mathrm{P}[S_n \leq -t\cdot\sqrt{n}-2] \\
&= \mathrm{P}[\frac{S_n}{\sqrt{n}} \leq t] - \mathrm{P}[\frac{S_n}{\sqrt{n}} + \frac{2}{\sqrt{n}} \leq -t] \\
&\longrightarrow \mathrm{P}'[W_1 \leq t] - \mathrm{P}'[W_1 \leq -t] \quad \text{für } n \to \infty \\
&\equiv \mathrm{P}'[|W_1| \leq t]
\end{aligned}
$$

Da nach obigem $\max\limits_{0\leq k\leq n} \dfrac{S_k}{\sqrt{n}} \longrightarrow \max\limits_{0\leq t\leq 1} W_t$ in Verteilung, $(n \to \infty)$, hat man als Folgerung:

$\max\limits_{0\leq t\leq 1} W_t$ ist verteilungsgleich mit $|W_1|$.

Soweit zunächst die Erörterung der Brown'schen Bewegung. Aus diesem Abschnitt steht noch die wichtige Frage nach der Folgenkompaktheit der Familie $\{ P(X^n)^{-1} : n \in \mathbb{N} \}$ im Raume. Zu deren Beantwortung dient der nächste Abschnitt:

Abschnitt 9: Straffheit auf C[0,∞)

Zur Erinnerung: Ist (S,d) ein metrischer Raum mit der Borel-σ-Algebra $\mathcal{B}(S)$ versehen, so gilt:
Eine Menge $\Gamma \subset \mathbb{P}(S)$ heißt straff $\iff \forall \epsilon > 0$ existiert eine kompakte Menge $K \subset S$ mit $P[K] \geq 1-\epsilon \quad \forall P \in \Gamma$

Für das Hauptziel dieses Abschnittes, nämlich den Beweis von (8.11), wird nun der zentrale Satz hergeleitet:

Satz von Prohoroff

In der W-Theorie (Satz 30.7) wurde schon der Spezialfall $S = \mathbb{R}^d$ bewiesen. Im Hinblick auf (8.11) und den dort zugrundeliegenden metrischen Raum ist es jedoch notwendig, das schon bekannte Resultat für den Fall eines polnischen Raumes herzuleiten. Zuvor noch zwei vorbereitende Lemmata und eine Vorüberlegung:

Lemma 9.1: Seien S, S' metrische Räume und $h : S \to S'$ eine stetige Funktion (bzgl. der durch die Metriken erzeugten Topologien). Dann gilt:
Ist $\Gamma \subset \mathbb{P}(S)$ straff, so auch $\Gamma h^{-1} := \{ Ph^{-1} : P \in \Gamma \}$.

Beweis: Das stetige Bild einer kompakten Menge ist wieder kompakt. Somit ist mit K in S auch $h(K)$ in S' wieder kompakt. Ferner gilt $K \subset h^{-1}\big(h(K)\big)$, d.h. $Ph^{-1}[h(K)] \geq P[K]$

q.e.d.

Vorüberlegung zu (9.2) und (9.3):

Der Raum $\mathbb{R}^k$ werde mit der Metrik $d_k(x,y) := \sum_{i=1}^{k} \frac{1}{2^i} \wedge |x^i - y^i|$, $k \in \overline{\mathbb{N}}$, welche die übliche Topologie $\mathcal{O}_k$ auf $\mathbb{R}^k$ induziert. Diese Topologie hat als Basis das System $\{ B^k(x,\epsilon) : x \in \mathbb{R}^k,\ \epsilon > 0 \}$, wo $B^k(x,\epsilon)$ definiert durch $B^k(x,\epsilon) := \{ y \in \mathbb{R}^k : d_k(x,y) < \epsilon \}$. Bezeichnet jeweils $\mathcal{F}_i = \mathcal{B}_1$ die σ-Algebra der borelschen Mengen von $\mathbb{R}$ bzgl. d_1, so gilt: $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$ ist die σ-Algebra der borelschen Mengen von $\mathbb{R}^\infty$ bzgl. d_∞.
Dies besteht, da $\{ A_1 \times \ldots \times A_k \times \mathbb{R} \times \ldots : A_i \in \mathcal{O}_1,\ 1 \leq i \leq k,\ k \in \mathbb{N} \}$ eine Basis der von d_∞ induzierten Topologie ist. Nach (2.1) erzeugt $\{ A_1 \times \ldots \times A_k \times \mathbb{R} \times \ldots : A_i \in \mathcal{F}_1 \equiv \sigma(\mathcal{O}_1),\ 1 \leq i \leq k,\ k \in \mathbb{N} \}$ die σ-Algebra $\bigotimes_{i=1}^{\infty} \mathcal{F}_i$. Daraus folgt die Gleichheit $\bigotimes_{i=1}^{\infty} \mathcal{F}_i = \sigma(\mathcal{O}_\infty)$ (vgl. auch Abschnitt 8).

Nun zu

Lemma 9.2: Seien für $k \in \mathbb{N}$ $\Pi_k \equiv (\pi_1, \ldots, \pi_k) : \mathbb{R}^\infty \to \mathbb{R}^k$ die Projektionen auf die ersten k Komponenten. Dann gilt für eine Folge $(P_n)_{n \in \overline{\mathbb{N}}} \subset \mathbb{P}(\mathbb{R}^\infty)$:

$$P_n \xrightarrow{w} P_\infty \iff P_n \Pi_k^{-1} \xrightarrow{w} P_\infty \Pi_k^{-1},\ (n \to \infty),\ \forall k \in \mathbb{N}$$

Beweis: $\mathbb{R}^k$ sei wieder mit der Metrik $\sum_{i=1}^{k} \frac{1}{2^i} \wedge |x^i - y^i| \equiv d_k(x,y)$ versehen, $k \in \overline{\mathbb{N}}$.

"$\Rightarrow$": Diese Richtung folgt mit W-Theorie 17.8, da Π_k stetig bzgl. der Produkttopologie ist.

"$\Leftarrow$": Definiere für $k \in \overline{\mathbb{N}}$, $\epsilon > 0$ und $x \in \mathbb{R}^k$ die offenen Kugeln $B^k(x,\epsilon)$ gemäß $B^k(x,\epsilon) := \{ y = (y^1, \ldots, y^k) \in \mathbb{R}^k : d_k(x,y) < \epsilon \}$. Mit dieser Definition gilt:

$$B^\infty(x,\epsilon) \overset{(1)}{\subset} \Pi_k^{-1} B^k\big(\Pi_k(x), \epsilon\big) \overset{(2)}{\subset} B^\infty(x, 2\epsilon) \quad \text{falls} \sum_{i=k+1}^{\infty} \frac{1}{2^i} < \epsilon, \text{ d.h. } k \geq k_0(\epsilon).$$

Beweis von (1): $y \in B^\infty(x,\epsilon) \Longrightarrow \sum_{i=1}^{\infty} \frac{1}{2^i} \wedge |x^i - y^i| < \epsilon$

$$\Longrightarrow \sum_{i=1}^{k} \frac{1}{2^i} \wedge |x^i - y^i| < \epsilon\,, \text{ d.h. } \Pi_k(y) \in B^k\big(\Pi_k(x),\epsilon\big)$$

Beweis von (2): $y \in \Pi_k^{-1} B^k\big(\Pi_k(x),\epsilon\big)$

$$\Longrightarrow \Pi_k(y) \in B^k\big(\Pi_k(x),\epsilon\big) \Longrightarrow \sum_{i=1}^{k} \frac{1}{2^i} \wedge |x^i - y^i| < \epsilon$$

$$\Longrightarrow 2\epsilon > \sum_{i=1}^{k} \frac{1}{2^i} \wedge |x^i - y^i| + \sum_{i=k+1}^{\infty} \frac{1}{2^i} \geq \sum_{i=1}^{\infty} \frac{1}{2^i} \wedge |x^i - y^i|$$

d.h. $y \in B^\infty(x, 2\epsilon)$

Nach Portemanteau (W-Theorie 29.3) ist z.z.: $\underline{\lim}_{n\to\infty} \mathrm{P}_n[\,G\,] \geq \mathrm{P}_\infty[\,G\,]\,,\ G \subset \mathbb{R}^\infty$ offen.

Sei $G \subset \mathbb{R}^\infty$ offen. Dann existiert zu $x \in G$ ein $\epsilon(x) > 0$ mit $B^\infty\big(x, 2\epsilon(x)\big) \subset G$.

Somit ist $G = \bigcup_{x\in G} B^\infty\big(x,\epsilon(x)\big) = \bigcup_{x \in G} B^\infty\big(x, 2\epsilon(x)\big)$. Da $\mathbb{R}^\infty$ ein Lindelöf-Raum ist (d.h. jede offene Überdeckung enthält eine abzählbare Teilüberdeckung, somit ist jeder Raum mit abzählbarer Basis ein Lindelöf-Raum), existiert eine Folge $(x_i)_{i\in\mathbb{N}} \subset G$ mit $G = \bigcup_{i=1}^{\infty} B^\infty(x_i,\epsilon_i)\ \big(= \bigcup_{i=1}^{\infty} B^\infty(x_i, 2\epsilon_i)\ \big)$, $\epsilon_i := \epsilon(x_i)$. Damit erhält man die Kette

$$\begin{aligned}
\underline{\lim}_{n\to\infty} \mathrm{P}_n[\,G\,] &\equiv \underline{\lim}_{n\to\infty} \mathrm{P}_n[\bigcup_{i=1}^{\infty} B^\infty(x_i, 2\epsilon_i)\,] \\
&\geq \underline{\lim}_{n\to\infty} \mathrm{P}_n[\bigcup_{i=1}^{m} B^\infty(x_i, 2\epsilon_i)\,]\,,\ m \in \mathbb{N} \\
&\overset{(2)}{\geq} \underline{\lim}_{n\to\infty} \mathrm{P}_n[\bigcup_{i=1}^{m} \Pi_k^{-1} B^k\big(\Pi_k(x_i),\epsilon_i\big)\,] \text{ mit } k := \max_{i\leq m} k_0(\epsilon_i) \\
&= \underline{\lim}_{n\to\infty} \mathrm{P}_n[\,\Pi_k^{-1} \bigcup_{i=1}^{m} B^k\big(\Pi_k(x_i),\epsilon_i\big)\,] \\
&\geq \mathrm{P}_\infty \Pi_k^{-1}[\bigcup_{i=1}^{m} B^k\big(\Pi_k(x_i),\epsilon_i\big)\,] \text{ nach Portemanteau und Voraussetzung} \\
&= \mathrm{P}_\infty[\bigcup_{i=1}^{m} \Pi_k^{-1} B^k\big(\Pi_k(x_i),\epsilon_i\big)\,] \\
&\overset{(2)}{\geq} \mathrm{P}_\infty[\bigcup_{i=1}^{m} B^\infty(x_i,\epsilon_i)\,] \uparrow \mathrm{P}_\infty[\,G\,]\,,\ (n \to \infty)
\end{aligned}$$

q.e.d.

Nun zum zentralen

Satz 9.3: (Prohoroff) Sei S polnisch und $\Gamma \subset \mathbb{P}(S)$. Dann sind äquivalent:

(i) Γ ist relativ folgenkompakt

(ii) Γ ist straff

Beweis: (i) $\Rightarrow$(ii): S ist ein Lindelöf-Raum, da S separabel. Somit existiert für alle $n \in \mathbb{N}$ eine Folge $(x_i)_{i\in\mathbb{N}} \subset S$ mit $S = \bigcup_{i=1}^{\infty} B(x_i, \frac{1}{n})$. Sei nun $\epsilon > 0$ vorgegeben.

$(*)$ Beh.: $\forall\, n \in \mathbb{N}\ \exists\, k_n \in \mathbb{N}\ \forall\, \mathrm{P} \in \Gamma:\ \ \mathrm{P}[\bigcup_{i\leq k_n} B(x_i, \frac{1}{n})] > 1 - \frac{\epsilon}{2^n}$

Beweis: Die Behauptung ist für $|\Gamma| = 1$ klar, da $\bigcup_{i=1}^{m} B(x_i, \frac{1}{n}) \uparrow S$, $(m \to \infty)$.

Für den allgemeinen Fall führt man einen Widerspruchsbeweis:

Annahme: $\exists\, n \in \mathbb{N}\ \forall\, k \in \mathbb{N}\ \exists\, \mathrm{P}_k \in \Gamma$ mit $\mathrm{P}_k[\bigcup_{i \le k} B(x_i, \frac{1}{n})] \le 1 - \frac{\epsilon}{2^n}$.

Da Γ relativ folgenkompakt ist, existiert eine Teilfolge $(k')_{k \in \mathbb{N}}$ von $\mathbb{N}$ und ein $\mu \in \mathbb{P}(S)$ mit $\mathrm{P}_{k'} \xrightarrow{w} \mu$, $(k' \to \infty)$. Mit Portemanteau gilt für $l \in \mathbb{N}$:

$$\begin{aligned}
\mu[\bigcup_{i \le l} B(x_i, \frac{1}{n})] &\le \varliminf_{k' \to \infty} \mathrm{P}_{k'}[\bigcup_{i \le l} B(x_i, \frac{1}{n})] \\
&\le \varliminf_{k' \to \infty} \mathrm{P}_{k'}[\bigcup_{i \le k'} B(x_i, \frac{1}{n})] \\
&\le 1 - \frac{\epsilon}{2^n} \quad \text{nach Annahme.}
\end{aligned}$$

Wegen $\mu[\bigcup_{i \le l} B(x_i, \frac{1}{n})] \uparrow 1$, $(l \to \infty)$, ergibt sich somit ein Widerspruch.

Setze nun $K := \bigcap_{n \in \mathbb{N}} \bigcup_{i \le k_n} B(x_i, \frac{1}{n})$. K ist total beschränkt (hier gleichbedeutend mit "präkompakt", da S metrisch), denn es existiert für jedes $\epsilon > 0$ eine endliche Überdeckung von K mit Mengen, deren Durchmesser $\le \epsilon$ ist. S ist vollständig, daher ist $\overline{K}$ kompakt. Für $\overline{K}^C$ gilt:

$$\begin{aligned}
\mathrm{P}[\overline{K}^C] &\le \mathrm{P}[K^C] = \mathrm{P}[\bigcup_{n \in \mathbb{N}} (\bigcup_{i \le k_n} B(x_i, \frac{1}{n}))^C] \\
&\le \sum_{n \in \mathbb{N}} \mathrm{P}[(\bigcup_{i \le k_n} B(x_i, \frac{1}{n}))^C] \le \sum_{n \in \mathbb{N}} \frac{\epsilon}{2^n} \quad \text{nach } (*) \\
&\equiv \epsilon
\end{aligned}$$

Mit $\overline{K}$ hat man die gewünschte kompakte Menge gefunden.

