

Teil I
Markov-Ketten auf abzählbaren
Zustandsräumen

Nach einem einjährigen Kurs über Wahrscheinlichkeitstheorie, in dem wir neben den notwendigen fundamentalen Begriffsbildungen vor allem Folgen unabhängiger Zufallsgrößen und ihre Partialsummen bereits recht eingehend studiert haben, wenden wir uns nun einer Klasse von stochastischen Prozessen zu, die durch eine besonders einfache Abhängigkeitsstruktur gekennzeichnet sind und als *Markov-Ketten* bezeichnet werden. Aufgrund ihrer zahllosen Anwendungen (☞ Abschnitte 2.1 und 2.9) in nahezu allen Disziplinen, in denen stochastische Modelle eine Rolle spielen, kommt ihnen zugleich eine große Bedeutung zu. Mit Fug und Recht lässt sich sagen, dass Markov-Ketten zum Grundwissen eines jeden Stochastikers zählen.

Nach einer Einführung in die theoretischen Grundlagen für allgemeine Markov-Ketten in Kapitel 1, konzentrieren wir uns in Kapitel 2 auf das Studium *diskreter Markov-Ketten*, deren zusätzliche Eigenschaft in einem höchstens abzählbaren Zustandsraum besteht. Ein Kernthema nach Bereitstellung der notwendigen Begriffsbildungen und Hilfsmittel ist die Untersuchung ihres Grenzverhaltens, die in der Herleitung des Ergodensatzes mittels der Kopplungstechnik gipfelt. Darüber hinaus befassen wir uns mit dem wichtigen Begriff der (*zeitlichen*) *Reversibilität* und seiner Charakterisierung. Eine große Zahl von Beispielen unterstreicht die obige Aussage der vielseitigen Anwendbarkeit diskreter Markov-Ketten. Kapitel 3 darf als Zugabe angesehen werden und wendet sich nur an diejenigen Leser, die nach weiterführenden Aussagen über das Grenzverhalten ergodischer Markov-Ketten streben. Es stellt eine Reihe von Techniken aus verschiedenen Bereichen vor, mit denen Konvergenzraten im Ergodensatz hergeleitet werden können.

Kapitel 1

Markov-Ketten: Theoretische Grundlagen

Basierend auf den im Skriptum [3] dargestellten Grundlagen der Wahrscheinlichkeitstheorie, geben wir im vorliegenden Kapitel zunächst eine Einführung über Markov-Ketten mit beliebigen Zustandsräumen, bevor wir uns in den nachfolgenden Kapiteln auf solche mit abzählbaren Zustandsräumen, genannt *diskrete* Markov-Ketten, konzentrieren werden. Den Grund für den allgemeinen und recht trockenen Einstieg, der zugegebenermaßen zu den eher freudlosen Beschäftigungen zählt, bildet die Tatsache, dass Definition und fundamentale Eigenschaften dieser Prozesse von der Art ihrer Zustandsräume vollständig unabhängig sind. Darüber hinaus vereinfacht sich die Darstellung im Falle der Abzählbarkeit nur unwesentlich, so dass eine Beschränkung unter diesem Aspekt eher künstlich erschiene. Die Begriffe “Filtration” und “Stopzeit”, die generell für die Untersuchung stochastischer Prozesse von großer Bedeutung sind, stellen wir in dem für unsere Zwecke gebotenen Umfang in Abschnitt 1.3 vor.

1.1 Definitionen und grundlegende Eigenschaften

Betrachten wir eine stochastische Folge $M = (M_n)_{n \geq 0}$ auf einem W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$, wobei jedes M_n Werte in einem beliebigen messbaren Raum $(\mathcal{S}, \mathfrak{S})$ annimmt. Die Elemente von \mathcal{S} bezeichnen wir als *Zustände* der Folge und \mathcal{S} selbst, oder präziser auch $(\mathcal{S}, \mathfrak{S})$, als ihren *Zustandsraum*. Den Folgenindex interpretieren wir als *Zeitparameter*, so dass M_n den Zustand der Folge zum Zeitpunkt n angibt. Sei $P_0 := \mathbb{P}^{M_0}$ die *Anfangsverteilung* von M und ferner vorausgesetzt, dass sämtliche bedingten Verteilungen

$$P_{n+1}((s_0, \dots, s_n), \cdot) := \mathbb{P}^{M_{n+1} | (M_0, \dots, M_n) = (s_0, \dots, s_n)}, \quad (1.1)$$

in diesem Kontext auch *Übergangskerne* genannt, existieren. Nach dem Satz von Ionescu Tulcea [3, Satz 54.4] ist die Verteilung von M durch $(P_n)_{n \geq 0}$ eindeutig bestimmt, und zwar vermöge

$$\mathbb{P}^M = \bigotimes_{n \geq 0} P_n. \quad (1.2)$$

Im folgenden wollen wir uns mit Folgen beschäftigen, die eine besonders einfache Abhängigkeitsstruktur besitzen, die durch die sogenannte *Markov-Eigenschaft* beschrieben wird.

Definition 1.1. Eine stochastische Folge $M = (M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{G})$ heißt *Markov-Kette (MK)*, wenn sie die *Markov-Eigenschaft* besitzt, definiert durch

$$\mathbb{P}^{M_{n+1}|M_0, \dots, M_n} = \mathbb{P}^{M_{n+1}|M_n} \quad \mathbb{P}\text{-f.s.} \quad (1.3)$$

für alle $n \geq 0$, oder auch

$$\mathbb{P}(M_{n+1} \in A | M_0, \dots, M_n) = \mathbb{P}(M_{n+1} \in A | M_n) \quad \mathbb{P}\text{-f.s.} \quad (1.3')$$

für alle $n \geq 0$ und $A \in \mathfrak{G}$. Hängen die bedingten Verteilungen nicht von n ab, existiert also ein stochastischer Kern P von $(\mathcal{S}, \mathfrak{G})$ nach $(\mathcal{S}, \mathfrak{G})$, so dass

$$\mathbb{P}^{M_{n+1}|M_n} = P(M_n, \cdot) \quad \mathbb{P}\text{-f.s.} \quad (1.4)$$

für alle $n \geq 0$, bezeichnet man die MK ferner als *zeitlich homogen*. Eine Markov-Kette mit abzählbarem bzw. endlichem Zustandsraum heißt *diskrete* bzw. *endliche Markov-Kette (DMK/EMK)*.

Eine Folge M genügt also der Markov-Eigenschaft, wenn die Verteilung ihres Zustands M_{n+1} zum Zeitpunkt $n+1$ bedingt unter der Vorgeschichte M_0, \dots, M_n immer nur vom gegenwärtigen Zustand M_n abhängt. Sie ist außerdem zeitlich homogen, wenn diese bedingte Verteilung nicht von n abhängt, wenn also die Wahrscheinlichkeit für den Übergang von einem Zustand $s \in \mathcal{S}$ in eine Menge $A \in \mathfrak{G}$ nicht davon abhängt, wann dieser Übergang stattfindet. Man spricht dann auch von einer MK mit *stationären Übergangswahrscheinlichkeiten*.

Gegeben die Markov-Eigenschaft, folgt in (1.1) offenkundig

$$P_{n+1}((s_0, \dots, s_n), \cdot) = P_{n+1}(s_n, \cdot)$$

für alle $n \geq 0$ und im Falle der zeitlichen Homogenität gemäß (1.3) weiter

$$\mathbb{P}^{M_{n+1}|M_n=s} = P_{n+1}(s, \cdot) = P(s, \cdot)$$

für alle $n \geq 0$ und \mathbb{P}^{M_n} -fast alle $s \in \mathcal{S}$. Wie der Leser nun leicht erkennt, ist die Verteilung von M dann vollständig durch ihre Anfangsverteilung $P_0 = \mathbb{P}^{M_0}$ und ihren (1-Schritt-)Übergangskern P determiniert, weil nämlich unter Hinweis auf (1.2)

$$\mathbb{P}^M = P_0 \otimes \left(\bigotimes_{n=1}^{\infty} P \right) = P_0 \otimes P^{\infty}.$$

Im Fall eines abzählbaren Zustandsraums ist P sogar bereits durch die Elementarwahrscheinlichkeiten

$$p_{ij} := P(i, \{j\}), \quad i, j \in \mathcal{S},$$

genannt (1-Schritt-)Übergangswahrscheinlichkeiten, vollständig festgelegt. Sie werden in der sogenannten Übergangsmatrix $P := (p_{ij})_{i,j \in \mathcal{S}}$ zusammengefasst (wir benutzen somit dasselbe Symbol für Übergangskern und Übergangsmatrix).

Lemma 1.2. Gegeben eine MK $M = (M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$, gilt für jedes $n \geq 0$ und jede $\mathbb{P}^{M_{n+1}}$ -quasi-integrierbare Funktion $f : \mathcal{S} \rightarrow \overline{\mathbb{R}}$

$$\mathbb{E}(f(M_{n+1}) | M_0, \dots, M_n) = \mathbb{E}(f(M_{n+1}) | M_n) \quad \mathbb{P}\text{-f.s.} \quad (1.5)$$

Ist M außerdem zeitlich homogen mit Übergangskern P , folgt weiter

$$\mathbb{E}(f(M_{n+1}) | M_n) = \int_{\mathcal{S}} f(s) P(M_n, ds) \quad \mathbb{P}\text{-f.s.} \quad (1.6)$$

für alle $n \geq 0$.