(ii) $\Rightarrow$ (i): 1.Fall: $S = \mathbb{R}^k$. Dieser Fall wurde in der W-Theorie (Satz 18.8) bereits bewiesen.

2.Fall: $S = \mathbb{R}^\infty$. Sei $\{\mathrm{P}_n\}_{n \in \mathbb{N}} \subset \Gamma$. Nach (9.1) ist $\Gamma\Pi_k^{-1}$ straff $\forall\, k \in \mathbb{N}$, daher (da $\Gamma\Pi_k^{-1} \subset \mathbb{P}(\mathbb{R}^k)$) nach Fall 1 auch relativ folgenkompakt. Mit einem Diagonalverfahren erhält man eine Teilfolge $(n')_{n \in \mathbb{N}}$ und für $k \in \mathbb{N}$ ein $\mu_k \in \mathbb{P}(\mathbb{R}^k)$, so daß $\mathrm{P}_{n'}\Pi_k^{-1} \xrightarrow{w} \mu_k$ $(n' \to \infty)$ $\forall\, k \in \mathbb{N}$.

Fasst man Π_k auch als Abbildung von $\mathbb{R}^{k+1}$ nach $\mathbb{R}^k$ auf, d.h. Π_k Projektion von $\mathbb{R}^{k+1}$ auf die ersten k Komponenten, so ist die Menge $\{\mu_k\}_{k \in \mathbb{N}}$ konsistent im Sinne von (2.7) (Kolmogoroff): Für $k \in \mathbb{N}$ hat man

$$\mu_{k+1}\Pi_k^{-1} \equiv \operatorname*{w-lim}_{n' \to \infty}(\mathrm{P}_{n'}\Pi_{k+1}^{-1}\Pi_k^{-1}) = \operatorname*{w-lim}_{n' \to \infty} \mathrm{P}_{n'}\Pi_k^{-1} = \mu_k.$$

Daher existiert nach (2.7) ein $\mathrm{P} \in \mathbb{P}(\mathbb{R}^\infty)$ (vgl. obige Vorüberlegung) mit der Eigenschaft $\mathrm{P}\Pi_k^{-1} = \mu_k$. Somit gilt $\mathrm{P}_{n'}\Pi_k^{-1} \xrightarrow{w} \mathrm{P}\Pi_k^{-1}$, $(n' \to \infty)$, d.h. mit (9.2): $\mathrm{P}_{n'} \xrightarrow{w} \mathrm{P}$, $(n' \to \infty)$.

3.Fall: Sei S nun ein beliebiger polnischer Raum und $\{\mathrm{P}_n\}_{n \in \mathbb{N}} \subset \Gamma$. Dann existiert ein Homöomorphismus $h : S \to B$, wo B eine G_δ-Menge in $[0, 1]^\infty$ ist.

(Eine G_δ-Menge ist ein abzählbarer Durchschnitt von offenen Mengen und daher meßbar.) Mit $\check{h} : B \to S$ werde die Inverse zu h bezeichnet, $j : B \to \mathbb{R}^\infty$ sei die Injektion (stetig in der Unterraumtopologie). Da h und j stetig sind, gelten $\Gamma h^{-1} \subset \mathbb{P}(B)$ und $\Gamma(j \circ h)^{-1} \subset \mathbb{P}(\mathbb{R}^\infty)$ mit $\mathrm{P}_{n'}(j \circ h) \xrightarrow{w} \mu$ für $n' \to \infty$.

Das Problem wird nun auf $\mathbb{P}(B)$ verlagert. Aufgrund der Straffheit von Γh^{-1} hat man für alle $\epsilon > 0$ ein kompaktes K in B (und somit im $\mathbb{R}^\infty$) mit $\mathrm{P}h^{-1}[K] > 1 - \epsilon \quad \forall \mathrm{P} \in \Gamma$.

$$\Longrightarrow \mu[B] \geq \mu[K] \geq \overline{\lim_{n' \to \infty}} \mathrm{P}_{n'}(j \circ h)^{-1}[K] \quad \text{nach Portemanteau}$$
$$= \overline{\lim_{n' \to \infty}} \mathrm{P}_{n'} h^{-1}[K] \geq 1 - \epsilon$$

Folglich ist μ auf B konzentriert, d.h. $\mu[B] = 1$.

Damit kann man den allgemeinen Fall lösen, d.h. es erfolgt nun eine Verlagerung auf $\mathbb{P}(S)$. Setze dazu $\mathrm{P}_\infty := \mu(\check{h})^{-1}$ $\big(\in \mathbb{P}(S)$ wegen $\mu \in \mathbb{P}(B) \big)$.

<u>Beh.:</u> Es gilt $\mathrm{P}_{n'} \xrightarrow{w} \mathrm{P}_\infty$, $(n' \to \infty)$.

<u>Beweis:</u> U offen in $S \iff U = h^{-1}(G')$ für ein G' in B, G' offen in B

G' offen in $B \iff G' = B \cap G = j^{-1}(G)$ für ein offenes G in $\mathbb{R}^\infty$

Somit wieder mit Portemanteau:

$$\underline{\lim_{n' \to \infty}} \mathrm{P}'_n[U] \equiv \underline{\lim_{n' \to \infty}} \mathrm{P}_{n'}(j \circ h)^{-1}[G] \geq \mu[G]$$
$$= \mu[G'] \quad \text{da } \mu \text{ auf } B \text{ konzentriert}$$
$$= \mu(\check{h})^{-1}[U]$$

q.e.d.

<u>Bemerkung 9.4:</u> Die Richtung (ii) $\Rightarrow$ (i) in (9.3) gilt in beliebigen metrischen Räumen. Dieser Fall kann auf den Fall "S σ-kompakt" (S σ-kompakt $\iff$ S kann durch eine Folge kompakter Mengen überdeckt werden) zurückgeführt werden.

Nun wird ein Kriterium dafür gegeben, wann eine Teilmenge $K \subset C$ relativ (folgen-) kompakt ist. Mache dazu die

<u>Definition 1:</u> Für $x \in C$, $T > 0$ und $\delta > 0$ bezeichne $m^T(x, \delta)$ den <u>*Stetigkeitsmodul von x auf $[0,T]$*</u>, d.h. $m^T(x,\delta) := \max\limits_{\substack{0 \leq s,t \leq T \\ |s-t| \leq \delta}} |x(s) - x(t)|$

Mit Hilfe des Stetigkeitsmoduls hat man das bereits aus der Analysis her wohlbekannte Kriterium

<u>Proposition 9.5:</u> (Arzéla-Ascoli)

Eine Menge $K \subset C$ ist genau dann relativ (folgen-) kompakt, wenn gilt:

(i) $\sup\limits_{x \in K} |x(0)| < \infty$

(ii) $\lim\limits_{\delta \downarrow 0} \sup\limits_{x \in K} m^T(x, \delta) = 0 \quad \forall T \in \mathbb{N}$.

<u>Beweis:</u> Zum Beweis vgl. z.B. Karatzas & Shreve p.62 oder Querenburg p.169 .

□

<u>Bemerkung 9.6:</u> Die Bedingung (ii) aus (9.5) besagt, daß K <u>*gleichgradig stetig auf $[0,T]$*</u> ist $\forall T > 0$. Aus den Bedingungen (i) und (ii) in (9.5) folgt, daß $\sup\limits_{x \in K} \sup\limits_{0 \leq t \leq T} |x(t)| < \infty \quad \forall T > 0$.

(vgl. dazu Karatzas & Shreve, Querenburg)

Mit Hilfe von Arzéla-Ascoli ist es nun möglich, die Straffheit von Mengen von Wahrscheinlichkeitsmaßen auf $(C,\mathcal{L})$ zu charakterisieren:

Satz 9.7: Eine Folge $(\mathrm{P}_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf $(C,\mathcal{L})$ ist genau dann straff, falls folgende Bedingungen erfüllt sind:

(i) $\lim\limits_{\alpha\to\infty}\overline{\lim\limits_{n\to\infty}}\,\mathrm{P}_n[\{x\in C:|x(0)|>\alpha\}]=0$

(ii) $\lim\limits_{\delta\downarrow 0}\overline{\lim\limits_{n\to\infty}}\,\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]=0\quad\forall T\in\mathbb{N},\ \epsilon>0$

Beweis: a.) Sei $(\mathrm{P}_n)_{n\in\mathbb{N}}$ straff. Dann existiert zu $\eta>0$ ein kompaktes $K\subset C$ mit der Eigenschaft $\mathrm{P}_n[K^C]\le\eta\quad\forall n\in\mathbb{N}$. Nach (9.5) existiert zu diesem K ein $\alpha>0$ derart, daß $\sup\limits_{x\in K}|x(0)|\le\alpha$, d.h. es besteht die Inklusion $\{x\in C:|x(0)|>\alpha\}\subset K^C$.

Mithin hat man sogar $\lim\limits_{\alpha\to\infty}\sup\limits_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:|x(0)|>\alpha\}]\le\eta$, d.h. mit $\eta\to 0$:

(i)* $\lim\limits_{\alpha\to\infty}\sup\limits_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:|x(0)|>\alpha\}]=0.$

(i)* impliziert (i). Analog dazu existiert nach (9.5) für alle $T>0$, $\epsilon>0$ ein $\delta>0$ mit $\{x\in C:m^T(x,\delta)>\epsilon\}\subset K^C$, wo K wie oben gewählt. Damit erhält man ($\eta\to 0$):

(ii)* $\lim\limits_{\delta\downarrow 0}\sup\limits_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]=0.$

(ii)* impliziert (ii).

b.) Es gelten nun (i) und (ii).

<u>Beh.:</u> Es gelten sogar (i)* und (ii)*.

<u>Beweis:</u> (nur für (ii)*, der Beweis für (i)* verläuft analog)

Mit Hilfe von Epsilontik zeigt man leicht:

$$\text{(ii)}\iff\inf_{\delta>0}\inf_{N\in\mathbb{N}}\sup_{n\ge N}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]=0$$
$$\iff\inf_{N\in\mathbb{N}}\lim_{\delta\downarrow 0}\sup_{n\ge N}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]=0$$

Damit kann man den Term aus (ii)* wie folgt abschätzen:

$$\lim_{\delta\downarrow 0}\sup_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]\le\lim_{\delta\downarrow 0}\sup_{n<N}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]$$
$$+\lim_{\delta\downarrow 0}\sup_{n\ge N}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]$$

Dabei wird der zweite Term der rechten Seite nach obigen Äquivalenzen beliebig klein durch die Wahl von N. Die Menge $\{\mathrm{P}_n:n<N\}$ ist folgenkompakt (leicht), daher nach (9.3) straff. Deshalb gilt gemäß Teil a.): $\lim\limits_{\delta\downarrow 0}\sup\limits_{n<N}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\epsilon\}]=0$, womit man (ii)* bewiesen hat.

c.) Es gelten (i)* und (ii)*. Somit existiert nach (i)* zu $\eta>0$ ein $\alpha<\infty$ mit der Eigenschaft $\sup\limits_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:|x(0)|>\alpha\}]\le\frac{\eta}{2}$. Ferner existiert nach (ii)* zu gleichem η für alle $T>0$ und alle $k\in\mathbb{N}$ ein $\delta_k^T>0$ derart, daß $\sup\limits_{n\in\mathbb{N}}\mathrm{P}_n[\{x\in C:m^T(x,\delta)>\frac{1}{k}\}]\le\frac{\eta}{2^{T+1+k}}$.

Setze $A:=\{x\in C:|x(0)|\le\alpha,\ m^T(x,\delta_k^T)\le\frac{1}{k},\ k,T\in\mathbb{N}\}$. Aufgrund von $\lim\limits_{\delta\downarrow 0}$ wird irgendwann $\delta\le\delta_k^T$ sein für alle $k,T\in\mathbb{N}$. Deshalb folgt mit Arzéla-Ascoli, daß A relativ kompakt ist, d.h. $\overline{A}$ ist kompakt. (Hier ist sogar $A=\overline{A}$.)

Für $\overline{A}^C$ erhält man:

$$\mathrm{P}_n[\,\overline{A}^C\,] \leq \mathrm{P}_n[\,A^C\,] \leq \mathrm{P}_n[\,\{\,x \in C : |x(0)| > \alpha\,\}\,] + \sum_{k,T \in \mathbb{N}} \mathrm{P}_n[\,\{\,x \in C : m^T(x, \delta_k^T) > \frac{1}{k}\,\}\,]$$

$$\leq \frac{\eta}{2} + \sum_{k,T \in \mathbb{N}} \frac{\eta}{2^{T+1+k}} \leq \frac{\eta}{2} + \frac{\eta}{2} = \eta.$$

q.e.d.

Man ist damit am Ziel dieses Abschnitts angelangt, denn man ist jetzt in der Lage, die Straffheit der Polygonprozesse aus Abschnitt 8 nachzuweisen, d.h. es wird nun (8.11) mit Hilfe von (9.7) bewiesen:

Die Bedingung (i) von (9.7) ist klar, da $X_0^n = 0$ und somit $\mathrm{P}(X^n)^{-1}[\,\{\,x \in C : |x(0)| > \alpha\,\}\,] = 0 \quad \forall\, \alpha \geq 0$. Für die Untersuchung von Bedingung (ii) werden nun anstelle von $m^T(X^n, \delta)$ zuerst die Fluktuationen der Irrfahrt untersucht, d.h. $\max_{1 \leq i \leq m} |S_{k+i} - S_k|\,,\ k \in \mathbb{N}_0\,,\ m \in \mathbb{N}$. Diese sind aufgrund der Verteilungsgleichheit der Z_i verteilungsgleich mit $\max_{1 \leq i \leq m} |S_i|$.

Als ein Resultat erhält man

Lemma 9.8: Seien $(Z_i)_{i \in \mathbb{N}}$ und $(S_i)_{i \in \mathbb{N}}$ wie in Abschnitt 8 definiert. Dann gilt:

$$\mathrm{P}[\,\max_{1 \leq i \leq m} \frac{|S_i|}{\sqrt{m}} > \alpha\,] \leq 2 \cdot \mathrm{P}[\,\frac{|S_m|}{\sqrt{m}} > \alpha - \sqrt{2}\,]\,,\ \alpha \in \mathbb{R}.$$

Bemerkung 9.9: Die Kolmogoroff'sche Ungleichung $\mathrm{P}\big[\,\max_{1 \leq i \leq m} |S_i - \mathrm{E}[\,S_i\,]| \geq \alpha\,\big] \leq \dfrac{\mathrm{Var}(S_m)}{\alpha^2}$, welche zu einem Beweis des starken Gesetzes der großen Zahlen verwendet werden kann (vgl. z.B. Ash), liefert hier $\big($ da $\mathrm{E}[\,S_i\,] = 0\,,\ \mathrm{Var}(S_m) = m$ $\big)$ als rechte Seite $\dfrac{1}{\alpha^2}$, was für die weiteren Betrachtungen jedoch nicht ausreicht.

Beweis von 9.8: Sei $\beta < \alpha\,,\ S_i' := \dfrac{S_i}{\sqrt{m}}$. Dann gilt für $i \leq m$:

$$(+) \quad \mathrm{Var}(S_m' - S_i') = \sum_{j=i+1}^{m} \mathrm{Var}(\frac{Z_j}{\sqrt{m}}) \quad \text{da die } Z_j \text{ unabhängig}$$
$$\leq 1\,,\ \text{da } \mathrm{Var}(Z_j) = 1.$$

Mit $\tau := \inf\{\,i \in \mathbb{N} : |S_i'| > \alpha\,\}$ hat man ferner $\{\,\max_{1 \leq i \leq m} |S_i'| > \alpha\,\} = \{\,\tau \leq m\,\}$ (leicht).
Setzt man noch $a := \mathrm{P}[\,|S_m'| > \beta\,]$, so ergibt sich:

$$\mathrm{P}[\,\tau \leq m\,] = \mathrm{P}[\,\tau \leq m\,,\ |S_m'| > \beta\,] + \mathrm{P}[\,\tau \leq m\,,\ |S_m'| \leq \beta\,] \leq a + \sum_{i=1}^{m} \mathrm{P}[\,\tau = i\,,\ |S_m'| \leq \beta\,]$$

$$\leq a + \sum_{i=1}^{m} \mathrm{P}[\,\tau = i\,,\ |S_m' - S_i'| > \alpha - \beta\,] \quad \text{da } |S_i'| > \alpha \text{ nach Defintion von } \tau$$

$$= a + \sum_{i=1}^{m} \mathrm{P}[\,\tau = i\,] \cdot \mathrm{P}[\,|S_m' - S_i'| > \alpha - \beta\,] \quad \text{da } \{\,\tau = i\,\} = \bigcap_{j=1}^{i-1} \{\,|S_i'| > \alpha, |S_j'| \leq \alpha\,\}$$

$$\leq a + \sum_{i=1}^{m} \mathrm{P}[\,\tau = i\,] \cdot \frac{1}{(\alpha - \beta)^2} \quad \text{nach Tschebyscheff und } (+)$$

$$\equiv a + \mathrm{P}[\,\tau \leq m\,] \cdot \frac{1}{(\alpha - \beta)^2}$$

$$\Longrightarrow \mathrm{P}[\,\tau \leq m\,] \leq a \cdot \big(1 - \frac{1}{(\alpha - \beta)^2}\big)^{-1}.$$

Mit $\beta := \alpha - \sqrt{2}$ erhält man die Behauptung.

q.e.d.

Mit diesem Lemma kann nun der letzte Schritt des Beweises von (8.11) vollzogen werden:
Es war noch die Bedingung (ii) von (9.7) für $\{ \mathrm{P}(X^n)^{-1} \}_{n\in\mathbb{N}}$ nachzuweisen, d.h. für alle $T \in \mathbb{N}$ und $\epsilon > 0$: $\quad \lim_{\delta\downarrow 0} \overline{\lim_{n\to\infty}} \mathrm{P}[\{ m^T(X^n,\delta) > \epsilon \}] = 0$

Sei dazu $\epsilon > 0$. Zu $\delta > 0$ wähle man ein $\delta_n \in \{ \frac{k}{n}, \; k \in \mathbb{N} \}$ mit der Eigenschaft $\delta \le \delta_n < \delta + \frac{1}{n}$ und zu s,t mit $0 \le s < t \le s+\delta$ ein $m \in \mathbb{N}_0$ derart, daß $m \cdot \delta_n \le s < (m+1)\cdot\delta_n$. Dann ist $t < (m+2)\cdot\delta_n$.

Fall $t \le (m+1)\cdot\delta_n$: Zerlege $|X_t^n - X_s^n| \le |X_s^n - X_{m\cdot\delta_n}^n| + |X_t^n - X_{m\cdot\delta_n}^n|$

Fall $(m+1)\cdot\delta_n < t < (m+2)\cdot\delta_n$: Zerlege $|X_t^n - X_s^n| \le |X_s^n - X_{m\cdot\delta_n}^n| + |X_t^n - X_{(m+1)\cdot\delta_n}^n| + |X_{(m+1)\cdot\delta_n}^n - X_{m\cdot\delta_n}^n|$

In beiden Fällen gilt: $\quad |X_t^n - X_s^n| \le 3 \cdot \max_{\substack{k\in\mathbb{N}_0 \\ k\cdot\delta_n \le t}} \max_{t'\in[k\cdot\delta_n,(k+1)\cdot\delta_n]} |X_{t'}^n - X_{k\cdot\delta_n}^n| =: (*)$

Für festes k ist $(|X_t^n - X_{k\cdot\delta_n}^n|)_{t\in[k\cdot\delta_n,(k+1)\cdot\delta_n]}$ ebenfalls ein Polygonprozeß mit Richtungsänderungen an Vielfachen von $\frac{1}{n}$ (weil dort X_t^n die Richtung ändert) und beim Erreichen von 0 (aufgrund des Betrags).
Daher genügt es, die Vielfachen von $\frac{1}{n}$ im Intervall $[\, k\cdot\delta_n, (k+1)\cdot\delta_n \,]$ zu untersuchen, d.h. es ergibt sich $\;(*) = 3 \cdot \max_{\substack{k \le \frac{t}{\delta_n} \\ k\in\mathbb{N}_0}} \max_{\substack{\frac{i}{n} \le \delta_n \\ i\in\mathbb{N}}} |X_{k\cdot\delta_n+\frac{i}{n}}^n - X_{k\cdot\delta_n}^n|$. Damit erhält man

$$(++) \qquad \{ m^T(X^n,\delta) > \epsilon \} \subset \bigcup_{0\le k\le \frac{T}{\delta_n}} \{ \max_{\substack{i\le n\cdot\delta_n \\ i\in\mathbb{N}}} |X_{k\cdot\delta_n+\frac{i}{n}}^n - X_{k\cdot\delta_n}^n| > \frac{\epsilon}{3} \}$$

Denn aus $m^T\big(X^n(\omega),\delta\big) > \epsilon$ folgt: $\exists\, s,t$ mit $0 \le s < t \le T$ und $|s-t| < \delta$, so daß $|X_t^n(\omega) - X_s^n(\omega)| > \epsilon$

$$\Longrightarrow \max_{\substack{k \le \frac{t}{\delta_n} \\ k\in\mathbb{N}_0}} \max_{\substack{\frac{i}{n} \le \delta_n \\ i\in\mathbb{N}}} |X_{k\cdot\delta_n+\frac{i}{n}}^n(\omega) - X_{k\cdot\delta_n}^n(\omega)| > \frac{\epsilon}{3}.$$

Aus (++) folgt:
$$\begin{aligned}
\mathrm{P}[\, m^T(X^n,\delta) > \epsilon \,] &\le \sum_{0\le k\le\frac{T}{\delta_n}} \mathrm{P}[\max_{\frac{i}{n}\le\delta_n} |\sum_{j=1}^{i} \frac{Z_{k\cdot n\cdot\delta_n+j}}{\sqrt{n}}| > \frac{\epsilon}{3}] \\
&= \sum_{0\le k\le\frac{T}{\delta_n}} \mathrm{P}[\max_{\frac{i}{n}\le\delta_n} \frac{|S_i|}{\sqrt{n}} > \frac{\epsilon}{3}] \quad \text{da die } Z_j \text{ identisch verteilt} \\
&\le (1+\frac{T}{\delta}) \cdot \mathrm{P}[\max_{\frac{i}{n}\le\delta_n} \frac{|S_i|}{\sqrt{n}} > \frac{\epsilon}{3}] \quad \text{da } \frac{T}{\delta_n} \le \frac{T}{\delta} \\
&\le (1+\frac{T}{\delta}) \cdot \mathrm{P}[\max_{\frac{i}{n}\le\delta_n} \frac{|S_i|}{\sqrt{n\cdot\delta_n}} > \frac{\epsilon}{3\sqrt{2\delta}}] \quad \text{für } n \ge \frac{1}{\delta}, \text{ da dann } \delta_n \le 2\delta \\
&\le O(\frac{1}{\delta}) \cdot \mathrm{P}[\frac{|S_{n\cdot\delta_n}|}{\sqrt{n\cdot\delta_n}} > \frac{1}{O(\sqrt{\delta})}] \quad \text{nach (9.8), } n \ge \frac{1}{\delta} \\
&\longrightarrow O(\frac{1}{\delta}) \cdot \mathrm{P}[\, |Y| \ge \frac{1}{O(\sqrt{\delta})} \,], \; (n\to\infty) \quad \text{mit } Y \sim \mathrm{N}(0,1) \\
&\le O(\frac{1}{\delta}) \cdot \big(O(\sqrt{\delta})\big)^k \cdot \mathrm{E}[\, |Y|^k \,] \quad \text{für } k > 2 \text{ nach Markoff} \\
&\longrightarrow 0\,, \; (\delta \to 0).
\end{aligned}$$

q.e.d.

Mit diesem Beweis ist die Betrachtung der standardisierten Brown'schen Bewegung zunächst abgeschlossen, sie wird später noch verallgemeinert.
Nun werden als Vorbereitung für die nachfolgenden Abschnitte allgemeine Eigenschaften von stochastischen Prozessen untersucht sowie der schon aus dem zeitdiskreten Fall her bekannte Begriff der Stoppzeit erweitert.

Nun werden als Vorbereitung für die nachfolgenden Abschnitte allgemeine Eigenschaften stochastischer Prozesse untersucht sowie der schon aus dem zeitdiskreten Fall her bekannte Begriff der Stoppzeit erweitert.

Abschnitt 10: Stochastische Prozesse und Stoppzeiten

Für die folgenden Betrachtungen sei ein metrischer Raum (S,d) gegeben, $\mathcal{S}$ bezeichne die Borel-σ-Algebra bzgl. d. Falls nichts anderes gesagt wird, ist $X = (X_t)_{t\geq 0}$ ein stochastischer Prozeß auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Werten in $(S, \mathcal{S})$.

Eine zu untersuchende Eigenschaft stochastischer Prozesse ist z.B. die Gleichheit, d.h. die Beantwortung der Frage: Wann sind zwei stochastische Prozesse $X = (X_t)_{t\geq 0}$, $Y = (Y_t)_{t\geq 0}$ mit gleichem Zustandsraum S als gleich anzusehen ?

Zur Lösung dieser Frage definiert man

Definition 1: Seien X, Y stochastische Prozesse auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$.
Dann heißt Y *Modifikation* von X, wenn $X_t = Y_t$ P-fast sicher $\forall\, 0 \leq t < \infty$.
X und Y heißen *nicht unterscheidbar*, falls $X_t = Y_t \quad \forall\, 0 \leq t < \infty$ P-fast sicher.

Definition 2: Seien X, Y stochastische Prozesse. X sei auf $(\Omega, \mathcal{F}, \mathrm{P})$, Y auf $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathrm{P}})$ definiert. Man sagt, daß X und Y die gleichen *endlich-dimensionalen Randverteilungen* besitzen, wenn gilt: $\mathrm{P}(X_{t_1}, \ldots, X_{t_n})^{-1} = \widetilde{\mathrm{P}}(Y_{t_1}, \ldots, Y_{t_n})^{-1} \quad \forall\, 0 \leq t_1 < \ldots < t_n < \infty\,,\ n \in \mathbb{N}$.

Zwischen diesen Begriffen bestehen die nachstehenden Relationen:

Lemma 10.1: Es gelten die Implikationen (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) mit

(i) X und Y sind nicht unterscheidbar

(ii) Y ist Modifikation von X

(iii) X und Y haben die gleichen endlich-dimensionalen Randverteilungen.

Haben ferner X und Y rechtsstetige (bzw. linksstetige) Pfade, so gilt sogar (i) $\Leftrightarrow$ (ii).

Beweis: (i) $\Rightarrow$ (ii): Sind X, Y nicht unterscheidbar, so existiert eine Menge vom Maß 1 derart, daß $X_t = Y_t \quad \forall\, 0 \leq t < \infty$ auf dieser Menge. Ist X eine Modifikation von Y, so steht für jedes $t \geq 0$ eine Menge vom Maß 1 zur Verfügung, so daß $X_t = Y_t$ auf dieser Menge. Die erste Bedingung ist also stärker als die zweite, man verwende im zweiten Fall als jeweilige Menge vom Maß 1 die globale Menge aus dem ersten Fall.

(ii) $\Rightarrow$ (iii): Zu vorgegebenen $t_1, \ldots, t_n$ gilt nach Voraussetzung $X_{t_m} = Y_{t_m}\,,\ 1 \leq m \leq n$ P-fast sicher. Daraus folgt die Gleichheit der Bildmaße.

(ii) $\Rightarrow$ (i): Nach (ii) gilt $X_t = Y_t \quad \forall\, t \in \mathbb{Q}^+$ P-fast sicher (abzählbarer Schnitt der jeweiligen Mengen vom Maß 1). Für $r \in \mathbb{R}^+ \setminus \mathbb{Q}^+$ sei $(t_i)_{i\in\mathbb{N}} \subset \mathbb{Q}^+$ eine Folge mit der Eigenschaft $t_i \downarrow r\,,\ (i \to \infty)$
$\Longrightarrow Y_r = \lim\limits_{t_i \downarrow r} Y_{t_i} = \lim\limits_{t_i \downarrow r} X_{t_i} = X_r$ P-fast sicher.

q.e.d.