Beweis. Unter Benutzung von [3, Satz 53.6] ergibt sich (1.5) vermöge

$$\begin{aligned} \mathbb{E}(f(M_{n+1}) | M_0, \dots, M_n) &= \int_{\mathcal{S}} f(s) \mathbb{P}^{M_{n+1} | (M_0, \dots, M_n)}(ds) \\ &= \int_{\mathcal{S}} f(s) \mathbb{P}^{M_{n+1} | M_n}(ds) \\ &= \mathbb{E}(f(M_{n+1}) | M_n), \end{aligned}$$

woraus im Falle der zeitlichen Homogenität unter Benutzung von (1.4) sofort (1.6) folgt. \square

Kombiniert man Lemma 1.2 mit einer Induktion über k , so erhält man leicht die folgenden intuitiv zu erwartenden Verallgemeinerungen der Markov-Eigenschaft (1.3) sowie von (1.4):

Satz 1.3. Gegeben eine MK $M = (M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$, gilt

$$\begin{aligned} \mathbb{P}^{(M_n, \dots, M_{n+k}) | M_0, \dots, M_n} &= \mathbb{P}^{(M_n, \dots, M_{n+k}) | M_n} \\ &= \delta_{M_n} \otimes \left(\bigotimes_{i=1}^k P_{n+i} \right) \quad \mathbb{P}\text{-f.s.} \end{aligned} \quad (1.7)$$

– mit P_1, P_2, \dots gemäß (1.1) – und daher insbesondere

$$\mathbb{P}^{M_{n+k} | M_0, \dots, M_n} = \mathbb{P}^{M_{n+k} | M_n} \quad \mathbb{P}\text{-f.s.}$$

für alle $k, n \geq 0$. Ist M außerdem zeitlich homogen, d.h. $P_1 = P_2 = \dots = P$, folgt weiter

$$\mathbb{P}^{(M_n, \dots, M_{n+k})|M_n} = \delta_{M_n} \otimes P^k \quad \mathbb{P}\text{-f.s.} \quad (1.8)$$

für alle $k, n \geq 0$ sowie die Existenz stochastischer Kerne $P^{(k)}, k \geq 0$, von $(\mathcal{S}, \mathfrak{S})$ nach $(\mathcal{S}, \mathfrak{S})$, so dass

$$\mathbb{P}^{M_{n+k}|M_n} = P^{(k)}(M_n, \cdot) \quad \mathbb{P}\text{-f.s.}$$

für alle $k, n \geq 0$, wobei offensichtlich $P^{(0)}(s, \cdot) = \delta_s$ und $P^{(1)} = P$.

Beweis. Wir demonstrieren das Vorgehen durch einen exemplarischen Beweis von (1.7) und (1.8) für $k = 2$. Zur Abkürzung setzen wir $M_{0:n} = (M_0, \dots, M_n)$. Dann gilt für alle $n \geq 0, A_0, A_1, A_2 \in \mathfrak{S}$ sowie $\mathbb{P}^{M_{0:n}}$ -fast alle $(s_0, \dots, s_n) \in \mathcal{S}^{n+1}$

$$\begin{aligned} & \mathbb{P}(M_{n+2} \in A_2, M_{n+1} \in A_1, M_n \in A_0 | M_{0:n} = (s_0, \dots, s_n)) \\ &= \delta_{s_n}(A_0) \mathbb{P}(M_{n+2} \in A_2, M_{n+1} \in A_1 | M_{0:n} = (s_0, \dots, s_n)) \\ &= \delta_{s_n}(A_0) \mathbb{E}(\mathbf{1}_{\{M_{n+1} \in A_1\}} \mathbb{P}(M_{n+2} \in A_2 | M_{0:n+1}) | M_{0:n} = (s_0, \dots, s_n)) \\ &= \delta_{s_n}(A_0) \mathbb{E}(\mathbf{1}_{\{M_{n+1} \in A_1\}} \mathbb{P}(M_{n+2} \in A_2 | M_{n+1}) | M_{0:n} = (s_0, \dots, s_n)) \\ &= \delta_{s_n}(A_0) \mathbb{E}(\mathbf{1}_{\{M_{n+1} \in A_1\}} \mathbb{P}(M_{n+2} \in A_2 | M_{n+1}) | M_n = s_n) \quad (1.9) \\ &= \delta_{s_n}(A_0) \mathbb{P}(M_{n+2} \in A_2, M_{n+1} \in A_1 | M_n = s_n) \\ &= \mathbb{P}(M_{n+2} \in A_2, M_{n+1} \in A_1, M_n \in A_0 | M_n = s_n) \\ &= \delta_{s_n} \otimes P_{n+1} \otimes P_{n+2}(A_0 \times A_1 \times A_2), \end{aligned}$$

wobei die Markov-Eigenschaft in der vierten Zeile und deren Erweiterung (Lemma 1.2) in der fünften Zeile verwendet wurde. Für die letzte Zeile beachte, dass

$$\begin{aligned} \mathbb{P}^{(M_n, M_{n+1}, M_{n+2})|M_n=s} &= \mathbb{P}^{M_n|M_n=s} \otimes \mathbb{P}^{M_{n+1}|M_n=\bullet} \otimes \mathbb{P}^{M_{n+2}|(M_n, M_{n+1})=\bullet} \\ &= \delta_s \otimes P_{n+1} \otimes P_{n+2} \end{aligned}$$

für \mathbb{P}^{M_n} -fast alle $s \in \mathcal{S}$. Ist $(M_n)_{n \geq 0}$ außerdem zeitlich homogen, vereinfacht sich die letzte Zeile von (1.9) weiter zu $\delta_{s_n} \otimes P^2(A_0 \times A_1 \times A_2)$, weil $P_n = P$ für alle $n \geq 1$. \square

Korollar 1.4. Gegeben eine MK $M = (M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$, gilt

$$\mathbb{P}^{(M_k)_{k \geq n} | M_0, \dots, M_n} = \mathbb{P}^{(M_k)_{k \geq n} | M_n} = \delta_{M_n} \otimes \left(\bigotimes_{k=1}^{\infty} P_{n+k} \right) \quad \mathbb{P}\text{-f.s.} \quad (1.10)$$

für alle $k, n \geq 0$. Ist M außerdem zeitlich homogen, entspricht der letzte Ausdruck in (1.10) $\delta_{M_n} \otimes P^\infty$.

Beweis. Wenn überhaupt eine regulär bedingte Verteilung $\mathbb{P}^{(M_k)_{k \geq n} | M_n}$ existiert, so folgt unter Benutzung von (1.7) für jede Zylindermenge $Z = A \times \mathcal{S}^\infty$, $A \in \mathfrak{G}^{m+1}$ und $m \geq 0$

$$\begin{aligned} \mathbb{P}^{(M_k)_{k \geq n} | M_n}(Z) &= \mathbb{P}((M_k)_{k \geq n} \in Z | M_n) \\ &= \mathbb{P}((M_n, \dots, M_{n+m}) \in A | M_n) \\ &= \mathbb{P}((M_n, \dots, M_{n+m}) \in A | M_0, \dots, M_n) \\ &= \mathbb{P}((M_k)_{k \geq n} \in Z | M_0, \dots, M_n) \quad \mathbb{P}\text{-f.s.} \end{aligned}$$

und dann mittels eines Dynkin-System-Arguments allgemein

$$\mathbb{P}^{(M_k)_{k \geq n} | M_n}(B) = \mathbb{P}((M_k)_{k \geq n} \in B | M_0, \dots, M_n) \quad \mathbb{P}\text{-f.s.}$$

für alle $B \in \mathfrak{G}^\infty$, da die Zylindermengen eine Algebra bilden. $\mathbb{P}^{(M_k)_{k \geq n} | M_n}$ bildet somit auch eine regulär bedingte Verteilung von $(M_k)_{k \geq n}$ gegeben M_0, \dots, M_n , was die erste Gleichung in (1.10) beweist.

Dass tatsächlich eine regulär bedingte Verteilung $\mathbb{P}^{(M_k)_{k \geq n} | M_n}$ existiert, erhält man auf ähnliche Weise per Nachweis der zweiten Gleichung in (1.10): Für Z wie oben liefert nämlich wiederum (1.7)

$$\begin{aligned} \mathbb{P}((M_k)_{k \geq n} \in Z | M_n) &= \mathbb{P}((M_n, \dots, M_{n+m}) \in A | M_n) \\ &= \delta_{M_n} \otimes \left(\bigotimes_{k=1}^m P_{n+k} \right) (A) \\ &= \delta_{M_n} \otimes \left(\bigotimes_{k=1}^{\infty} P_{n+k} \right) (Z) \quad \mathbb{P}\text{-f.s.} \end{aligned}$$

was zusammen mit einem weiteren Dynkin-System-Argument

$$\mathbb{P}((M_k)_{k \geq n} \in B | M_n) = \delta_{M_n} \otimes \left(\bigotimes_{k=1}^{\infty} P_{n+k} \right) (B) \quad \mathbb{P}\text{-f.s.}$$

für alle $B \in \mathfrak{G}^\infty$ ergibt. Der stochastische Kern $\delta_{M_n} \otimes \left(\bigotimes_{k=1}^{\infty} P_{n+k} \right)$ ist also eine regulär bedingte Verteilung von $(M_k)_{k \geq n}$ gegeben M_n . \square

Der anschließende Satz gibt eine wichtige äquivalente Charakterisierung der Markov-Eigenschaft.