Es werden nun die schon aus Abschnitt 4 her bekannten Begriffe wie Filterung und Adaptiertheit übertragen und einige Eigenschaften stochastischer Prozesse zusammengestellt:

Definition 3: Ein stochastischer Prozeß X auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit Zustandsraum $(S, \mathcal{S})$ heißt *meßbar*, falls die Abbildung $(t,\omega) \mapsto X_t(\omega) \quad [0,\infty) \cap \mathcal{B} \otimes \mathcal{F} - \mathcal{S}$-meßbar ist.
Eine *Filterung* (*in $\mathcal{F}$*) ist eine Familie von σ-Algebren $\mathbb{F} \equiv (\mathcal{F}_\mathrm{t})_{\mathrm{t}\geq 0}$ mit der Eigenschaft $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \quad \forall\, 0 \leq s \leq t < \infty$. Dabei definiert man als Abschluß nach rechts die übergeordnete σ-Algebra $\mathcal{F}_\infty := \sigma(\bigcup\limits_{t\geq 0} \mathcal{F}_t) \subset \mathcal{F}$.

Eine Filterung $\mathbb{F}$ wird als <u>rechtsstetig</u> bezeichnet, falls $\mathbb{F} = \mathbb{F}_+$, wobei $\mathbb{F}_+ = (\mathcal{F}_{t+})_{t \geq 0}$ gegeben wird durch $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$.

Analog zum zeitdiskreten Fall heißt X an $\mathbb{F}$ <u>adaptiert</u>, wenn X_t $\mathcal{F}_t - \mathcal{S}$-meßbar ist $\forall\, t \geq 0$. In diesem Zusammenhang wird mit $\mathbb{F}^X$ wieder die <u>kanonische Filterung</u> notiert, d.h. mit anderen Worten: $\mathcal{F}_t^X := \sigma(X_s\,,\ s \leq t)$.

X heißt <u>$\mathbb{F}$-progressiv</u> $:\Longleftrightarrow$ Die Abbildung $(s, \omega) \mapsto X_s(\omega)$ von $[0,t] \times \Omega$ nach S ist $[0,t] \cap \mathcal{B} \otimes \mathcal{F}_t - \mathcal{S}$-meßbar $\forall\, t \geq 0$.

Bemerkung: "Filterung" und "Adaptiertheit" werden im zeitdiskreten Fall näher erläutert, vgl. dort. Die Filterung $\mathbb{F}_+$ (in (10.2) wird gezeigt, daß $\mathbb{F}_+$ eine Filterung ist) ist geringfügig feiner als die Ausgangsfilterung. Somit besagt die Rechtsstetigkeit einer Filterung, daß sich die Ereignismenge bei infinitesimaler Zeitverschiebung in Richtung Zukunft hin nicht verändert.
Die Meßbarkeit eines Prozesses X ist z.B. für die Anwendbarkeit des Satzes von Fubibni wichtig, während die Progressivität von X, wie nachher gezeigt werden wird, eine stärkere Eigenschaft ist.

Eine Zusammenstellung von Eigenschaften der Begriffe aus Definition 3 beinhaltet

Lemma 10.2: a.) Ist $\mathbb{F}$ eine Filterung, so ist $\mathbb{F}_+$ ebenfalls eine Filterung.

b.) $\mathbb{F}_+$ ist rechtsstetig.

c.) Ist X $\mathbb{F}$-progressiv, so ist X meßbar und an $\mathbb{F}$ adaptiert.

d.) X ist meßbar $\Longleftrightarrow$ X ist $\mathbb{F}$-progressiv mit $\mathcal{F}_t := \mathcal{F} \quad \forall\, t \geq 0$.

Beweis: a.) klar, da $\bigcap_{s>t} \mathcal{F}_t$ eine σ-Algebra. Mit $t_1 \geq t_2$ folgt $\bigcap_{s>t_2} \mathcal{F}_s \subset \bigcap_{s>t_1} \mathcal{F}_s$.

b.) z.z.: $(\mathbb{F}_+)_+ = \mathbb{F}_+$

$$A \in \mathcal{F}_{t+} \Longleftrightarrow A \in \bigcap_{s>t} \mathcal{F}_s \Longleftrightarrow A \in \mathcal{F}_s \quad \forall\, s > t \Longleftrightarrow A \in \bigcap_{r>s} \mathcal{F}_r \quad \forall\, s > t$$
$$\Longleftrightarrow A \in \bigcap_{s>t} \bigcap_{r>s} \mathcal{F}_r \equiv (\mathcal{F}_{t+})_+$$

c.) $$X^{-1}(\mathcal{S}) = X^{-1}(\mathcal{S}) \cap \Big(\bigcup_{n \in \mathbb{N}} [0,n] \times \Omega \Big) = \bigcup_{n \in \mathbb{N}} \big(X^{-1}(\mathcal{S}) \cap ([0,n] \times \Omega) \big)$$
$$\subset \bigcup_{n \in \mathbb{N}} ([0,n] \cap \mathcal{B} \otimes \mathcal{F}_n) \quad \text{aufgrund der Progressivität}$$
$$\subset [0,\infty) \cap \mathcal{B} \otimes \mathcal{F}\,, \text{ d.h. } X \text{ ist meßbar.}$$

d.) "$\Leftarrow$" ist klar nach c.)
"$\Rightarrow$": $X^{-1}(\mathcal{S}) \subset [0,\infty) \cap \mathcal{B} \otimes \mathcal{F} \Longrightarrow (X\,|_{[0,t] \times \Omega})^{-1}(\mathcal{S}) \subset [0,t] \cap \mathcal{B} \otimes \mathcal{F}$.

q.e.d.

Beispiel: Ist $S = \mathbb{R}$ und besitzt X rechtsstetige Pfade, so gilt: $\{ \lim_{h \to 0} \frac{1}{h} \cdot (X_{t+h} - X_t) \geq 0 \} \in \mathcal{F}_{t+}^X$

<u>Beweis:</u> Aufgrund von Rechtsstetigkeit ergibt sich:

$$\lim_{h \to 0} \frac{1}{h} \cdot (X_{t+h} - X_t) \equiv \sup_{h>0} \inf_{s \in (0,h)} \frac{1}{s} \cdot (X_{t+s} - X_t)$$
$$= \sup_{\substack{h=\frac{1}{n} \\ n \in \mathbb{N}\,,\, n \geq n_0}} \inf_{s \in (0,h) \cap \mathbb{Q}} \frac{1}{s} \cdot (X_{t+s} - X_t) \in \mathcal{F}^X_{t+\frac{1}{n_0}} \quad \forall\, n_0 \in \mathbb{N}$$

q.e.d.

Eine weitere Kennzeichnung von Progressivität zeigt

Proposition 10.3: X sei an $\mathbb{F}$ adaptiert. Sind zusätzlich alle Pfade von X rechtsstetig oder alle Pfade linksstetig, so ist X $\mathbb{F}$-progressiv.

Beweis: Der Beweis wird hier nur für den rechtsstetigen Fall geführt, der Beweis für den anderen Fall ist durch eine analoge Konstruktion führbar. Sei o.E. $t > 0$ (im Fall $t = 0$ ist nichts zu zeigen). Für $n \in \mathbb{N}$ definiert man sich eine Abbildung von $[0,t] \times \Omega$ nach S gemäß der Vorschrift $(s,\omega) \mapsto X_s^n(\omega) := X_{\frac{k}{n}\cdot t}(\omega)$ für $s \in [\frac{k-1}{n} \cdot t, \frac{k}{n} \cdot t)$. Die Abbildung $(s,\omega) \mapsto X_s^n(\omega)$ ist $[0,t] \cap \mathcal{B} \otimes \mathcal{F}_t - \mathcal{S}$-meßbar, da $(X^n)^{-1}(\mathcal{S}) = \bigcup_{k=1}^{n} [\frac{k-1}{n} \cdot t, \frac{k}{n} \cdot t) \times X_{\frac{k}{n}\cdot t}^{-1}(\mathcal{S}) \subset [0,t] \cap \mathcal{B} \otimes \mathcal{F}_t$.

Da $d\big(X_s(\omega), X_s^n(\omega)\big) \leq \sup_{k \leq \frac{t}{n}} d\big(X_s(\omega), X_{s+k}(\omega)\big) \longrightarrow 0$, $(n \to \infty)$, gilt mithin $\lim_{n\to\infty} X_s^n(\omega) = X_s(\omega)$ und somit die Behauptung.

q.e.d.

Als Vorstufe zum Begriff der Stoppzeit definiert man

Definition 4: Eine *Zufallszeit* τ ist eine $[0,\infty]$-wertige $\mathcal{F} - \overline{\mathcal{B}} \cap [0,\infty]$-meßbare Zufallsvariable. Mit τ definiere als Vorbereitung auf gestoppte Prozesse $X_\tau(\omega) := \begin{cases} X_{\tau(\omega)}(\omega) & \text{auf } \{\tau < \infty\} \\ e \notin S & \text{auf } \{\tau = \infty\} \end{cases}$

Beachte: τ ist $\mathcal{F} - \overline{\mathcal{B}} \cap [0,\infty]$-meßbar $\iff \{\tau \leq t\} \in \mathcal{F} \quad \forall\, 0 \leq t < \infty$
(W-Theorie 17.2)

Für den Umgang mit X_τ nun zwei wichtige Eigenschaften:

Lemma 10.4: Sei X meßbar und τ eine Zufallszeit. Dann gelten:

a.) X_τ ist eine Zufallsvariable

b.) $\sigma(X_\tau) = \big\{\{X_\tau \in B, \tau < \infty\},\ \{\tau = \infty\} \cup \{X_\tau \in B, \tau < \infty\},\ B \in \mathcal{S}\big\}$

Beweis: a.) Die Abbildung $g : \{\tau < \infty\} \to S$ mit $g(\omega) := X_{\tau(\omega)}(\omega)$, ist $\mathcal{F} - \mathcal{S}$-meßbar, denn: $f : \{\tau < \infty\} \to [0,\infty) \times \Omega$ mit $f(\omega) := \big(\tau(\omega), \omega\big)$ ist nach W-Theorie 8.10 $\mathcal{F} - \mathcal{B} \cap [0,\infty) \otimes \mathcal{F}$-meßbar. Da X meßbar ist, ist die Abbildung $h : [0,\infty) \times \Omega \to S$ mit $h(s,\omega) := X_s(\omega)$ $[0,\infty) \cap \mathcal{B} \otimes \mathcal{F} - \mathcal{S}$-meßbar, daher ist wegen $g = h \circ f$ obiges klar.
Auf $\{\tau = \infty\} \in \mathcal{F}$ ist X_τ konstant, also ist X_τ eine Zufallsvariable.

b.) Man unterscheidet die folgenden zwei Fälle ($B \in \mathcal{S}$):
1.Fall: $A := B \implies X_\tau^{-1}(A) = \{X_\tau \in B, \tau < \infty\}$.
2.Fall: $A := B \cup \{e\} \implies X_\tau^{-1}(A) = \{X_\tau \in B, \tau < \infty\} \cup \{\tau = \infty\}$.

q.e.d.

Nun eine Verschärfung des Begriffs "Zufallszeit", nämlich

Definition 5: Sei $\mathbb{F}$ eine Filterung in $\mathcal{F}$. Eine Abbildung $\tau : \Omega \to [0,\infty]$ heißt eine *$\mathbb{F}$-Stoppzeit* $:\iff \{\tau \leq t\} \in \mathcal{F}_t \quad \forall\, t \geq 0$. Beachte: Jede Stoppzeit ist eine Zufallszeit.

Einige nützliche Eigenschaften einer $\mathbb{F}$-Stoppzeit beschreibt

Lemma 10.5: Es gelten die Implikationen (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Leftrightarrow$ (iv) mit

(i) τ ist konstant

(ii) τ ist eine Stoppzeit

(iii) $\{\tau < t\} \in \mathcal{F}_t \quad \forall\, t \geq 0$

(iv) τ ist $\mathbb{F}_+$-Stoppzeit

Beweis: (i) $\Rightarrow$ (ii): $\tau = a \in [0, \infty] \Longrightarrow \{ \tau \le t \} = \begin{cases} \emptyset \in \mathcal{F}_t & \text{für } t < a \\ \Omega \in \mathcal{F}_t & \text{für } t \ge a \end{cases}$

(ii) $\Rightarrow$ (iii): $\{ \tau < t \} = \bigcup\limits_{n=1}^{\infty} \{ \tau \le t - \frac{1}{n} \}$, dabei ist $\{ \tau \le t - \frac{1}{n} \} \in \mathcal{F}_{(t-\frac{1}{n})} \subset \mathcal{F}_t$

(iii) $\Rightarrow$ (iv): $\{ \tau \le t \} = \bigcap\limits_{\substack{n \ge m \\ n \in \mathbb{N}}} \{ \tau < t + \frac{1}{n} \} \in \mathcal{F}_{(t+\frac{1}{m})} \quad \forall\, m \in \mathbb{N}$

$$\Longrightarrow \{ \tau \le t \} \in \bigcap_{m \in \mathbb{N}} \mathcal{F}_{(t+\frac{1}{m})} = \mathcal{F}_{t+}$$

(iv) $\Rightarrow$ (iii): $\{ \tau < t \} = \bigcup\limits_{n=1}^{\infty} \{ \tau \le t - \frac{1}{n} \}$, dabei ist $\{ \tau \le t - \frac{1}{n} \} \in \mathcal{F}_{(t-\frac{1}{n})+} \subset \mathcal{F}_t$

q.e.d.

In Abschnitt 4 wurde beim Ruinproblem und im Zusammenhang mit der Doob'schen Ungleichung schon der Begriff "Eintrittszeit" ohne nähere präzise Definition verwendet. Diese wird nun hier nachgeholt.

Definition 6: Für $B \in \mathcal{S}$ sei $\tau_B(\omega) := \inf\{ t \ge 0 : X_t(\omega) \in B \}$ mit $\inf \emptyset := \infty$. Dann wird τ_B als *Eintrittszeit von X in B* bezeichnet. Als weiteren Begriff hat man damit den der *Kontaktzeit von X zu B*, welcher durch $\widetilde{\tau_B} := \tau_B \wedge \inf\{ t \ge 0 : X_{t-} \in B \}$ definiert wird (setze $X_{0-} := X_0$), falls X linksseitige Limites besitzt.
Beachte: Ist X linksseitig stetig, so folgt $\tau_B = \widetilde{\tau_B}$.