Satz 1.5. Eine stochastische Folge $(M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{G})$ besitzt genau dann die Markov-Eigenschaft, wenn für jedes $n \geq 0$ (M_0, \dots, M_n) und $(M_k)_{k \geq n}$ bedingt unter M_n stochastisch unabhängig sind, d.h.

$$\mathbb{P}^{((M_0, \dots, M_n), (M_k)_{k \geq n}) | M_n} = \mathbb{P}^{(M_0, \dots, M_n) | M_n} \otimes \mathbb{P}^{(M_k)_{k \geq n} | M_n} \quad \mathbb{P}\text{-f.s.} \quad (1.11)$$

für alle $n \geq 0$.

Weniger formal, aber griffig formuliert besagt dieser Satz:

Eine stochastische Folge besitzt genau dann die Markov-Eigenschaft, wenn zu jedem Zeitpunkt Vergangenheit und Zukunft der Folge bedingt unter der Gegenwart stochastisch unabhängig sind.

Beweis. Zur Abkürzung setzen wir wieder $M_{0:n} = (M_0, \dots, M_n)$ und außerdem $M^{(n)} = (M_k)_{k \geq n} = M_{n:\infty}$.

“ \Rightarrow ” Gilt die Markov-Eigenschaft, so folgt für alle $A \in \mathfrak{S}^{n+1}, B \in \mathfrak{S}$ und $C \in \mathfrak{S}^\infty$

$$\begin{aligned}
 & \int_{\{M_n \in B\}} \mathbb{P}^{(M_{0:n}, M^{(n)})|M_n}(A \times C) d\mathbb{P} \\
 &= \mathbb{P}(M_{0:n} \in A, M_n \in B, M^{(n)} \in C) \\
 &= \int_{\{M_{0:n} \in A, M_n \in B\}} \mathbb{P}^{M^{(n)}|M_{0:n}}(C) d\mathbb{P} \\
 &= \int_{\{M_{0:n} \in A, M_n \in B\}} \mathbb{P}^{M^{(n)}|M_n}(C) d\mathbb{P} \\
 &= \int_{\{M_n \in B\}} \mathbb{E}(\mathbf{1}_{\{M_{0:n} \in A\}} \mathbb{P}^{M^{(n)}|M_n}(C) | M_n) d\mathbb{P} \\
 &= \int_{\{M_n \in B\}} \mathbb{P}(M_{0:n} \in A | M_n) \mathbb{P}^{M^{(n)}|M_n}(C) d\mathbb{P} \\
 &= \int_{\{M_n \in B\}} \mathbb{P}^{M_{0:n}|M_n}(A) \mathbb{P}^{M^{(n)}|M_n}(C) d\mathbb{P}
 \end{aligned}$$

und somit $\mathbb{P}^{(M_{0:n}, M^{(n)})|M_n} = \mathbb{P}^{M_{0:n}|M_n} \otimes \mathbb{P}^{M^{(n)}|M_n}$ \mathbb{P} -f.s., d.h. (1.11).

“ \Leftarrow ” Bei Gültigkeit von (1.11) erhalten wir für alle A, C wie oben

$$\begin{aligned}
 & \int_{\{M_{0:n} \in A\}} \mathbb{P}^{M^{(n)}|M_{0:n}}(C) d\mathbb{P} = \mathbb{P}(M_{0:n} \in A, M^{(n)} \in C) \\
 &= \int_{\Omega} \mathbb{P}^{(M_{0:n}, M^{(n)})|M_n}(A \times C) d\mathbb{P} \\
 &= \int_{\Omega} \mathbb{P}^{M_{0:n}|M_n}(A) \mathbb{P}^{M^{(n)}|M_n}(C) d\mathbb{P} \\
 &= \int_{\Omega} \mathbb{E}(\mathbf{1}_{\{M_{0:n} \in A\}} \mathbb{P}^{M^{(n)}|M_n}(C) | M_n) d\mathbb{P} \\
 &= \int_{\{M_{0:n} \in A\}} \mathbb{P}^{M^{(n)}|M_n}(C) d\mathbb{P}
 \end{aligned}$$

und folglich $\mathbb{P}^{M^{(n)}|M_{0:n}} = \mathbb{P}^{M^{(n)}|M_n}$ \mathbb{P} -f.s. □

Zum Ende dieses Abschnitts wenden wir uns, gegeben eine *zeitlich homogene* MK $(M_n)_{n \geq 0}$ mit n -Schritt Übergangskernen $P^{(n)}$, einer wichtigen Eigenschaft der Familie $(P^{(n)})_{n \geq 0}$ zu: Für beliebige Kerne Q_1, Q_2 von $(\Omega_0, \mathfrak{A}_0)$ nach $(\Omega_1, \mathfrak{A}_1)$ bzw. $(\Omega_1, \mathfrak{A}_1)$ nach $(\Omega_2, \mathfrak{A}_2)$ definieren wir zuvor die *Hintereinanderschaltung von Q_1 und Q_2* durch

$$Q_1 \circ Q_2(\omega_0, A_2) := Q_1 \otimes Q_2(\omega_0, \Omega_1 \times A_2) = \int_{\Omega_1} Q_2(\omega_1, A_2) Q_1(\omega_0, d\omega_1).$$

$Q_1 \circ Q_2$ bildet folglich einen Kern von $(\Omega_0, \mathfrak{A}_0)$ nach $(\Omega_2, \mathfrak{A}_2)$. Ist Q_1 unabhängig von ω_0 , also ein Maß auf $(\Omega_1, \mathfrak{A}_1)$, so hängt auch $Q_1 \circ Q_2$ nicht von ω_0 ab und bildet deshalb ein Maß auf $(\Omega_2, \mathfrak{A}_2)$. Für $n \geq 3$ Kerne Q_j von $(\Omega_{j-1}, \mathfrak{A}_{j-1})$ nach $(\Omega_j, \mathfrak{A}_j)$, $1 \leq j \leq n$, ergibt sich induktiv

$$\begin{aligned} & Q_1 \circ \dots \circ Q_n(\omega_0, A_n) \\ &= Q_1 \otimes \dots \otimes Q_n(\omega_0, \Omega_1 \times \dots \times \Omega_{n-1} \times A_n) \\ &= \int_{\Omega_1} \dots \int_{\Omega_{n-1}} Q_n(\omega_{n-1}, A_n) Q_{n-1}(\omega_{n-2}, d\omega_{n-1}) \dots Q_1(\omega_0, d\omega_1). \end{aligned} \quad (1.12)$$

Die Verknüpfung ist außerdem assoziativ, d.h.

$$Q_1 \circ \dots \circ Q_n = Q_1 \circ (Q_2 \circ \dots \circ Q_n) = (Q_1 \circ \dots \circ Q_{n-1}) \circ Q_n.$$

Statt $Q_1 \circ Q_2$ schreiben wir im folgenden kurz $Q_1 Q_2$.

Gegeben eine MK $(M_n)_{n \geq 0}$ mit Anfangsverteilung $P_0 = \mathbb{P}^{M_0}$ und bedingten Verteilungen $P_n(s, \cdot) = \mathbb{P}^{M_n | M_{n-1}=s}$, $n \geq 1$, sieht man nun mittels (1.12) sofort ein, dass

$$P_0 \dots P_n = \mathbb{P}^{M_n} \quad \text{und} \quad P_{n+1} \dots P_{n+k} = \mathbb{P}^{M_{n+k} | M_n = \bullet} \quad \mathbb{P}^{M_n}\text{-f.s.}$$

für alle $k, n \geq 0$. Im zeitlich homogenen Fall führt dies schließlich zu

$$P^{(k)} = \mathbb{P}^{M_{n+k} | M_n = \bullet} = P^{\circ k} := \underbrace{P \dots P}_{k\text{-mal}} \quad \mathbb{P}^{M_n}\text{-f.s.} \quad (1.13)$$

für alle $k, n \geq 0$, wobei $\mathbb{P}^{\circ 0}(x, \cdot) := \delta_x(\cdot)$. $P^{\circ k}$ definiert somit für jedes k eine Version des k -Schritt-Übergangskerns, und es folgt in diesem Fall sofort:

Satz 1.6. Für eine zeitlich homogene MK $(M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$ bilden die n -Schritt-Übergangskerne $P^{(n)} := P^{\circ n}$, $n \geq 0$, eine Halbgruppe bezüglich der Hintereinanderschaltung, d.h.

$$P^{(m+n)} = P^{(m)} P^{(n)}.$$

für alle $m, n \geq 0$. Dies bedeutet, dass

$$\mathbb{P}^{M_{k+m+n}|M_k=s}(A) = \int_{\mathcal{S}} \mathbb{P}^{M_{k+m+n}|M_{k+m}=t}(A) \mathbb{P}^{M_{k+m}|M_k=s}(dt)$$

für alle $k, m, n \geq 0$, $A \in \mathfrak{S}$ und \mathbb{P}^{M_k} -fast alle $s \in \mathcal{S}$.

Beweis. Unter Hinweis auf (1.13) und die Assoziativität der Hintereinanderschaltung ist nichts mehr zu zeigen. \square

Im diskreten Fall (\mathcal{S} abzählbar) ist $P^{(k)}$ genau wie $P = P^{(1)}$ durch die einfachen k -Schritt-Übergangswahrscheinlichkeiten

$$p_{ij}^{(k)} := P^{(k)}(i, \{j\})$$

determiniert, die in der k -Schritt-Übergangsmatrix $P^{(k)} := (p_{ij}^{(k)})_{i,j \in \mathcal{S}}$ zusammengefasst werden. Setzen wir wie im obigen Satz $P^{(k)} = P^{ok}$, so ergibt sich unter Hinweis auf (1.12) bei gleichzeitiger Ausnutzung der Abzählbarkeit von \mathcal{S}

$$p_{ij}^{(k)} = \sum_{(i_1, \dots, i_{k-1}) \in \mathcal{S}^{k-1}} p_{i i_1} p_{i_1 i_2} \cdots p_{i_{k-2} i_{k-1}} p_{i_{k-1} j},$$

was nichts anderes bedeutet als

$$P^{(k)} = P^k = \underbrace{P \cdots P}_{k\text{-mal}}$$

für alle $k \geq 0$, wobei $P^{(0)} = \mathbf{I} := (\delta_{ij})_{i,j \in \mathcal{S}}$ die Einheitsmatrix bezeichnet. Aus der Hintereinanderschaltung wird im diskreten Fall also eine gewöhnliche Matrizenmultiplikation. Die entsprechende, nunmehr triviale Halbgruppeneigenschaft notieren wir abschließend in

Korollar 1.7. [Kolmogorov-Chapman-Gleichungen] Gegeben eine DMK mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$ und Übergangsmatrizen $P^{(k)} = (p_{ij}^{(k)})_{i,j \in \mathcal{S}} = P^k$, $k \geq 0$, gilt

$$P^{(m+n)} = P^{(m)} P^{(n)}, \quad \text{d.h.} \quad p_{ij}^{(m+n)} = \sum_{k \in \mathcal{S}} p_{ik}^{(m)} p_{kj}^{(n)}$$

für alle $m, n \geq 0$ und $i, j \in \mathcal{S}$.

1.2 Das Standardmodell

Wir haben im vorherigen Abschnitt bereits festgestellt, dass die Verteilung einer zeitlich homogenen MK $M = (M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$ durch ihre Anfangsverteilung $\lambda = \mathbb{P}^{M_0}$ und ihren Übergangskern (Übergangsmatrix, falls \mathcal{S} abzählbar) P vollständig determiniert ist. Eine Analyse dieser Verteilung ist somit gleichbedeutend mit einer Analyse des Paares (λ, P) und hängt im Ergebnis von der expliziten Definition des Prozesses M gar nicht ab. Wir dürfen uns deshalb ohne weiteres ein geeignetes Modell auf der Basis von (λ, P) wählen, in dem es sich besonders angenehm rechnen lässt. Dabei erweist es sich als sinnvoll, wie wir bald einsehen werden, beliebige Anfangsverteilungen λ zu berücksichtigen und nur den Übergangskern P als festen Parameter zugrunde zu legen.

Wir beschreiben als nächstes ein Standardmodell, was diesen Anforderungen Rechnung trägt: Gegeben einen Übergangskern P auf $(\mathcal{S}, \mathfrak{S})$, seien $(\Omega, \mathfrak{A}) = (\mathcal{S}^\infty, \mathfrak{S}^\infty)$ und $M = (M_n)_{n \geq 0}$ die Identität auf Ω , genannt *kanonischer Prozess* oder auch *Koordinatenprozess* ($\mathfrak{U}^{\mathfrak{S}}$ [3, Bemerkung 54.2(c)]). $M_n(\omega)$ bezeichnet demnach die Projektion von $\omega \in \Omega$ auf die n -te Komponente. Für *jede* Verteilung λ auf $(\mathcal{S}, \mathfrak{S})$ definieren wir nun weiter auf (Ω, \mathfrak{A}) das nach dem Satz von Ionescu Tulcea eindeutig bestimmte W-Maß

$$\mathbb{P}_\lambda = \lambda \otimes P^\infty$$

mit der Eigenschaft ($\mathfrak{U}^{\mathfrak{S}}$ (54.1) in [3])

$$\begin{aligned} \mathbb{P}_\lambda \left(\prod_{k=0}^n A_k \times \mathcal{S}^\infty \right) &= \mathbb{P}_\lambda (M_0 \in A_0, \dots, M_n \in A_n) \\ &= \int_{A_0} \int_{A_1} \dots \int_{A_n} P(s_{n-1}, ds_n) \dots P(s_0, ds_1) \lambda(ds_0) \end{aligned} \quad (1.14)$$

für alle $A_0, \dots, A_n \in \mathfrak{S}$ und $n \geq 0$, was sich im diskreten Fall zu

$$\mathbb{P}_\lambda \left(\prod_{k=0}^n A_k \times \mathcal{S}^\infty \right) = \sum_{i_0 \in A_0, \dots, i_n \in A_n} \lambda_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n}$$

vereinfacht mit $p_{ij} = P(i, \{j\})$ und $\lambda = (\lambda_i)_{i \in \mathcal{S}}$, wobei $\lambda_i := \lambda(\{i\})$. Insbesondere gilt dann

$$\mathbb{P}_\lambda (M_0 = i_0, M_1 = i_1, \dots, M_n = i_n) = \lambda_{i_0} \prod_{k=1}^n p_{i_{k-1} i_k} \quad (1.15)$$

für alle $(i_0, \dots, i_n) \in \mathcal{S}^{n+1}$ und $n \geq 0$. $(M_n)_{n \geq 0}$ bildet also unter \mathbb{P}_λ eine MK mit Anfangsverteilung λ und Übergangskern P bzw. Übergangsmatrix $P = (p_{ij})_{i, j \in \mathcal{S}}$. Ferner gilt offenbar $\mathbb{P}_\lambda^{M_n} = \lambda P^{on}$ für alle $n \geq 0$.

Definition 1.8. Gegeben einen Übergangskern P auf $(\mathcal{S}, \mathfrak{S})$, nennen wir das zuvor spezifizierte Modell

$$(\mathcal{S}^\infty, \mathfrak{S}^\infty, (M_n)_{n \geq 0}, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})}),$$

$\mathcal{P}(\mathcal{S})$ die Menge der Verteilungen auf $(\mathcal{S}, \mathfrak{S})$, das zu P gehörende kanonische Modell. Ferner heißt jedes Modell

$$(\Omega, \mathfrak{A}, (M_n)_{n \geq 0}, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})}),$$

so dass $M_n : (\Omega, \mathfrak{A}) \rightarrow (\mathcal{S}, \mathfrak{S})$, $n \geq 0$, unter \mathbb{P}_λ eine MK mit Startverteilung λ und Übergangskern P definiert, ein *Standardmodell* zu P .

In einem Standardmodell haben wir es somit nur mit *einem* Prozess zu tun, wobei sich verschiedene Anfangsverteilungen durch Zugrundelegung verschiedener \mathbb{P}_λ ergeben. Dies erweist sich bei den nachfolgenden Untersuchungen als wesentlich zweckmäßiger als bei jedem Wechsel der Anfangsverteilung immer auch den Prozess wechseln zu müssen.

Startet $(M_n)_{n \geq 0}$ in einem Punkt $s \in \mathcal{S}$, gilt also $\lambda = \delta_s$, so schreiben wir auch \mathbb{P}_s für \mathbb{P}_{δ_s} . Offensichtlich gilt dann für beliebiges $\lambda \in \mathcal{P}(\mathcal{S})$

$$\mathbb{P}_\lambda(\cdot) = \int_{\mathcal{S}} \mathbb{P}_s(\cdot) \lambda(ds) \quad \text{bzw.} \quad \mathbb{P}_\lambda(\cdot) = \sum_{s \in \mathcal{S}} \lambda_s \mathbb{P}_s(\cdot), \quad (1.16)$$

im diskreten Fall. Jedes \mathbb{P}_λ ergibt sich somit als im Allgemeinen unendliche konvexe Kombination der \mathbb{P}_s , $s \in \mathcal{S}$. (1.16) bleibt auch für σ -endliche λ sinnvoll. Wir erhalten dann ein σ -endliches Maß \mathbb{P}_λ auf (Ω, \mathfrak{A}) , das weiter durch (1.14) und (1.15) charakterisiert ist. Diese Erweiterung benötigen wir später bei der Betrachtung sogenannter stationärer Maße, die im Allgemeinen lediglich σ -endlich sind (Abschnitt 5).

Markov-Eigenschaft und zeitliche Homogenität lassen sich in einem Standardmodell wie folgt formulieren:

Satz 1.9. Gegeben einen Übergangskern P auf $(\mathcal{S}, \mathfrak{S})$ samt eines zugehörigen Standardmodells $(\Omega, \mathfrak{A}, M, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})})$ mit $M = (M_n)_{n \geq 0}$, gilt für alle $\lambda \in \mathcal{P}(\mathcal{S})$ und $n \in \mathbb{N}_0$

$$\mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_0, \dots, M_n} = \mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_n} = P_{M_n}^M \quad \mathbb{P}_\lambda\text{-f.s.}, \quad (1.17)$$

oder expliziter

$$\mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_0, \dots, M_n}(\omega, A) = \mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_n}(\omega, A) = P_{M_n(\omega)}^M(A) \quad (1.31')$$

für \mathbb{P}_λ -fast alle $\omega \in \Omega$ und $A \in \mathfrak{G}^\infty$.