Zur Beziehung Eintrittszeit-Stoppzeit bzw. Kontaktzeit-Stoppzeit dient die

Proposition 10.6: Sei X an $\mathbb{F}$ adaptiert. Dann gelten:

a.) Hat X rechtsstetige Pfade und ist B offen, so ist τ_B eine $\mathbb{F}_+$-Stoppzeit.

b.) Hat X rechtsstetige Pfade mit linksseitigen Limites, so gilt:
Im Fall B abgeschlossen ist $\widetilde{\tau_B}$ eine $\mathbb{F}$-Stoppzeit.

c.) Hat X stetige Pfade und ist B abgeschlossen, so ist τ_B eine $\mathbb{F}$-Stoppzeit.

Beweis: a.) Wegen (10.5) ist nur z.z.: $\{ \tau_B < t \} \in \mathcal{F}_t \quad \forall\, t \ge 0$.

$$\{ \tau_B < t \} \overset{(+)}{=} \bigcup_{s \in [0,t)} \{ X_s \in B \} \overset{(*)}{=} \bigcup_{s \in [0,t) \cap \mathbb{Q}} \{ X_s \in B \}$$

Der Beweis von (+) ist klar. Beweis von ($*$):

"$\supseteq$" ist klar.

"$\subseteq$": Sei $\omega \in \bigcup\limits_{s \in [0,t)} \{ X_s \in B \} \Longrightarrow \omega \in \{ X_s \in B \}$ für ein $s \in [0, t)$.

Der Fall $s \in \mathbb{Q}$ ist klar. Sei daher $s \in \mathbb{R} \setminus \mathbb{Q}$.
B offen $\Longrightarrow \exists$ Umgebung U von $X_s(\omega)$ mit $U \subset B$.
$X_\bullet(\omega)$ rechtsstetig $\Longrightarrow \exists\, \epsilon > 0$ mit $X_{s'}(\omega) \in U \subset B \quad \forall\, s' \in [s, s+\epsilon)\,,\ s + \epsilon < t$.
Unter diesen s' befinden sich auch rationale Zahlen, damit folgt die Teilbehauptung.

Da X an $\mathbb{F}$ adaptiert und $B \in \mathcal{S}$ ist, folgt $\bigcup\limits_{s \in [0,t) \cap \mathbb{Q}} \{ X_s \in B \} \in \mathcal{F}_t$.

b.) 1.) Aus $s_n \downarrow s\,,\ (n \to \infty)\,,\ s \ge 0$, folgt $X_{s_n-} \longrightarrow X_s(\omega)\,,\ (n \to \infty)$. Denn:
X rechtsstetig $\Longrightarrow \forall\, \epsilon > 0\ \exists\, \delta > 0$, so daß $d\big(X_t(\omega), X_s(\omega)\big) < \epsilon \quad \forall\, t \in (s, s+\delta]$
$\Longrightarrow d\big(X_{t-}(\omega), X_s(\omega)\big) \le \epsilon \quad \forall\, t \in (s, s+\delta]$

Mit $t := s_n$, wo $n \ge n_0(\epsilon)$, folgt diese Teilbehauptung.

2.) Sei $B_n := \{ x \in S : d(x,B) < \frac{1}{n} \}$. B_n ist offen. Dann gilt

$$(**) \quad \{ \widetilde{\tau_B} \le t \} = \{ X_t \in B \} \cup \bigcap_{n \in \mathbb{N}} \{ \tau_{B_n} < t \} \quad \forall t \ge 0.$$

Hat man $(**)$ bewiesen, so erhält man die Behauptung mit Teil a.).

3.) Beweis von $(**)$:

"$\subseteq$": $s := \widetilde{\tau_B}(\omega) \le t \implies \exists (s_n)_{n \in \mathbb{N}}$ mit $s_n \downarrow s$, $(n \to \infty)$, und mindestens einer der beiden Eigenschaften (wegen der Rechtsstetigkeit)

i.) $X_{s_n}(\omega) \in B$ ii.) $X_{s_n-}(\omega) \in B$.

$\implies X_s(\omega) \in B$ (im Fall ii.) mit 1.)), da B abgeschlossen.

<u>Fall 1:</u> $s = t$: Dann ist $X_t(\omega) \in B$.

<u>Fall 2:</u> $s < t$: $\implies \forall n \in \mathbb{N}\ \exists \epsilon_n > 0$ mit $s + \epsilon_n < t$ und $X_{s+\epsilon_n}(\omega) \in B_n$, da die Pfade rechtsstetig sind.

$$\Big(\forall n \in \mathbb{N}\ \exists \delta_n \text{ mit } d\big(X_s(\omega), X_t(\omega)\big) < \frac{1}{n} \quad \forall t \in [s, s+\delta_n]$$

$$\implies d\big(X_t(\omega), B\big) \le d\big(X_s(\omega), X_t(\omega)\big) + d\big(X_s(\omega), B\big)$$

$$= d\big(X_s(\omega), X_t(\omega)\big) < \frac{1}{n}, \quad \text{d.h. } X_t(\omega) \in B_n \Big)$$

$\implies \forall n \in \mathbb{N}$ gilt: $\tau_{B_n} \le s + \epsilon_n < t$.

"$\supseteq$": <u>Fall 1:</u> Für $X_t(\omega) \in B$ ist diese Richtung klar.

<u>Fall 2:</u> Für alle $n \in \mathbb{N}$ gelte $\tau_{B_n}(\omega) < t$. Dann existiert eine Folge $(s_n)_{n \in \mathbb{N}}$ mit $s_n < t$ und $X_{s_n}(\omega) \in B_n$. Mit s werde ein Häufungspunkt von $(s_n)_{n \in \mathbb{N}}$ bezeichnet, $s \in [0,t]$.

<u>Fall 2a:</u> Es existiert eine Teilfolge $(n')_{n \in \mathbb{N}}$ mit $s_{n'} \uparrow s$, $(n' \to \infty)$
$\implies X_{s-}(\omega) \in B$ (leicht)

<u>Fall 2b:</u> Es existiert eine Teilfolge $(n')_{n \in \mathbb{N}}$ mit $s_{n'} \downarrow s$, $(n' \to \infty)$
$\implies X_s(\omega) \in B$ (leicht)

c.) Da X stetige Pfade hat, gilt $\widetilde{\tau_B} = \tau_B$ und mit b.) folgt die Behauptung.

q.e.d.

Das nachstehende Lemma ist wohlbekannt, vgl. z.B. Abschnitt 6:

<u>Lemma 10.7:</u> Sind σ, τ $\mathbb{F}$-Stoppzeiten, so sind auch $\tau \wedge \sigma$ und $\tau \vee \sigma$ $\mathbb{F}$-Stoppzeiten.

<u>Beweis:</u> $\{ \tau \wedge \sigma \le t \} = \{ \tau \le t \} \cup \{ \sigma \le t \}$; $\{ \tau \vee \sigma \le t \} = \{ \tau \le t \} \cap \{ \sigma \le t \}$.

q.e.d.

Für $\mathbb{F}_+$-Stoppzeiten hat man noch

<u>Lemma 10.8:</u> a.) τ ist eine $\mathbb{F}_+$-Stoppzeit $\iff \tau \wedge t$ ist $\mathcal{F}_t - \mathcal{B}$-meßbar $\forall t \ge 0$.

b.) Ist τ eine $\mathbb{F}_+$-Stoppzeit und $s > 0$, so ist $\tau + s$ eine $\mathbb{F}$-Stoppzeit.

<u>Beweis:</u> a.) "$\Rightarrow$": $\{ \tau \wedge t \le s \} = \begin{cases} \{ \tau \le s \} \in \mathcal{F}_{s+} \subset \mathcal{F}_t & \text{falls } s < t \\ \Omega \in \mathcal{F}_t & \text{falls } s \ge t \end{cases}$

"$\Leftarrow$": z.z.: $\{ \tau < t \} \in \mathcal{F}_t \quad \forall t \ge 0$. $\{ \tau < t \} = \{ \tau \wedge t < t \} \in \mathcal{F}_t$

b.) $\{ \tau + s \le t \} = \{ \tau \le t - s \} \in \mathcal{F}_{(t-s)+} \subset \mathcal{F}_t$

q.e.d.

In Abschnitt 12 werden starke Markoff-Prozesse vorgestellt, bei denen die Markoff-Eigenschaft mit Stoppzeiten anstelle von festen Zeitpunkten als Gegenwart verwendet wird. Für diese Betrachtungen benötigt man

Definition 7: Ist τ eine $\mathbb{F}$-Stoppzeit, so definiere $\mathcal{F}_\tau := \{ A \in \mathcal{F}_\infty : A \cap \{ \tau \le t \} \in \mathcal{F}_t \quad \forall t \ge 0 \}$. $\mathcal{F}_\tau$ heißt <u>*σ-Algebra der τ-Vergangenheit*</u> (d.h. $\mathcal{F}_\tau$ beinhaltet die Ereignisse bis zum Zeitpunkt τ)

Die Namensgebung "σ-Algebra der τ-Vergangenheit" ist berechtigt. Dies und andere Eigenschaften zeigt

Lemma 10.9: Sei τ eine $\mathbb{F}$-Stoppzeit. Dann gelten:

a.) $\mathcal{F}_\tau$ ist eine σ-Algebra.

b.) τ ist $\mathcal{F}_\tau - \overline{\mathcal{B}}$-meßbar.

c.) Ist $\tau = t_0 = \text{const.} \implies \mathcal{F}_\tau = \mathcal{F}_{t_0}$

d.) Aus $\sigma \le \tau$ folgt $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$, falls σ eine $\mathbb{F}$-Stoppzeit ist.

Beweis: a.) $\Omega \in \mathcal{F}_\tau$; $A \in \mathcal{F}_\tau \implies A^C \cap \{ \tau \le t \} = \{ \tau \le t \} - A \cap \{ \tau \le t \} \in \mathcal{F}_t$, d.h. $A^C \in \mathcal{F}_\tau$.

$(A_i)_{i \in \mathbb{N}} \subset \mathcal{F}_\tau \implies (\bigcup_{i=1}^{\infty} A_i) \cap \{ \tau \le t \} = \bigcup_{i=1}^{\infty} (A_i \cap \{ \tau \le t \}) \in \mathcal{F}_t$, d.h. $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}_\tau$.

b.) z.z.: $\{ \tau \le s \} \in \mathcal{F}_\tau \quad \forall s \ge 0$. $\{ \tau \le s \} \cap \{ \tau \le t \} = \{ \tau \le s \wedge t \} \in \mathcal{F}_t \quad \forall s \ge 0$.

c.) Für $A \in \mathcal{F}_\infty$ ist $A \cap \{ \tau \le t \} = \begin{cases} A & \text{falls } t \ge t_0 \\ \emptyset & \text{falls } t < t_0 \end{cases} \implies \mathcal{F}_\tau = \bigcap_{t \ge t_0} \mathcal{F}_t = \mathcal{F}_{t_0}$

d.) Für $A \in \mathcal{F}_\infty$ gilt: $A \cap \{ \tau \le t \} = A \cap \{ \sigma \le t \} \cap \{ \tau \le t \}$, da $\{ \tau \le t \} \subset \{ \sigma \le t \}$

Ist $A \in \mathcal{F}_\tau \implies A \cap \{ \sigma \le t \} \in \mathcal{F}_t \implies A \cap \{ \tau \le t \} \equiv A \cap \{ \sigma \le t \} \cap \{ \tau \le t \} \in \mathcal{F}_t$

q.e.d.

Ebenfalls im Hinblick auf Abschnitt 12 die

Proposition 10.10: Ist X $\mathbb{F}$-progressiv und τ eine $\mathbb{F}$-Stoppzeit, so gelten:

a.) X_τ ist $\mathcal{F}_\tau - \sigma(\mathcal{S} \cup \{ e \})$-meßbar.

b.) Der Prozeß $(X_{\tau \wedge t})_{t \ge 0}$ ist $\mathbb{F}$-progressiv.

Beweis: a.) X $\mathbb{F}$-progressiv $\implies X$ ist nach (10.2) meßbar.

$\implies X_\tau$ ist nach (10.4) a.) eine Zufallsvariable, d.h. $\{ X_\tau \in B \} \in \mathcal{F} \quad \forall B \in \sigma(\mathcal{S} \cup \{ e \})$.

<u>Fall 1:</u> $B \in \mathcal{S} \implies \{ X_\tau \in B \} \cap \{ \tau \le t \} = \{ X_{\tau \wedge t} \in B \} \cap \{ \tau \le t \}$ (leicht)

Nach b.) ist $\{ X_{\tau \wedge t} \in B \} \in \mathcal{F}_t$, da aus der dortigen Produktmeßbarkeit die Meßbarkeit der t-Schnitte folgt.

<u>Fall 2:</u> $B \in \sigma(\mathcal{S} \cup \{ e \}) \setminus \mathcal{S}$, d.h. $B = A \cup \{ e \}$ für ein $A \in \mathcal{S}$.

$\implies \{ X_\tau \in B \} = \{ X_\tau \in A \} \cup \{ X_\tau = e \} = \{ X_\tau \in A \} \cup \{ \tau = \infty \}$.

$\{ \tau = \infty \} \in \mathcal{F}_\tau$, da $\{ \tau = \infty \} \cap \{ \tau \le t \} = \emptyset \quad \forall t \ge 0$.

$\{ X_\tau \in A \} \in \mathcal{F}_\tau$ gemäß Fall 1.

b.) Nach (10.8) a.) ist $\tau \wedge t$ $\mathcal{F}_t - \mathcal{B}$-meßbar, da τ aufgrund von (10.5) eine $\mathbb{F}_+$-Stoppzeit ist.

Setze $(D, \mathcal{D}) := ([0,t] \times \Omega, [0,t] \cap \mathcal{B} \otimes \mathcal{F}_t)$. Die Abbildung $g : D \to S$ definiert durch $g(s, \omega) := X_{\tau(\omega) \wedge s}(\omega)$ ist die Komposition von $f : D \to D$ und $h : D \to S$, wo $f(s, \omega) := \big(\tau(\omega) \wedge s, \omega\big)$ sowie $h(s, \omega) := X_s(\omega)$.