Beweis. Die Aussage ergibt sich sofort, weil unter Hinweis auf Korollar 1.4

$$\mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_0, \dots, M_n} = \mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_n} = \delta_{M_n} \otimes P^\infty = P_{M_n}^M \quad \mathbb{P}_\lambda\text{-f.s.}$$

gilt. □

Da die bedingte Verteilung von $(M_k)_{k \geq n}$ gegeben M_n unter \mathbb{P}_λ gar nicht von der Anfangsverteilung λ abhängt, schreiben wir im folgenden einfach $\mathbb{P}^{(M_k)_{k \geq n} | M_n}$ statt $\mathbb{P}_\lambda^{(M_k)_{k \geq n} | M_n}$. Entsprechend bedeute “ \mathbb{P} -f.s.” in einem Standardmodell, dass die betreffende Aussage \mathbb{P}_λ -f.s. für alle $\lambda \in \mathcal{P}(\mathcal{S})$ Gültigkeit hat. Abschließend sei noch notiert, dass \mathbb{E}_s und \mathbb{E}_λ die Erwartungswertoperatoren unter \mathbb{P}_s bzw. \mathbb{P}_λ bezeichnen. Es gilt dann

$$\mathbb{E}(f(M_n, M_{n+1}, \dots) | M_n = s) = \mathbb{E}_s f(M_0, M_1, \dots)$$

für \mathbb{P}^{M_n} -fast alle $s \in \mathcal{S}$ und jede \mathbb{P}^{M_n} -quasi-integrierbare Funktion $f: \mathcal{S}^\infty \rightarrow \overline{\mathbb{R}}$.

1.3 Filtrationen und Stopzeiten

Auf dem Weg zu einer wichtigen Verschärfung der Markov-Eigenschaft im nächsten Abschnitt bedarf es zunächst der kurzen Einführung der Begriffe “Filtration” und “Stopzeit”, denen in der Theorie stochastischer Prozesse auch allgemein große Bedeutung zukommt. Da hierfür der zuvor gesteckte, sehr spezielle Rahmen bedeutungslos ist, begeben wir uns für einen Moment in die generische Situation eines gegebenen stochastischen Prozesses in diskreter Zeit.

1.3.1 Filtrationen

Sei (Ω, \mathfrak{A}) ein beliebiger messbarer Raum. Obgleich wir dabei im Grunde einen W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$ im Auge haben, spielt das W-Maß \mathbb{P} zunächst keine Rolle. Dennoch werden wir uns aus Interpretationsgründen die Freiheit nehmen, messbare Mengen auch Ereignisse zu nennen.

Definition 1.10. Eine aufsteigende Folge $(\mathcal{F}_n)_{n \geq 0}$ von Unter- σ -Algebren von \mathfrak{A} heißt *Filtration des Raums* (Ω, \mathfrak{A}) .

Stellen wir uns n als Zeitparameter vor, so können wir \mathcal{F}_n als das *System der bis zum Zeitpunkt n beobachtbaren Ereignisse* interpretieren. Mit anderen Worten, \mathcal{F}_n bildet die Gesamtheit aller Ereignisse, von denen ein Beobachter zum Zeitpunkt n entscheiden kann, ob sie eingetreten sind oder nicht. Man nennt \mathcal{F}_n deshalb manchmal etwas vager auch *die zum Zeitpunkt n für den Beobachter verfügbare Information*. Die σ -Algebra

$$\mathcal{F}_\infty := \sigma \left(\bigcup_{n=0}^{\infty} \mathcal{F}_n \right)$$

beinhaltet offenkundig alle jemals vom betreffenden Beobachter entscheidbaren Ereignisse, seine asymptotische Gesamtinformation also.

Verschiedene Beobachter können natürlich verschiedene Informationen erhalten. Ihnen sind dann verschiedene Filtrationen $(\mathcal{F}_n)_{n \geq 0}$ und $(\mathcal{G}_n)_{n \geq 0}$ zugeordnet. Wir schreiben $(\mathcal{G}_n)_{n \geq 0} \subset (\mathcal{F}_n)_{n \geq 0}$, falls $\mathcal{G}_n \subset \mathcal{F}_n$ für alle $n \geq 0$. In diesem Fall hat also der “ \mathcal{G} -Beobachter” zu jedem Zeitpunkt n höchstens genausoviel Information wie der “ \mathcal{F} -Beobachter”.

Betrachten wir als nächstes eine Folge $(X_n)_{n \geq 0}$ messbarer Abbildungen auf (Ω, \mathfrak{A}) , die also bei zusätzlich gegebenem W-Maß \mathbb{P} einen stochastischen Prozess in diskreter Zeit bildet.

Definition 1.11. Eine Folge $(X_n)_{n \geq 0}$ messbarer Abbildungen auf (Ω, \mathfrak{A}) heißt *adaptiert bzgl. der Filtration $(\mathcal{F}_n)_{n \geq 0}$* oder einfach *$(\mathcal{F}_n)_{n \geq 0}$ -adaptiert*, wenn X_n \mathcal{F}_n -messbar ist für jedes $n \geq 0$.

Offensichtlich ist $(X_n)_{n \geq 0}$ genau dann adaptiert bezüglich $(\mathcal{F}_n)_{n \geq 0}$, wenn

$$\mathcal{G}_n := \sigma(X_0, \dots, X_n) \subset \mathcal{F}_n$$

für alle $n \geq 0$ gilt. $(\mathcal{G}_n)_{n \geq 0}$ bildet offenkundig selbst eine Filtration, und zwar gerade die kleinste, bezüglich der $(X_n)_{n \geq 0}$ adaptiert ist. Sie heißt *kanonische Filtration von $(X_n)_{n \geq 0}$* . Erwähnen wollen wir noch, dass eine bezüglich $(\mathcal{F}_n)_{n \geq 0}$ adaptierte Folge $(X_n)_{n \geq 0}$ als Vektor \mathcal{F}_∞ -messbar ist, was im Falle reellwertiger oder numerischer X_n insbesondere die \mathcal{F}_∞ -Messbarkeit der Abbildungen

$$\inf_{n \geq 0} X_n, \quad \sup_{n \geq 0} X_n, \quad \liminf_{n \rightarrow \infty} X_n \quad \text{und} \quad \limsup_{n \rightarrow \infty} X_n$$

impliziert.

1.3.2 Stopzeiten

Bei der Untersuchung stochastischer Prozesse $(X_n)_{n \geq 0}$ spielen häufig Zufallszeiten der Form

$$\tau = \inf\{n \geq 0 : (X_0, \dots, X_n) \in A_n\} \quad (1.18)$$

für geeignete messbare Mengen A_n eine wichtige Rolle. Wir setzen dabei immer $\tau = \infty$, falls das Infimum über die leere Menge gebildet wird. Stellen wir uns vor, τ bezeichnet den Zeitpunkt, zu dem ein Beobachter aufhört, den Prozess $(X_n)_{n \geq 0}$ zu verfolgen. Das typische an τ ist, dass es nicht auf Information über die Folge zurückgreift, die erst in der Zukunft verfügbar würde. Mit anderen Worten, das Ereignis, zum Zeitpunkt n zu stoppen, hängt nur von den Werten X_0, \dots, X_n ab für jedes $n \geq 0$. Man sagt auch, τ ist *nicht antizipierend*. Ein Beispiel einer antizipierenden Zufallszeit im Fall reellwertiger X_n bildet etwa

$$v = \sup\{n \geq 0 : X_n \leq 0\} \quad [\sup \emptyset := 0],$$

Unter v bedarf es nämlich zur Entscheidung darüber, zum Zeitpunkt n zu stoppen, der vollständigen Realisierung von $(X_n)_{n \geq 0}$.

Wie zuvor bemerkt, lässt sich verfügbare Information zu sukzessiven Zeitpunkten formal mittels Filtrationen beschreiben, was zu folgender allgemeinen Definition nicht antizipierender Zufallszeiten führt:

Definition 1.12. Sei $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration des messbaren Raums (Ω, \mathfrak{A}) . Dann heißt eine messbare Abbildung $\tau : \Omega \rightarrow \mathbb{N}_0 \cup \{\infty\}$ *Stopzeit bezüglich $(\mathcal{F}_n)_{n \geq 0}$* oder auch *$(\mathcal{F}_n)_{n \geq 0}$ -Zeit*, wenn

$$\{\tau = n\} \in \mathcal{F}_n \quad (1.19)$$

für alle $n \in \mathbb{N}_0$ gilt. Im Fall $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ für eine Folge $(X_n)_{n \geq 0}$ messbarer Abbildungen nennt man τ auch *Stopzeit bezüglich $(X_n)_{n \geq 0}$* . Die σ -Algebra

$$\mathcal{F}_\tau := \left\{ A \in \mathfrak{A} : A \cap \{\tau = n\} \in \mathcal{F}_n \text{ für alle } n \in \overline{\mathbb{N}}_0 \right\} \quad (1.20)$$

bezeichnet man als *σ -Algebra der τ -Vergangenheit* (gegeben $(\mathcal{F}_n)_{n \geq 0}$).

Anmerkung 1.13. Jede konstante Abbildung $\tau \equiv n$, $n \in \overline{\mathbb{N}}_0$, ist selbstverständlich Stopzeit bezüglich jeder Filtration $(\mathcal{F}_n)_{n \geq 0}$ des zugrundeliegenden messbaren Raumes, und es gilt dann $\mathcal{F}_\tau = \mathcal{F}_n$.

Anmerkung 1.14. Bedingung (1.19) gilt auch für $n = \infty$, denn

$$\{\tau = \infty\} = \left(\sum_{n \in \mathbb{N}_0} \underbrace{\{\tau = n\}}_{\in \mathcal{F}_n \subset \mathcal{F}_\infty} \right)^c \in \mathcal{F}_\infty. \quad (1.21)$$

Anmerkung 1.15. Äquivalent zur Bedingung (1.19) ist offensichtlich sowohl

$$\{\tau \leq n\} \in \mathcal{F}_n \quad (1.22)$$

für alle $n \in \mathbb{N}_0$ als auch (Komplementbildung)

$$\{\tau > n\} \in \mathcal{F}_n$$

für alle $n \in \mathbb{N}_0$.

Anmerkung 1.16. Jede Stopzeit τ bezüglich einer Folge $(X_n)_{n \geq 0}$ hat die Form (1.18). Dazu beachte man, dass $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ gerade aus den Urbildern messbarer Mengen unter (X_0, \dots, X_n) besteht. Für jedes $n \geq 0$ impliziert demnach $\{\tau \leq n\} \in \mathcal{F}_n$ die Existenz einer messbaren Menge A_n , so dass $\{\tau \leq n\} = \{(X_0, \dots, X_n) \in A_n\}$, was offenbar (1.18) für diese A_n liefert.

Anmerkung 1.17. Dass \mathcal{F}_τ tatsächlich eine σ -Algebra bildet, wie in der obigen Definition einfach konstatiert wird, und dass sich dieselbe σ -Algebra ergibt, wenn man dort die Mengen $\{\tau = n\}$, $n \in \mathbb{N}_0$, durch $\{\tau \leq n\}$ ersetzt, kann der Leser mühelos selbst nachweisen.

Die grundlegenden Fakten über Stopzeiten und die zugehörigen σ -Algebren fassen wir in folgendem Satz zusammen.

Satz 1.18. *Gegeben eine Filtration $(\mathcal{F}_n)_{n \geq 0}$ des messbaren Raums (Ω, \mathfrak{A}) , messbare Abbildungen X, X_0, X_1, \dots auf diesem sowie Stopzeiten $\sigma, \tau, \tau_1, \tau_2, \dots$ bezüglich $(\mathcal{F}_n)_{n \geq 0}$, gelten folgende Aussagen:*

- (a) $\sigma \wedge \tau, \sigma \vee \tau, \sigma + \tau$ sind Stopzeiten bezüglich $(\mathcal{F}_n)_{n \geq 0}$.
- (b) $\inf_{n \geq 1} \tau_n, \sup_{n \geq 1} \tau_n, \liminf_{n \rightarrow \infty} \tau_n, \limsup_{n \rightarrow \infty} \tau_n$ und $\lim_{n \rightarrow \infty} \tau_n$ (falls existent) sind Stopzeiten bezüglich $(\mathcal{F}_n)_{n \geq 0}$.
- (c) $\{\sigma = \tau\}, \{\sigma \leq \tau\} \in \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.
- (d) $\sigma \leq \tau$ impliziert $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$. Insbesondere folgt $\mathcal{F}_\sigma \subset \mathcal{F}_{\sigma+1} \subset \dots \subset \mathcal{F}_\infty$, und $\tau - \sigma$ bildet eine $(\mathcal{F}_{\sigma+n})_{n \geq 0}$ -Zeit, sofern $\sigma < \infty$.
- (e) $\mathcal{F}_{\sigma \wedge \tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$ und $\mathcal{F}_{\sigma \vee \tau} = \sigma(\mathcal{F}_\sigma \cup \mathcal{F}_\tau)$.
- (f) Aus $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ und $\tau < \infty$ folgt $\mathcal{F}_\tau = \sigma(\tau, X_0, \dots, X_\tau)$.
- (g) Aus $(\mathcal{F}_n)_{n \geq 0} \subset (\mathcal{G}_n)_{n \geq 0}$ folgt $\mathcal{F}_\tau \subset \mathcal{G}_\tau$.
- (h) Eine Zufallsgröße X ist genau dann \mathcal{F}_τ -messbar, wenn $X \mathbf{1}_{\{\tau=n\}}$ \mathcal{F}_n -messbar ist für alle $n \in \overline{\mathbb{N}}_0$.

Beweis. (a) Dass $\sigma \wedge \tau, \sigma \vee \tau$ und $\sigma + \tau$ wieder Stopzeiten bilden, folgt unter Hinweis auf (1.20) und (1.22), denn

$$\begin{aligned} \{\sigma \wedge \tau > n\} &= \{\sigma > n\} \cap \{\tau > n\} \in \mathcal{F}_n, \\ \{\sigma \vee \tau \leq n\} &= \{\sigma \leq n\} \cap \{\tau \leq n\} \in \mathcal{F}_n, \\ \text{und } \{\sigma + \tau = n\} &= \sum_{k=0}^n \{\sigma = k\} \cap \{\tau = n - k\} \in \mathcal{F}_n \end{aligned}$$

für jedes $n \geq 0$.

(b) Hier betrachten wir nur $\liminf_{n \rightarrow \infty} \tau_n$ und notieren, dass

$$\begin{aligned} \left\{ \liminf_{n \rightarrow \infty} \tau_n > m \right\} &= \{ \omega : \exists k = k(\omega) : \forall n \geq k : \tau_n(\omega) > m \} \\ &= \bigcup_{k \geq 0} \bigcap_{n \geq k} \{ \tau_n > m \} \in \mathcal{F}_m \end{aligned}$$

für alle $m \in \mathbb{N}_0$.

(c) überlassen wir dem Leser.

(d) $\sigma \leq \tau$ und $A \in \mathcal{F}_\sigma$ implizieren

$$A \cap \{ \tau = n \} = \sum_{0 \leq k \leq n} \underbrace{(A \cap \{ \sigma = k \})}_{\in \mathcal{F}_k \subset \mathcal{F}_n} \cap \{ \tau = n \} \in \mathcal{F}_n$$

für alle $n \in \bar{\mathbb{N}}_0$ und somit auch $A \in \mathcal{F}_\tau$. Falls $\sigma < \infty$, so gilt für jedes $n \in \mathbb{N}_0$, dass

$$\{ \tau - \sigma = n \} \cap \{ \sigma + n = k \} = \{ \sigma = k - n \} \cap \{ \tau = n \} \in \mathcal{F}_k$$

für alle $k \geq n$ und folglich $\{ \tau - \sigma = n \} \in \mathcal{F}_{\sigma+n}$. Also ist $\tau - \sigma$ wie behauptet eine Stopzeit bezüglich $(\mathcal{F}_{\sigma+n})_{n \geq 0}$.

(e) Hier notieren wir lediglich als Hinweis für die zweite Aussage, dass sich jedes $A \in \mathcal{F}_{\sigma \vee \tau}$ in die Mengen $A \cap \{ \sigma \leq \tau \} \in \mathcal{F}_\tau$ und $A \cap \{ \sigma > \tau \} \in \mathcal{F}_\sigma$ zerlegen lässt. Den vollständigen Beweis empfehlen wir dem Leser als Übung.

(f) Falls $X_n : (\Omega, \mathfrak{A}) \rightarrow (\Omega_n, \mathfrak{A}_n)$ und $\tau < \infty$, folgt

$$(\tau, X_0, \dots, X_\tau) : (\Omega, \mathfrak{A}) \rightarrow (\Omega', \mathfrak{A}')$$

mit

$$\begin{aligned} \Omega' &:= \sum_{n \geq 0} \{n\} \times \Omega_0 \times \dots \times \Omega_n, & \mathfrak{A}' &:= \sigma(\mathfrak{E}'), \\ \mathfrak{E}' &:= \left\{ \{n\} \times A_n : n \in \mathbb{N}_0, A_n \in \mathfrak{A}_0 \otimes \dots \otimes \mathfrak{A}_n \right\}. \end{aligned} \tag{1.23}$$

Also ist $\sigma(\tau, X_0, \dots, X_\tau) = (\tau, X_0, \dots, X_\tau)^{-1}(\mathfrak{A}') = \sigma((\tau, X_0, \dots, X_\tau)^{-1}(\mathfrak{E}'))$ unter Hinweis auf Lemma 6.1 in [3]. Wie man sofort sieht, gilt $(\tau, X_0, \dots, X_\tau)^{-1}(\mathfrak{E}') \subset \mathcal{F}_\tau$ und damit $\sigma(\tau, X_0, \dots, X_\tau) \subset \mathcal{F}_\tau$. Umgekehrt impliziert $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ für jedes $A \in \mathcal{F}_\tau$ die Existenz eines $A_n \in \mathfrak{A}_0 \otimes \dots \otimes \mathfrak{A}_n$, so dass

$$A \cap \{ \tau = n \} = \{ (X_0, \dots, X_n) \in A_n \} = \{ (\tau, X_0, \dots, X_\tau) \in \{n\} \times A_n \}.$$

Summation über alle $n \geq 0$ liefert dann

$$A = \sum_{n \geq 0} A \cap \{\tau = n\} = \left\{ (\tau, X_0, \dots, X_\tau) \in \sum_{n \geq 0} \{n\} \times A_n \right\} \in \sigma(\tau, X_0, \dots, X_\tau),$$

d.h. die umgekehrte Inklusion $\mathcal{F}_\tau \subset \sigma(\tau, X_0, \dots, X_\tau)$.

(g) kann der Leser wiederum leicht selbst nachweisen.

(h) “ \Rightarrow ” Nach Definition von \mathcal{F}_τ ist $\mathbf{1}_A$ genau dann messbar bezüglich dieser σ -Algebra, wenn $\mathbf{1}_{A \cap \{\tau = n\}} = \mathbf{1}_A \mathbf{1}_{\{\tau = n\}}$ \mathcal{F}_n -messbar ist für alle $n \in \bar{\mathbb{N}}_0$. Damit erhält man sofort, dass jede \mathcal{F}_τ -messbare Elementarfunktion X die Behauptung erfüllt, was schließlich mittels eines Funktions-Erweiterungsarguments auf alle Zufallsgrößen ausgedehnt werden kann.