Dabei ist f $\mathcal{D} - \mathcal{D}$-meßbar, da $\Pi_2 : D \to \Omega$ nach Definition von $\otimes$ $\mathcal{D} - \mathcal{F}_t$-meßbar und die Abbildung $f^* : D \to [0,t]$ mit $f^*(s, \omega) := \tau(\omega) \wedge s$ $\mathcal{D} - [0,t] \cap \mathcal{B}$-meßbar ist

($\{ \tau(\omega) \wedge s \le r \} = [0,r] \times \Omega \cup [0,t] \times \{ \tau \le r \}$ für $r \in [0,t]$)

Wegen $f = (f^*, \Pi_2)$ folgt die behauptete Meßbarkeit von f aus W-Theorie 8.10 .

h ist zudem $\mathcal{D} - \mathcal{S}$-meßbar, da X progressiv ist.

q.e.d.

Analog zur Einführung der σ-Algebra $\mathcal{F}_{t+}$ zu $\mathcal{F}_t$ nun

Definition 8: Ist τ eine $\mathbb{F}_+$-Stoppzeit, so wird die *σ-Algebra $\mathcal{F}_{\tau+}$ der $(\tau+)$-Vergangenheit* definiert als σ-Algebra der τ-Vergangenheit bzgl. $\mathbb{F}_+$. Mit anderen Worten:
$\mathcal{F}_{\tau+} := \{ A \in \mathcal{F}_\infty : A \cap \{ \tau \le t \} \in \mathcal{F}_{t+} \quad \forall t \ge 0 \}$.

Bemerkung 10.11: Ist τ eine $\mathbb{F}_+$-Stoppzeit, so gilt $\mathcal{F}_{\tau+} = \bigcap_{\epsilon>0} \mathcal{F}_{\tau+\epsilon}$

Beweis: "$\subseteq$": Sei $A \in \mathcal{F}_{\tau+}$. Annahme: $\exists\, \epsilon > 0$ mit $A \notin \mathcal{F}_{\tau+\epsilon}$
$\Longrightarrow \exists\, t_0 \ge 0$ mit $A \cap \{ \tau \le t_0 - \epsilon \} \notin \mathcal{F}_{t_0}$
$A \in \mathcal{F}_{\tau+} \Longrightarrow A \cap \{ \tau \le t_0 - \epsilon \} \in \mathcal{F}_{(t_0-\epsilon)+} = \bigcap_{s>t_0-\epsilon} \mathcal{F}_s \subset \mathcal{F}_{t_0}$ und somit Widerspruch.

"$\supseteq$": Sei $A \in \bigcap_{\epsilon>0} \mathcal{F}_{\tau+\epsilon}$. Annahme: $A \notin \mathcal{F}_{\tau+}$
$\Longrightarrow \exists\, t_0 \ge 0$ mit $A \cap \{ \tau \le t_0 \} \notin \mathcal{F}_{t_0+} = \bigcap_{s>t_0} \mathcal{F}_s \Longrightarrow \exists\, s_0 > t_0$ mit $A \cap \{ \tau \le t_0 \} \notin \mathcal{F}_{s_0}$
$A \in \bigcap_{\epsilon>0} \mathcal{F}_{\tau+\epsilon} \Longrightarrow A \in \mathcal{F}_{\tau+\epsilon} \quad \forall \epsilon > 0 \Longrightarrow A \cap \{ \tau \le t - \epsilon \} \in \mathcal{F}_t \quad \forall t \ge 0, \epsilon > 0.$
Mit $t := s_0$ und $\epsilon := s_0 - t_0$ ergibt sich ein Widerspruch.

q.e.d.

Als letztes noch eine Charakterisierung der Menge $\mathcal{F}_{\tau+}$:

Lemma 10.12: Für eine $\mathbb{F}_+$-Stoppzeit τ hat man die Aussagen

a.) $\mathcal{F}_{\tau+} = \{ A \in \mathcal{F}_\infty : A \cap \{ \tau < t \} \in \mathcal{F}_t \quad \forall t \ge 0 \}$

b.) Ist τ sogar eine $\mathbb{F}$-Stoppzeit, so folgt $\mathcal{F}_\tau \subset \mathcal{F}_{\tau+}$

Beweis: a.) Der Beweis verläuft analog zu dem von (10.5) (iii)$\Longleftrightarrow$(iv), man schneide dort nur mit $A \in \mathcal{F}_\infty$.

b.) Dies ist klar, da $\mathcal{F}_t \subset \mathcal{F}_{t+}$

q.e.d.

Soweit diese Einstimmung auf die Theorie der zeitstetigen Prozesse.
Es folgt nun, nachdem in Abschnitt 5 schon als Spezialfall eines Markoff-Prozesses die Markoff-Kette behandelt wurde, eine Ausdehnung der Theorie dieser Art von Prozessen.
Da sich die Terminologien in beiden Fällen entsprechen, schlage man für deren Interpretationen in Abschnitt 5 nach.

Abschnitt 11: Markoff-Prozesse

Sei im folgenden ein metrischer Raum (S, d) gegeben, welcher mit der Borel-σ-Algebra $\mathcal{S}$ versehen ist.
Ein wesentliches Element der Betrachtungen in den Abschnitten 5 und 6 war die interne Markoff-Familie.
Nun soll der allgemeine Begriff der Markoff-Familie eingeführt werden, dazu benötigt man:

Definition 1: Sei X ein stochastischer Prozeß auf $(\Omega, \mathcal{F})$. Dann heißt $\{ \vartheta_s : s \ge 0 \}$ eine Familie von *Shift-Operatoren*, falls die ϑ_s Abbildungen von Ω in sich sind mit der Eigenschaft

(11.1) $X_{t+s}(\omega) = X_s\big(\vartheta_t(\omega)\big) \quad \forall\, \omega \in \Omega\,,\ s, t \ge 0$

Damit hat man

Definition 2: $(X, \{ \mathrm{P}_x \}_{x \in S})$ ist eine *Markoff-Familie*, falls folgende Bedingungen erfüllt sind:
X ist ein stochastischer Prozeß auf $(\Omega, \mathcal{F})$ mit Zustandsraum S, $\{ \mathrm{P}_x \}_{x \in S}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathcal{F})$, so daß mit der kanonischen Filterung $\mathbb{F}^{\mathrm{X}}$ gilt:

(i) $\mathrm{P}_x[X_0 = x] = 1 \quad \forall\, x \in S$

(ii) Die Abbildung $x \mapsto \mathrm{P}_x[X_s \in B]$ ist $\mathcal{S}-\mathcal{B}$-meßbar $\forall\, B \in \mathcal{S}$, $s \geq 0$.

(iii) $\mathrm{P}_{X_t}[X_s \in B]$ ist eine Version von $\mathrm{P}_x[X_{t+s} \in B \mid \mathcal{F}_t^X] \quad \forall\, x \in S$, $t, s \geq 0$, $B \in \mathcal{S}$.

(iv) Es existiert eine Familie von Shift-Operatoren $\{\vartheta_s : s \geq 0\}$.

Bemerkungen: 1.) Eine Familie von Shift-Operatoren erlaubt eine Berechnung des Zustands eines Systems im Zeitpunkt $t+s$ schon im Zeitpunkt s.

2.) Man beachte, daß die interne Markoff-Familie aus Abschnitt 5 dieser Definition genügt (nur mit diskreter Zeitparametermenge): Punkt (i) ist dort mit $\mathrm{P}_i[X_0^0 = i] = 1$, $i \in S$, formuliert. Punkt (ii) ist dort wegen der Abzählbarkeit des Zustandsraumes erfüllt, (iii) wurde damals in (5.7) gezeigt. Der Punkt (iv) wird durch eine Indexverschiebung erhalten (X^0 war ein Projektionsprozeß !).

Die Untersuchung der Vergangenheitsbedingung in Definition 2 kann vereinfacht werden, man hat nämlich

Lemma 11.2: In Definition 2 kann Punkt (iii) ersetzt werden durch

(iii)$'$ $\mathrm{P}_{X_t}[X_s \in B]$ ist eine Version von $\mathrm{P}_x[X_{t+s} \in B \mid X_{t_0}, \ldots, X_{t_n}, X_t]$ $\forall\, s,t \geq 0$, $0 \leq t_0 < \ldots < t_n \leq t$, $n \in \mathbb{N}_0$

Beweis: (iii) $\Rightarrow$(iii)$'$ ist klar, da $\sigma(X_{t_0}, \ldots, X_{t_n}, X_t) \subset \mathcal{F}_t^X \quad \forall\, 0 \leq t_0 < \ldots < t_n \leq t$, $n \in \mathbb{N}_0$.

(iii)$'$ $\Rightarrow$(iii) folgt dadurch, daß $\mathcal{L} := \bigcup\limits_{0 \leq t_0 < \ldots < t_n \leq t} \sigma(X_{t_0}, \ldots, X_{t_n}, X_t)$ ein durchschnittstabiles Erzeugendensystem von $\mathcal{F}_t^X$ ist. $\mathcal{D} := \{A \in \mathcal{F}_t^X : \int\limits_A 1_{\{X_{t+s} \in B\}}\, d\mathrm{P}_x = \int\limits_A \mathrm{P}_{X_t}[X_s \in B]\, d\mathrm{P}_x\}$ ist ein Dynkinsystem (leicht). Aus $\mathcal{L} \subset \mathcal{D}$ folgt dann die Behauptung.

q.e.d.

Als ein Beispiel für eine Markoff-Familie folgt nun die

Konstruktion einer d-dimensionalen Brown'schen Familie

Mache dazu die folgenden Schritte:

1.) $C[0,\infty)^d := \bigtimes\limits_{j=1}^{d} C[0,\infty)$. $C[0,\infty)^d$ kann als die Menge $\{\omega : [0,\infty) \to \mathbb{R}^d \mid \omega \text{ ist stetig}\}$ aufgefaßt werden. $C[0,\infty)^d$ werde mit der σ-Algebra $\mathcal{L}_d := \bigotimes\limits_{j=1}^{d} \mathcal{B}(C[0,\infty))$ versehen.

2.) $W^{(i)} : C[0,\infty)^d \to C[0,\infty)$ bezeichne die i-te Projektion, $W_t^{(i)}$ die Auswertung derselben an der Stelle t, d.h. $W_t^{(i)}(\omega^{(1)}, \ldots, \omega^{(d)}) := \omega^{(i)}(t)$.

3.) $W := (W^{(1)}, \ldots, W^{(d)})$ ist die Identität auf $C[0,\infty)^d$, W_t das Analogon zu $W_t^{(i)}$: $W_t := (W_t^{(1)}, \ldots, W_t^{(d)})$.

4.) μ bezeichne das Wienermaß auf $C[0,\infty)$, $\mathrm{P}_0 := \bigtimes\limits_{j=1}^{d} \mu$.

5.) Nach der allgemeinen Konstruktion von Wahrscheinlichkeitsmodellen (vgl. W-Theorie §13 bzw. Abschnitt 2 oben) folgt somit: $W^{(1)}, \ldots, W^{(d)}$ sind unabhängig und $W^{(i)} \sim \mu$.
Nach den Ergebnissen aus Abschnitt 8 (speziell Definition 8 und (8.14)) gilt daher: $W^{(1)}, \ldots, W^{(d)}$ sind unabhängige 1-dimensionale standardisierte Brown'sche Bewegungen. Aus Defintion 8 in Abschnitt 8 und den Ergebnissen aus der W-Theorie über Unabhängigkeit (siehe dort §13, vor allem Satz 13.1), folgt relativ leicht: W ist eine d-dimensionale standardisierte Brown'sche Bewegung.

6.) Zu $x \in \mathbb{R}^d$ sei $x + W$ der translatierte Prozeß $(x + W_t)_{t \geq 0}$ und $\mathrm{P}_x := \mathrm{P}_0(x + W)^{-1}$, also ist P_x ein Wahrscheinlichkeitsmaß auf $\big(C[0,\infty)^d, \mathcal{L}_d\big)$. Somit gilt $\mathrm{P}_x W^{-1} \equiv \mathrm{P}_0(x + W)^{-1}$.

7.) $x + W$ hat offensichtlich die gleichen Zuwächse wie W.

8.) Für $x \in \mathbb{R}^d$ gilt: $\mathrm{P}_x[\,W_0 = x\,] \equiv \mathrm{P}_0(x+W)^{-1}[\,W_0 = x\,] = \mathrm{P}_0[\,W_0 \circ (x+W) = x\,]$
$= \mathrm{P}_0[\,x + W_0 = x\,] = \mathrm{P}_0[\,W_0 = 0\,] = \underline{1}$

9.) Als Familie von Shift-Operatoren $\{\,\vartheta_s : s \geq 0\,\}$ definiert man $(\vartheta_t\omega)(\,\cdot\,) := \omega(t + \cdot\,)$. ϑ_t ist eine Abbildung von $C[0,\infty)^d$ in sich, denn mit ω ist auch $\omega(t + \cdot\,)$ stetig. Für $s, t \geq 0$ hat man $W_s(\vartheta_t\omega) \equiv W_s\big(\omega(t + \cdot\,)\big) \equiv \omega(t + s) \equiv W_{t+s}(\omega)$.

Damit ist eine Brown'sche Familie gemäß nachstehender Definition konstruiert:

Definition 3: a.) Ist μ Wienermaß auf $C[0,\infty)$, so heißt $\bigtimes_{j=1}^{d} \mu$ *Wienermaß auf $C[0,\infty)^d$*.

b.) $(W, \{\,\mathrm{P}_x\,\}_{x \in \mathbb{R}^d})$ heißt *d-dimensionale Brown'sche Familie*, falls W ein stochastischer Prozeß mit Zustandsraum $\mathbb{R}^d$ und $\{\,\mathrm{P}_x\,\}_{x \in \mathbb{R}^d}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathcal{F})$ ist mit den Eigenschaften

(0) W hat stetige Pfade.

(i) $\mathrm{P}_x[\,W_0 = x\,] = 1 \quad \forall\, x \in \mathbb{R}^d$.

(ii) W hat stationäre und unabhängige Zuwächse unter $\mathrm{P}_x \quad \forall\, x \in \mathbb{R}^d$.

(iii) $W_t - W_s \sim \bigtimes_{j=1}^{d} \mathrm{N}(0, t-s) \quad \forall\, 0 \leq s \leq t$ unter $\mathrm{P}_x \quad \forall\, x \in \mathbb{R}^d$.

(iv) Es existiert eine Familie von Shift-Operatoren $\{\,\vartheta_s : s \geq 0\,\}$.

Bemerkung: In obiger Konstruktion ist Punkt (0) evident, Punkt (i) wurde in 8.) gezeigt. Die Bedingung (ii) folgt für W aus 5.) und 6.) wegen 7.), (iii) ebenso. (iv) wird von 9.) abgedeckt.