“ \Leftarrow ” Ist $X \mathbf{1}_{\{\tau = n\}}$ \mathcal{F}_n -messbar für alle $n \in \bar{\mathbb{N}}_0$, so folgt für alle $x \in \bar{\mathbb{R}}$ und $n \in \mathbb{N}_0$

$$\{X \leq x\} \cap \{\tau = n\} = \{X \mathbf{1}_{\{\tau = n\}} \leq x\} \cap \{\tau = n\} \in \mathcal{F}_n,$$

also $\{X \leq x\} \in \mathcal{F}_\tau$ für alle $x \in \bar{\mathbb{R}}$, was die \mathcal{F}_τ -Messbarkeit von X impliziert. \square

Wenn $\{\tau = \infty\} \neq \emptyset$, so sind $(\tau, X_0, \dots, X_\tau)$ und die *Post- τ -Folge* $X^{(\tau)} := (X_{\tau+n})_{n \geq 0}$ nur auf der Spur $(\Omega \cap \{\tau < \infty\}, \mathfrak{A} \cap \{\tau < \infty\})$ wohldefiniert. Um mit diesem Umstand formal sauber umzugehen, treffen wir folgende Definition der Vektoren auf der Menge $\{\tau = \infty\}$:

$$(\tau, X_0, \dots, X_\tau) := (\infty, (X_n)_{n \geq 0}) \quad \text{und} \quad X^{(\tau)} = X^{(\infty)} := (\Delta, \Delta, \dots)$$

für ein nicht weiter spezifiziertes Element Δ . Gegeben $X_n : (\Omega, \mathfrak{A}) \rightarrow (\Omega_n, \mathfrak{A}_n)$ für $n \geq 0$, folgt dann $X^{(\tau)} : (\Omega, \mathfrak{A}) \rightarrow (\Omega'', \mathfrak{A}'')$ mit (vgl. (1.23))

$$\begin{aligned} \Omega'' &:= \left(\sum_{n \geq 0} \{n\} \times \prod_{k \geq n} \Omega_k \right) \cup \{(\infty, \Delta, \Delta, \dots)\}, \quad \mathfrak{A}'' := \sigma(\mathfrak{E}''), \\ \mathfrak{E}'' &:= \left\{ \{n\} \times A_n : n \in \mathbb{N}_0, A_n \in \bigotimes_{k \geq n} \mathfrak{A}_k \right\} \cup \{(\infty, \Delta, \Delta, \dots)\}. \end{aligned}$$

Als triviale Konsequenz der Teile (e) und (h) des vorherigen Satzes notieren wir ohne Beweis:

Korollar 1.19. *In der Situation von Satz 1.18 sei ferner angenommen, dass $(X_n)_{n \geq 0}$ adaptiert ist bezüglich $(\mathcal{F}_n)_{n \geq 0}$. Dann gilt:*

- (a) τ und $(\tau, X_0, \dots, X_\tau)$ sind \mathcal{F}_τ -messbar.
- (b) $X^{(\tau)}$ ist adaptiert bezüglich der Filtration $(\mathcal{F}_{\tau+n})_{n \geq 0}$.

Dass bedingte Erwartungswerte und Verteilungen bezüglich \mathcal{F}_τ sich letztendlich wieder aus solchen bezüglich \mathcal{F}_n für $n \in \bar{\mathbb{N}}_0$, also zu festen Zeitpunkten, berechnen lassen, zeigt der nächste Satz.

Satz 1.20. Sei $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration des W -Raums $(\Omega, \mathfrak{A}, \mathbb{P})$, X eine Zufallsvariable auf diesem mit Werten in (Ω', \mathfrak{A}') sowie τ eine Stopzeit bezüglich $(\mathcal{F}_n)_{n \geq 0}$.

- (a) Existieren regulär bedingte Verteilungen $\mathbb{P}^{X|\mathcal{F}_n}$ für alle $n \in \bar{\mathbb{N}}_0$, so existiert auch $\mathbb{P}^{X|\mathcal{F}_\tau}$, und zwar gilt

$$\mathbb{P}^{X|\mathcal{F}_\tau} = \sum_{n \in \bar{\mathbb{N}}_0} \mathbf{1}_{\{\tau=n\}} \mathbb{P}^{X|\mathcal{F}_n} \quad \mathbb{P}\text{-f.s.}$$

- (b) Ist X eine quasi-integrierbare Zufallsgröße, folgt

$$\mathbb{E}(X|\mathcal{F}_\tau) = \sum_{n \in \bar{\mathbb{N}}_0} \mathbf{1}_{\{\tau=n\}} \mathbb{E}(X|\mathcal{F}_n) \quad \mathbb{P}\text{-f.s.}$$

Beweis. (a) Gemäß Satz 1.18(h) ist $Q(\cdot, A') := \sum_{n \in \mathbb{N}_0 \cup \{\infty\}} \mathbf{1}_{\{\tau=n\}} \mathbb{P}^{X|\mathcal{F}_n}(\cdot, A')$ für jedes $A' \in \mathfrak{A}'$ \mathcal{F}_τ -messbar. Da außerdem

$$\begin{aligned} \int_B Q(\omega, A') \mathbb{P}(d\omega) &= \sum_{n \in \bar{\mathbb{N}}_0} \int_{B \cap \{\tau=n\}} \mathbb{P}^{X|\mathcal{F}_n}(\omega, A') \mathbb{P}(d\omega) \\ &= \sum_{n \in \bar{\mathbb{N}}_0} \int_{B \cap \{\tau=n\}} \mathbf{1}_{A'}(X(\omega)) \mathbb{P}(d\omega) \\ &= \mathbb{P}(B \cap \{X \in A'\}) \end{aligned}$$

für alle $B \in \mathcal{F}_\tau$ und $A' \in \mathfrak{A}'$ gilt, folgt die Behauptung.

(b) Hier geht man analog zu (a) vor. Wir verzichten deshalb auf die nochmalige Angabe der Details. \square

Beachtet man, dass für alle $n \geq 0$

$$\begin{aligned} \mathbf{1}_{\{\tau=n\}} \mathbb{P}(X^{(\tau)} \in \cdot | \mathcal{F}_n) &= \mathbb{E}(\mathbf{1}_{\{\tau=n, X^{(\tau)} \in \cdot\}} | \mathcal{F}_n) \\ &= \mathbb{E}(\mathbf{1}_{\{\tau=n, X^{(n)} \in \cdot\}} | \mathcal{F}_n) \\ &= \mathbf{1}_{\{\tau=n\}} \mathbb{P}(X^{(n)} \in \cdot | \mathcal{F}_n) \quad \mathbb{P}\text{-f.s.} \end{aligned}$$

und damit

$$\mathbb{P}^{X^{(\tau)}|\mathcal{F}_n} = \mathbb{P}^{X^{(n)}|\mathcal{F}_n} \quad \mathbb{P}\text{-f.s. auf } \{\tau = n\}$$

gilt, so ergibt sich offenbar bei Anwendung des vorherigen Satzes auf die Post- τ -Folge $X^{(\tau)}$:

Korollar 1.21. Gegeben eine Filtration $(\mathcal{F}_n)_{n \geq 0}$ des W-Raums $(\Omega, \mathfrak{A}, \mathbb{P})$, eine Folge $(X_n)_{n \geq 0}$ von Zufallsvariablen auf diesem und eine Stopzeit τ bezüglich $(\mathcal{F}_n)_{n \geq 0}$, gilt

$$\mathbb{P}^{X^{(\tau)}|\mathcal{F}_\tau} = \sum_{n \geq 0} \mathbb{P}^{X^{(n)}|\mathcal{F}_n} \mathbf{1}_{\{\tau=n\}} + \delta_{(\Delta, \Delta, \dots)} \mathbf{1}_{\{\tau=\infty\}} \quad \mathbb{P}\text{-f.s.},$$

vorausgesetzt, die regulär bedingten Verteilungen $\mathbb{P}^{X^{(n)}|\mathcal{F}_n}$, $n \geq 0$, existieren.

Wir haben somit die a priori keineswegs selbstverständliche *Einsetzungsregel*

$$\mathbb{P}^{X^{(\tau)}|\mathcal{F}_\tau}(\omega, \cdot) = \mathbb{P}^{X^{(\tau(\omega))}|\mathcal{F}_{\tau(\omega)}}(\omega, \cdot)$$

für \mathbb{P} -fast alle $\omega \in \Omega$.

1.4 Die starke Markov-Eigenschaft

Wir kehren zurück zur Theorie der Markov-Ketten. Die Markov-Eigenschaft lässt sich in Kürze auch wie folgt formulieren: “Bedingt unter der Vergangenheit zu einem beliebigen, aber fest gewählten Zeitpunkt n , hängt das zukünftige Verhalten der betreffenden Folge nur von ihrem gegenwärtigen Zustand ab.” Eine wichtige Verschärfung dieser Eigenschaft besteht darin, dass sie auch bei Bedingen unter der Vergangenheit zu einer Stopzeit Gültigkeit behält. Dabei wollen wir im folgenden die Vergangenheit bis zu einem Zeitpunkt allgemeiner fassen als bisher. Auf dem Weg zu der besagten Verschärfung der Markov-Eigenschaft geben wir deshalb zunächst eine Erweiterung der Definition 1.1. Sei dazu $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration des zugrundeliegenden messbaren Raums und $(M_n)_{n \geq 0}$ adaptiert bezüglich $(\mathcal{F}_n)_{n \geq 0}$, d.h. M_n ist \mathcal{F}_n -messbar für jedes $n \geq 0$. Die kanonische Filtration von $(M_n)_{n \geq 0}$ bezeichnen wir mit $(\mathcal{G}_n)_{n \geq 0}$, also $\mathcal{G}_n = \sigma(M_0, \dots, M_n) \subset \mathcal{F}_n$ für alle $n \geq 0$.