Es wird nun, wie oben kurz angedeutet, gezeigt, daß eine Brown'sche Familie eine Markoff-Familie ist. Dazu dient als Hilfsmittel

Definition 4: Sei $(\Omega, \mathcal{F}, \mathrm{P})$ ein Wahrscheinlichkeitsraum, $\mathcal{F}_1$ und $\mathcal{F}_2$ Unter-σ-Algebren. Dann sind *$\mathcal{F}_1$ und $\mathcal{F}_2$ unabhängig* $:\Longleftrightarrow A_1$ und A_2 sind unabhängig $\forall\, A_i \in \mathcal{F}_i\,,\ i = 1, 2$.

Ist Y eine Zufallsvariable auf $(\Omega, \mathcal{F}, \mathrm{P})$, so ist *$Y$ unabhänigig von $\mathcal{F}_1$* $:\Longleftrightarrow \sigma(Y)$ und $\mathcal{F}_1$ sind unabhängig.

Als weitere Vorbereitung

Lemma 11.3: Sei X ein d-dimensionaler stochastischer Prozeß auf $(\Omega, \mathcal{F}, \mathrm{P})$ mit unabhängigen und stationären Zuwächsen. Mit $Q(s, x, B) := \mathrm{P}[\,x + X_s - X_0 \in B\,]$ für $s \geq 0\,,\ x \in \mathbb{R}^d$ und $B \in \mathcal{B}_d$ gelten:

a.) Die Abbildung $x \mapsto Q(s, x, B)$ ist $\mathcal{B}_d - \mathcal{B}$-meßbar $\forall\, s \geq 0\,,\ B \in \mathcal{B}_d$.

b.) $X_{t+s} - X_t$ ist unabhängig von $\mathcal{F}_t^X \quad \forall\, s, t \geq 0$.

c.) $Q(s, X_t, B)$ ist eine Version von $\mathrm{P}[\,X_{t+s} \in B \mid \mathcal{F}_t^X\,] \quad \forall\, s \geq 0\,,\ B \in \mathcal{B}_d$.

Beweis: a.) Die Behauptung gilt für jede $\mathbb{R}^d$-wertige Zufallsvariable Y, d.h. $x \mapsto \mathrm{P}[\,x + Y \in B\,]$ ist meßbar. Die Abbildung $(x, \omega) \mapsto 1_B\big(x + Y(\omega)\big)$ ist nämlich $\mathcal{B}_d \otimes \mathcal{F} - \mathcal{F}$-meßbar. Daher folgt die Aussage sofort mit (1.2).

b.) Sei $0 < t_1 < \ldots < t_n < t$, $n \in \mathbb{N}$. Dann sind nach Voraussetzung die Zufallsvariablen $X_0, X_{t_1} - X_0, \ldots, X_t - X_{t_n}, X_{t+s} - X_s$ unabhängig. Nun gilt für alle $m \in \{1, \ldots, n\}$

$$X_{t_m} = X_0 + \sum_{j=1}^{m}(X_{t_j} - X_{t_{j-1}}) \quad \text{und weiter } X_t = X_0 + (X_{t_1} - X_0) + \ldots + (X_t - X_{t_m}).$$

Mithin ist auch $X_{t+s} - X_s$ unabhängig von $X_0, X_{t_1}, \ldots, X_{t_n}, X_t$.
Der Rest folgt mit der gleichen Argumentation wie im Beweis von (11.2).

c.) $\mathrm{P}[\,X_{t+s} \in B \mid \mathcal{F}_t^X\,] \equiv \mathrm{P}[\,X_t + (X_{t+s} - X_t) \in B \mid \mathcal{F}_t^X\,] = Q(s, X_t, B)$ nach (1.10) h.).
Denn es ist $Q(s,x,B) = \mathrm{P}[\,x + X_{t+s} - X_t \in B\,]$ wegen der stationären Zuwächse

$$= \mathrm{P}[\,x + X_{t+s} - X_t \in B \mid \mathcal{F}_t^X\,] \quad \text{nach dem Spezialfall aus (1.7).}$$

Setze in (1.10) h.) für festes s, B: $h(x,\omega) := Q(s,x,B)$, $X := X_t$, $Y := X_{t+s} - X_t$

$$\Omega_1 = \Omega_2 := \mathbb{R}^d,\ \mathcal{G} := \mathcal{F}_t^X,\ T(x,y) := 1_B(x+y).$$

□

Damit gelingt nun der Nachweis der obigen Behauptung:

<u>Propositon 11.4:</u> Eine d-dimensionale Brown'sche Familie ist eine Markoff-Familie.

<u>Beweis:</u> Verifikation der vier Punkte aus Definition 2:

(i) $\mathrm{P}_x[\,W_0 = x\,] = 1$ nach Punkt (i) aus Definition 3 $\forall\, x \in \mathbb{R}^d$

(ii) $x \mapsto \mathrm{P}_x[\,W_s \in B\,] = \mathrm{P}_x[\,x + W_s - W_0 \in B\,]$ nach Punkt (i)
$\qquad = \mathrm{P}_0[\,x + W_s - W_0 \in B\,]$ nach (iii) in Definition 3
$\qquad =: Q(s,x,B)$

(11.3) a.) liefert die Meßbarkeit.

(iii) erhält man aus obiger Gleichungskette in (ii) in Verbindung mit (11.3) c.).

(iv) ist klar

q.e.d.

Mit Hilfe von Markoff-Familien wird nun der allgemeine Fall eines Markoff-Prozesses definiert.
Dazu erinnere man sich, daß im zeitdiskreten Fall (Abschnitt 5) die Markoff-Eigenschaft mit Hilfe der Übergangswahrscheinlichkeiten p_{ij} formuliert wurde, es existierte eine Markoff-Familie (die kanonische bzw. interne) nach dem Satz von Ionescu-Tulcea. Im zeitstetigen Fall treten an die Stelle der p_{ij} zeitlich abhängige Übergangswahrscheinlichkeiten $p_{ij}(t)$ oder noch allgemeiner $p(t,x,B)$, wo $B \in \mathcal{S}$.
Das Problem der Existenz einer zugehörigen Markoff-Familie wird durch die zeitstetige Version des Satzes von Kolmogoroff (vgl. z.B. Ash, zeitdiskrete Version siehe (2.7)) gelöst. Wie in Abschnitt 2 schon bemerkt, wird dieser Satz hier nicht behandelt. Man umgeht ihn, indem man die Existenz einer Markoff-Familie voraussetzt. Die Übergangswahrscheinlichkeiten können nur in den seltensten Fällen explizit angegeben werden.

Nun zur

<u>Definition 5:</u> $(X, \mathbb{F})$ heißt (<u>*homogener*</u>) <u>*Markoff-Prozeß*</u> auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathrm{P})$ mit Zustandsraum $(S, \mathcal{S})$, falls

1.) X ist ein stochastischer Prozeß auf $(\Omega, \mathcal{F}, \mathrm{P})$.

2.) $\mathbb{F}$ ist eine Filterung in $\mathcal{F}$.

3.) X ist an $\mathbb{F}$ adaptiert.

4.) Es existiert eine Markoff-Familie $(X^0, \{\mathrm{P}_x\}_{x \in S})$ mit der Eigenschaft:
$\mathrm{P}_{X_t}[\,X_s^0 \in B\,]$ ist eine Version von $\mathrm{P}[\,X_{t+s} \in B \mid \mathcal{F}_t\,]$ $\forall\, s,t \geq 0$, $B \in \mathcal{S}$.

$(X^0, \{\mathrm{P}_x\}_{x \in S})$ heißt <u>*interne Markoff-Familie*</u>, $\mathrm{P}X_0^{-1}$ <u>*Startverteilung*</u>.

Ein Spezialfall zeigt das nächste Lemma auf:

Lemma 11.5: Ist $(X, \{P_x\}_{x \in S})$ eine Markoff-Familie, so ist für $x_0 \in S$ $(X, \mathbb{F}^X)$ unter P_{x_0} ein Markoff-Prozeß mit Startverteilung δ_{x_0} und interner Markoff-Familie $(X, \{P_x\}_{x \in S})$.
Hier ist mithin $(X, \{P_x\}_{x \in S})$ interne Markoff-Familie zu sich selbst.

Beweis: X ist ein stochastischer Prozeß, $\mathbb{F}^X$ eine Filterung in $\mathcal{F}$ und X ist an $\mathbb{F}^X$ adaptiert.
Ferner ist nach Definition 2 $P_{x_0}[X_0 = x_0] = 1$ und $P_{X_t}[X_s \in B]$ eine Version von $P_{x_0}[X_{t+s} \in B \mid \mathcal{F}_t^X]$ $\forall B \in \mathcal{S}$, $s, t \geq 0$ gemäß (iii) aus Definition 2.

q.e.d.

Nach der Definition einer standardisierten d-dimensionalen Brown'schen Bewegung in Abschnitt 8 nun

Definition 6: Sei $d \in \mathbb{N}$ und W ein an $\mathbb{F}$ adaptierter stochastischer Prozeß auf $(\Omega, \mathcal{F}, P)$ mit Zustandsraum $\mathbb{R}^d$. Dann heißt $(W, \mathbb{F})$ *d-dimesionale Brown'sche Bewegung*, falls gilt:

(0) W hat stetige Pfade.

(i) Für $t \geq 0, s > 0$ sind die Zuwächse $W_{t+s} - W_s$ unabhängig von $\mathcal{F}_t$ unter P.

(ii) Für $t \geq 0, s > 0$ gilt $W_{t+s} - W_t \sim \bigtimes_{j=1}^{d} N(0, s)$ unter P

(d.h. insbesondere: Die Zuwächse sind stationär.)

Mit $b \in \mathbb{R}^d$ und $\sigma \in \mathbb{R}^{d,d}$ nicht singulär (d.h. invertierbar) definiert man für $t \geq 0$ die Größe $X_t := tb + \sigma W_t$. Dann heißt $(X, \mathbb{F})$ *d-dimensionale Brown'sche Bewegung mit Drift b und Dispersionskoeffizient σ*. (Drift: Treibfaktor; Dispersion: Verbreitung, Zerstreuung)

Beispiel 11.6: Ist $W = (W^{(1)}, \ldots, W^{(d)})$ eine d-dimensionale Brown'sche Familie bzgl. $\{P_x\}_{x \in \mathbb{R}^d}$ und $\mathbb{F} := \mathbb{F}^W$, so ist für $x_0 \in \mathbb{R}^d$ $(W^{(i)}, \mathbb{F})$ eine 1-dimensionale Brown'sche Bewegung unter P_{x_0}, $i \in \{1, \ldots, d\}$.

Beweis: Es ist $W_{t+s}^{(i)} - W_t^{(i)} = \pi_i(W_{t+s} - W_t)$. Die Stetigkeit von π_i impliziert die Stetigkeit von $W_t^{(i)}$.
Die Unabhängigkeit von $W_{t+s}^{(i)} - W_t^{(i)}$ von $\mathcal{F}_t$ folgt aus W-Theorie 13.6 . $W_{t+s}^{(i)} - W_t^{(i)} \sim N(0, s)$ ist klar.

□

Eine d-dimensionale Brown'sche Bewegung erweist sich als ein Beispiel für einen Markoff-Prozeß. Dies formuliert genauer

Proposition 11.7: Eine d-dimensionale Brown'sche Bewegung $(X, \mathbb{F})$ auf $(\Omega, \mathcal{F}, P)$ mit Drift b und Dispersionskoeffizient σ ist ein Markoff-Prozeß mit interner Markoff-Familie $X^0 := (bt + \sigma W_t^0)_{t \geq 0}$ und $\{ P_{\sigma^{-1}x} \}_{x \in \mathbb{R}^d}$, wo $(W^0, \{ P_x \}_{x \in \mathbb{R}^d})$ eine d-dimensionale Brown'sche Familie ist.

Beweis: 1.) Eine d-dimensionale Brown'sche Bewegung $(W, \mathbb{F})$ auf $(\Omega, \mathcal{F}, P)$ ist ein Markoff-Prozeß.
Für $B \in \mathcal{B}_d$ und $s, t \geq 0$ gilt nämlich $P[W_{t+s} \in B \mid \mathcal{F}_t] \equiv P[W_t + (W_{t+s} - W_t) \mid \mathcal{F}_t]$

$$= Q(s.W_t, B)$$

wie in (11.3) mit $Q(s, x, B) := P[x + W_{t+s} - W_t \in B] \equiv \bigtimes_{i=1}^{d} N(x_i, s)[B] \equiv P_x[W_s^0 \in B]$

wo $(W^0, \{ P_x \}_{x \in \mathbb{R}^d})$ eine (z.B. wie oben konstruierte) Brown'sche Familie ist.

2.) $P[X_{t+s} \in B \mid \mathcal{F}_t] \equiv P[b(t+s) + \sigma W_{t+s} \in B \mid \mathcal{F}_t]$

$$= P[W_{t+s} \in \widetilde{B} \mid \mathcal{F}_t] \quad \text{wo} \quad \widetilde{B} := \sigma^{-1}\big(B - b(t+s)\big)$$

$$= P_{W_t}[W_s^0 \in \widetilde{B}] \text{ nach Schritt 1.)}$$

$$\stackrel{!}{=} Q_{X_t}[X_s^0 \in B]$$

Ein solches Q_x und X_s^0 sind noch zu finden.

$$
\begin{aligned}
\mathrm{P}_x[\,b(t+s)+\sigma W_s^0 \in B\,] &\equiv \mathrm{P}_x[\,W_s^0 \in \widetilde{B}\,] = \mathrm{P}_x[\,W_s^0 - W_0^0 \in \widetilde{B}-x\,] \\
&\equiv \mathrm{P}_0[\,W_s^0 - W_0^0 \in \widetilde{B}-x\,] \ \text{ nach (iii) aus Definition 3} \\
&= \mathrm{P}_0[\,W_s^0 + x \in \widetilde{B}\,] \equiv \mathrm{P}_0[\,b(t+s)+\sigma(x+W_s^0) \in B\,] \quad (+) \\
&\equiv \mathrm{P}_0[\,b\,s+\sigma\big(\sigma^{-1}(b\,t+\sigma x)+W_s^0\big) \in B\,] \\
&= \mathrm{P}_{\sigma^{-1}(b\,t-\sigma x)}[\,b\,s+\sigma W_s^0 \in B\,] \ \text{ analog zu vorhin} \\
&\equiv Q_{b\,t+\sigma x}[\,X_s^0 \in B\,]
\end{aligned}
$$

wo $Q_x := \mathrm{P}_{\sigma^{-1}x}$, $X_s^0 := b\,s+\sigma W_s^0$.

3.) Es bleibt noch z.z.: $(X^0, \{\,Q_x\,\}_{x\in\mathbb{R}^d})$ ist eine Markoff-Familie.

Für $x \in \mathbb{R}^d$, $B \in \mathcal{B}_d$ und $s \geq 0$ gelten:

$Q_x[\,X_0^0 = x\,] \equiv Q_x[\,\sigma W_0^0 = x\,] \equiv \mathrm{P}_{\sigma^{-1}x}[\,W_0^0 = \sigma^{-1}x\,] = 1$ sowie

$$
\begin{aligned}
Q_x[\,X_s^0 \in B\,] &= \mathrm{P}_0[\,b\,s + x + \sigma W_s^0 \in B\,] \ \text{ nach } (+) \\
&= \mathrm{P}_0[\,x+\sigma W_s^0 \in \widetilde{B}\,]\,,\ \widetilde{B} := B - b\,s
\end{aligned}
$$

Damit ist die Abbildung $x \mapsto Q_x[\,X_s^0 \in B\,]$ nach (11.3) a.) meßbar. Obige Überlegungen in Schritt 2.) gelten auch für Q_x, X_{t+s}^0 und $\mathcal{F}_t^{X^0}$ anstelle von P, X_{t+s} und $\mathbb{F}$, da eine d-dimensionale Brown'sche Familie auch eine d-dimensionale Brown'sche Bewegung bzgl. kanonischer Filterung ist (verwende dazu (11.3) b.)). Mithin ist $Q_{X_t^0}[\,X_s^0 \in B\,]$ eine Version von $Q_x[\,X_{t+s}^0 \in B \mid \mathcal{F}_t^X\,]$. Die Existenz einer Shift-Operatorenfamilie ist durch die Existenz einer dergleichen für W^0 gesichert.

q.e.d.

Am Schluß dieses Abschnitts werden allgemeine Ereignisse der Zukunft (d.h. nicht nur die Zustände in einem bestimmten Zeitpunkt $t+s$) betrachtet. Dazu nun einige nützliche Abkürzungen/Konventionen:

Definition 7: Mit $S^{[0,\infty)}$ wird die Menge aller Abbildungen von $[0,\infty)$ nach S bezeichnet, d.h. $S^{[0,\infty)} := \{\,\xi :\ [0,\infty) \xrightarrow{\xi} S\,\}$. Wie schon von früher her bekannt, stehe π_t , $t \geq 0$ für die Projektion auf die t-te Stelle (d.h. die Auswertung an der Stelle t).

Mithin: $\pi_t : S^{[0,\infty)} \to S$, $\pi_t(\xi) := \xi(t)$. Mit $\bigotimes_{0\leq t<\infty} \mathcal{S}$ werde die Produkt-σ-Algebra der π_t bezeichnet, d.h. $\bigotimes_{0\leq t<\infty} \mathcal{S} := \sigma(\pi_t, t \geq 0)$. Ist X ein stochastischer Prozeß, so stehe $X_{t+\bullet}$ für den von t ausgehenden zukünftigen Prozeß $(X_{t+s})_{s\geq 0}$ (d.h. man startet mit den Betrachtungen/Beobachtungen im Zeitpunkt t).

Der Zusammenhang zwischen den allgemeinen Ereignissen der Zukunft und den punktuellen zeigt

Lemma 11.8: Sei $(\Omega, \mathcal{F})$ ein Meßraum. Dann sind für $X := (X_t)_{t\geq 0}$ mit $X_t : \Omega \to S$ äquivalent:

(i) X_t ist $\mathcal{F} - \mathcal{S}$-meßbar $\forall\, t \geq 0$ (mithin X ein stochastischer Prozeß)

(ii) $X : \Omega \to S^{[0,\infty)}$ ist $\mathcal{F} - \bigotimes_{0\leq t<\infty} \mathcal{S}$-meßbar.

(iii) $X : \Omega \to D$ ist $\mathcal{F} - D \cap \bigotimes_{0\leq t<\infty} \mathcal{S}$-meßbar $\forall\, D$ mit $X(\Omega) \subset D \subset S^{[0,\infty)}$.

Beweis: (i)$\Rightarrow$(ii): $X = (X_t)_{t\geq 0}$. Damit folgt (ii) aus der Tatsache, daß $\bigotimes\limits_{0\leq t<\infty} \mathcal{S} = \bigcup\limits_{\substack{T'\subset[0,\infty)\\ T'\text{abzählbar}}} \bigotimes\limits_{s\in T'} \mathcal{S}$, mit W-Theorie 17.6.

(ii)$\Rightarrow$(iii): $A \in \bigotimes\limits_{0\leq t<\infty} \mathcal{S} \Longrightarrow X^{-1}(D\cap A) = X^{-1}\big(X(\Omega)\cap A\big) = X^{-1}(A) \in \mathcal{F}$.

(iii)$\Rightarrow$(i): Setze $D := S^{[0,\infty)} \Longrightarrow$ Beh.

q.e.d.

Nun zur angekündigten Betrachtung von allgemeinen zukünftigen Ereignissen:

Satz 11.9: Sei $(X,\mathbb{F})$ ein Markoff-Prozeß auf $(\Omega,\mathcal{F},\mathrm{P})$ mit interner Markoff-Familie $(X^0,\{\,\mathrm{P}_x\,\}_{x\in S})$.

Dann gelten für alle $\Gamma \in \bigotimes\limits_{0\leq t<\infty} \mathcal{S}$:

(i) Die Abbildung $x \mapsto \mathrm{P}_x[\,X^0\in\Gamma\,]$ ist meßbar, d.h. $(x,A)\mapsto \mathrm{P}_x[\,A\,]$ ist eine Übergangswahrscheinlichkeit von S nach $\mathcal{F}_\infty^{X^0}$.

(ii) $\mathrm{P}_{X_t}[\,X^0\in\Gamma\,]$ ist eine Version von $\mathrm{P}[\,X_{t+\bullet}\in\Gamma\mid\mathcal{F}_t\,] \quad \forall\, t\geq 0$.

(iii) $\mathrm{P}[\,X\in\Gamma\,] = \int\limits_S \nu(dx)\,\mathrm{P}_x[\,X^0\in\Gamma\,]$, wo $\nu := \mathrm{P}X_0^{-1}$.

$\mathrm{P}X^{-1}$ ist daher durch die Startverteilung und die interne Markoff-Familie gegeben.

Beweis: Die Menge $\{\,\Gamma\in \bigotimes\limits_{0\leq t<\infty}\mathcal{S} :$ (i) und (ii) gelten $\}$ ist ein Dynkin-System (leicht).

Mithin genügt es, (i) und (ii) für endlich-dimensionale Zylindermengen der Form

$(*)\quad \Gamma = \{\,\xi\in S^{[0,\infty)} : \xi(s_1)\in B_1,\ldots,\xi(s_n)\in B_n\,\} \equiv (\pi_{s_1},\ldots,\pi_{s_n})^{-1}(B_1\times\ldots\times B_n)$

zu zeigen. ($B_j\in\mathcal{S}$, $s_j\geq 0$, $j=1,\ldots,n$, $n\in\mathbb{N}_0$). Zeige noch allgemeiner:

(11.10) $\forall\, n\in\mathbb{N}$, $0\leq s_1<\ldots<s_n<\infty$, $f_m : S\to[0,\infty)$ meßbar, $1\leq m\leq n$, gelten:

(i)′ $x\mapsto \mathrm{E}_x[\,\prod\limits_{m=1}^n f_m(X^0_{s_m})\,]$ ist meßbar.

(ii)′ $\mathrm{E}_{X_t}[\,\prod\limits_{m=1}^n f_m(X^0_{s_m})\,]$ ist eine Version von $\mathrm{E}[\,\prod\limits_{m=1}^n f_m(X_{t+s_m})\mid\mathcal{F}_t\,]$, $t\geq 0$.

Auf die Form $(*)$ kommt man durch die Verwendung der speziellen Funktionen $f_m := 1_{B_m}$.
Der Beweis von (11.10) wird nun induktiv geführt:

$n=1$: Der Beweis verläuft sehr leicht mit dem Beweisprinzip für Integrale. Für $f=1_B$ mit $B\in\mathcal{S}$ verwende man Definition 2 um (i)′ zu zeigen sowie Definition 5 für (ii)′.
(ii)′ gilt dabei auch für E_X, X^0 und $\mathbb{F}^{X^0}$ anstelle von E, X und $\mathbb{F}$ $\big($ mit (11.5) $\big)$.

$n\to n+1$: Sei dazu $A\in\mathcal{F}_t\subset\mathcal{F}_{t+s_m}$, $m\leq n$.

$$\begin{aligned}
\int\limits_A \prod_{m=1}^{n+1} f_m(X_{t+s_m})\,d\mathrm{P} &\equiv \int\limits_A \mathrm{E}[\,\prod_{m=1}^n f_m(X_{t+s_m})\cdot f_{n+1}(X_{t+s_{n+1}})\mid\mathcal{F}_{t+s_n}\,]\,d\mathrm{P}\\
&\overset{(1.10)g.)}{=} \int\limits_A \prod_{m=1}^n f_m(X_{t+s_m})\cdot \mathrm{E}[\,f_{n+1}(X_{t+s_{n+1}})\mid\mathcal{F}_{t+s_n}\,]\,d\mathrm{P}\\
&\overset{\text{Fall } n=1}{=} \int\limits_A \prod_{m=1}^n f_m(X_{t+s_m})\cdot \underbrace{\mathrm{E}_{X_{t+s_n}}[\,f_{n+1}(X^0_{s_{n+1}-s_n})\,]}_{=:\ g(X_{t+s_n})}\,d\mathrm{P}\\
&= \int\limits_A \mathrm{E}_{X_t}[\,\prod_{m=1}^n f_m(X^0_{s_m})\cdot g(X^0_{s_n})\,]\,d\mathrm{P}
\end{aligned}$$

Der letzte Schritt gilt nach Induktionsvoraussetzung mit $\widetilde{f_n} := f_n \cdot g$ statt f_n g ist nämlich nach Induktionsanfang meßbar, somit auch $\widetilde{f_n}$. Dabei ist nach obiger Bemerkung im Fall $n = 1$ und (1.10) g.) (analog zu oben):

$$\begin{aligned} \mathrm{E}_x[\prod_{m=1}^{n} f_m(X^0_{s_m}) \cdot g(X^0_{s_n})] &= \mathrm{E}_x[\prod_{m=1}^{n} f_m(X^0_{s_m}) \cdot \mathrm{E}_x[f_{n+1}(X^0_{s_n+1}) \mid \mathcal{F}^{X^0}_{s_n}]] \\ &= \mathrm{E}_x[\prod_{m=1}^{n+1} f_m(X^0_{s_m})] \end{aligned}$$

Die Meßbarkeit der Abbildung $x \mapsto \mathrm{E}_x[\prod_{m=1}^{n+1} f_m(X^0_{s_m})]$ folgt aus der von $x \mapsto \mathrm{E}_x[\prod_{m=1}^{n} f_m(X^0_{s_m}) \cdot g(X^0_{s_n})]$, welche nach Induktionsvoraussetzung meßbar ist (wieder mit $\widetilde{f_n} := f_n \cdot g$ statt f_n).

(iii): Nach (i) ist $x \mapsto \mathrm{P}_x[X^0 \in \Gamma]$ meßbar, somit das Integral wohldefiniert.

$$\begin{aligned} \mathrm{P}[X \in \Gamma] \equiv \mathrm{E}\big[\mathrm{P}[X \in \Gamma \mid \mathcal{F}_0]\big] &= \mathrm{E}\big[\mathrm{P}_{X_0}[X^0 \in \Gamma]\big] \quad \text{nach (ii)} \\ &= \int_S \nu(dx)\, \mathrm{P}_x[X^0 \in \Gamma] \quad \text{nach Bildmaßformel} \end{aligned}$$

q.e.d.

Als direkte Folgerung:

Korollar 11.11: Sei $(X, \{\mathrm{P}_x\}_{x \in S})$ eine Markoff-Familie auf $(\Omega, \mathcal{F})$. Ist $f : S^{[0,\infty)} \to \overline{\mathbb{R}}$ $\bigotimes_{0 \le s < \infty} \mathcal{S} - \overline{\mathcal{B}}$-meßbar und nach unten beschränkt, so gilt:

$\mathrm{E}_{X_t}[f \circ X]$ ist eine Version von $\mathrm{E}_x[f \circ X \circ \vartheta_t \mid \mathcal{F}_t]$ $\quad \forall t \ge 0, x \in S$.

(d.h. die Abhängigkeit vom Startpunkt x verschwindet vollständig, nur die Gegenwart X_t ist entscheident)

Beweis: $X \circ \vartheta_t \equiv (X_s)_{s \ge 0} \circ \vartheta_t \equiv (X_s \circ \vartheta_t)_{s \ge 0} = (X_{t+s})_{s \ge 0} \equiv X_{t+\bullet}$. Nun wieder das Beweisprinzip für Integrale: $f = 1_\Gamma$ mit $\Gamma \in \bigotimes_{0 \le t < \infty} \mathcal{S} \Longrightarrow f \circ X \circ \vartheta_t = 1_{\{X_{t+\bullet} \in \Gamma\}}$

Auf diese Funktion ist aber (11.9) (ii)$'$ anwendbar, da (ii)$'$ auch, wie dort bemerkt, für P_x statt P gilt. Der Rest ist einfach.

□

Abschließend noch ein Beispiel:

Beispiel: Sei $\tau \equiv \tau_B := \inf\{s \ge 0 : X_s \in B\}$, $B \in \mathcal{S}$.
$\Longrightarrow \tau \circ \vartheta_t = \inf\{s \ge 0 : X_s \circ \vartheta_t \in B\} = \inf\{s \ge 0 : X_{t+s} \in B\}$.
Somit ist $t + \tau \circ \vartheta_t = \inf\{s \ge t : X_s \in B\} =: \tau_t$. Damit erhält man $X_{\tau_t} = X_\tau \circ \vartheta_t$, denn $X_\tau \circ \vartheta_t = X_{\tau \circ \vartheta_t}(\vartheta_t) = X_{t + \tau \circ \vartheta_t} = X_{\tau_t}$.