Definition 1.22. Sei $(M_n)_{n \geq 0}$ eine stochastische Folge auf einem W-Raum $(\Omega, \mathfrak{A}, \mathbb{P})$ mit (abzählbarem/endem) Zustandsraum $(\mathcal{S}, \mathfrak{S})$. Sei ferner $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration, bzgl. der $(M_n)_{n \geq 0}$ adaptiert ist. Dann heißt $(M_n)_{n \geq 0}$ (diskrete/endlliche) Markov-Kette bzgl. $(\mathcal{F}_n)_{n \geq 0}$, wenn

$$\mathbb{P}^{M_{n+1}|\mathcal{F}_n} = \mathbb{P}^{M_{n+1}|M_n} \quad \mathbb{P}\text{-f.s.} \quad (1.24)$$

für alle $n \geq 0$. Wir schreiben dann auch, dass $(M_n, \mathcal{F}_n)_{n \geq 0}$ eine (diskrete/endlliche) MK bildet.

Anmerkung 1.23. Die in Abschnitt 1.1 gegebene Definition einer MK entspricht offensichtlich der obigen, wenn man dort $\mathcal{F}_n = \mathcal{G}_n$ wählt, also die kanonische Filtration zugrundelegt. Aufgrund der Iterationsregel für bedingte Erwartungswerte und Verteilungen ([3], (51.21) in Satz 51.6) folgt sofort, dass jede MK bezüglich $(\mathcal{F}_n)_{n \geq 0}$ auch eine solche bezüglich ihrer kanonischen Filtration $(\mathcal{G}_n)_{n \geq 0}$ bildet, weil $\sigma(M_n) \subset \mathcal{G}_n \subset \mathcal{F}_n$ für alle $n \geq 0$. Entsprechendes gilt für jede Filtration $(\mathcal{F}'_n)_{n \geq 0}$ mit $\mathcal{G}_n \subset \mathcal{F}'_n \subset \mathcal{F}_n$ für alle $n \geq 0$. Eine typische Situation, in der eine MK $(M_n)_{n \geq 0}$ die Markov-Eigenschaft bezüglich einer größeren als der kanonischen Filtration erfüllt, liegt vor im Fall $\mathcal{F}_n = \sigma(M_0, \dots, M_n, M'_0, \dots, M'_n)$ für eine beliebige, von $(M_n)_{n \geq 0}$ unabhängige Folge $(M'_n)_{n \geq 0}$.

Anmerkung 1.24. Mittels der gleichen Argumente wie in Abschnitt 1.1 erhält man die Äquivalenz von (1.24) und

$$\mathbb{P}^{(M_k)_{k \geq n} | \mathcal{F}_n} = \mathbb{P}^{(M_k)_{k \geq n} | M_n} \quad \mathbb{P}\text{-f.s.}$$

für alle $n \geq 0$.

Anmerkung 1.25. Die Definition der zeitlichen Homogenität bleibt von der obigen Verallgemeinerung unberührt.

Anmerkung 1.26. Unter Rückgriff auf die im vorherigen Abschnitt gegebene anschauliche Interpretation einer Filtration $(\mathcal{F}_n)_{n \geq 0}$ als eine aufsteigende Folge der zu den sukzessiven Zeitpunkten von einem Beobachter entscheidbaren Ereignissysteme bedeutet die Beziehung

$$\mathbb{P}^{(M_k)_{k \geq n} | \mathcal{F}_n} = \mathbb{P}^{(M_k)_{k \geq n} | \mathcal{G}_n} = \mathbb{P}^{(M_k)_{k \geq n} | M_n} \quad \mathbb{P}\text{-f.s.},$$

dass ein “ \mathcal{F} -Beobachter” gegenüber einem “ \mathcal{G} -Beobachter” zu keinem Zeitpunkt zusätzliche Information über das zukünftige Verhalten der MK $(M_n)_{n \geq 0}$ besitzt und dass dasselbe auch gegenüber dem “gedächtnislosen Beobachter” gilt, der zu jedem Zeitpunkt nur den augenblicklichen Zustand der Kette kennt.

Gegeben eine MK $(M_n)_{n \geq 0}$ bezüglich $(\mathcal{F}_n)_{n \geq 0}$ mit Übergangskernen P_n , d.h. $\mathbb{P}^{M_n | \mathcal{F}_{n-1}} = P_n(M_{n-1}, \cdot)$ \mathbb{P} -f.s. für alle $n \geq 1$, wollen wir nun zeigen, dass das zukünftige Verhalten der Kette bedingt unter der Vergangenheit bis zu einer Stopzeit τ bezüglich $(\mathcal{F}_n)_{n \geq 0}$ wiederum nur vom gegenwärtigen Zustand M_τ abhängt. Häufig benutzte Beispiele von Stopzeiten für $(M_n)_{n \geq 0}$ bilden

$$\tau(i) = \inf\{n \geq 1 : M_n = i\} \tag{1.25}$$

für irgendeinen Zustand $i \in \mathcal{S}$, genannt *Rückkehr- oder Rekurrenzzeit in den Zustand i* , oder auch allgemeiner

$$\tau(A) = \inf\{n \geq 1 : M_n \in A\} \tag{1.26}$$

für ein $A \in \mathfrak{S}$, genannt *Rückkehr- oder Rekurrenzzeit in die Menge A* , wobei stets $\inf \emptyset := \infty$ vereinbart sei. Wie im vorherigen Abschnitt eingeführt, sei

$$M^{(n)} := (M_{n+k})_{k \geq 0}$$

die Post- n -Folge(Kette) für $n \in \mathbb{N}_0$ und für $n = \infty$

$$M^{(\infty)} := (\Delta, \Delta, \dots), \quad \text{insbesondere } M_\infty := \Delta$$

gesetzt, wobei Δ irgendein Element bezeichne, das im Unterschied zu dort jedoch *nicht* zu \mathcal{S} gehöre. Wir können uns Δ als einen zusätzlichen *absorbierenden* Zustand der Kette vorstellen, der deshalb manchmal auch *Friedhof* genannt wird. Formal gesehen haben wir eine Modellerweiterung durchgeführt, indem wir $(M_n)_{n \geq 0}$ nun als MK auf dem erweiterten Zustandsraum

$$(\mathcal{S}_\Delta, \mathfrak{S}_\Delta) := (\mathcal{S} \cup \{\Delta\}, \sigma(\mathfrak{S}, \{\Delta\}))$$

mit Übergangskernen

$$P_n^{(\Delta)}(x, \cdot) := \begin{cases} P_n(x, \cdot), & \text{falls } x \in \mathcal{S} \\ \delta_\Delta, & \text{falls } x = \Delta \end{cases}$$

auffassen. Da dies aber nur aus Definitheitsgründen relevant ist, werden wir darauf auch nur, wenn notwendig, zurückgreifen.

Gegeben eine Stopzeit τ bezüglich $(\mathcal{F}_n)_{n \geq 0}$, erinnern wir daran, dass die σ -Algebra der τ -Vergangenheit durch

$$\mathcal{F}_\tau = \left\{ A \in \mathfrak{A} : A \cap \{\tau = n\} \in \mathcal{F}_n \text{ für alle } n \in \overline{\mathbb{N}}_0 \right\}$$

definiert ist. Sie enthält alle Ereignisse, die beim Stoppen zum Zeitpunkt τ entscheidbar sind. Für weitere Informationen verweisen wir auf den vorherigen Abschnitt.

Satz 1.27. Sei $(\mathcal{F}_n)_{n \geq 0}$ eine Filtration des W -Raums $(\Omega, \mathfrak{A}, \mathbb{P})$ und $M = (M_n)_{n \geq 0}$ eine MK bzgl. $(\mathcal{F}_n)_{n \geq 0}$ mit Übergangskernen P_1, P_2, \dots . Dann besitzt $(M_n, \mathcal{F}_n)_{n \geq 0}$ die **starke Markov-Eigenschaft**: Für jede $(\mathcal{F}_n)_{n \geq 0}$ -Zeit τ gilt \mathbb{P} -f.s.

$$\mathbb{P}^{M^{(\tau)} | \mathcal{F}_\tau} = \mathbb{P}^{M^{(\tau)} | \tau, M_\tau} = \begin{cases} \delta_{M_\tau} \otimes \bigotimes_{n \geq 1} P_{\tau+n}, & \text{falls } \tau < \infty, \\ \delta_{(\Delta, \Delta, \dots)}, & \text{falls } \tau = \infty \end{cases} \quad (1.27)$$

sowie im zeitlich homogenen Fall mit $P_1 = P_2 = \dots =: P$ speziell

$$\mathbb{P}^{M^{(\tau)} | \mathcal{F}_\tau} = \mathbb{P}^{M^{(\tau)} | M_\tau} = \begin{cases} \delta_{M_\tau} \otimes P^\infty, & \text{falls } \tau < \infty, \\ \delta_{(\Delta, \Delta, \dots)}, & \text{falls } \tau = \infty \end{cases} \quad \mathbb{P}\text{-f.s.} \quad (1.28)$$

Beweis. Aus Korollar 1.21 und der gewöhnlichen Markov-Eigenschaft folgt

$$\mathbb{P}^{M^{(\tau)}}|_{\mathcal{F}_\tau} = \sum_{n \in \bar{\mathbb{N}}_0} \mathbf{1}_{\{\tau=n\}} \mathbb{P}^{M^{(n)}}|_{\mathcal{F}_n} = \sum_{n \in \bar{\mathbb{N}}_0} \mathbf{1}_{\{\tau=n\}} \mathbb{P}^{M^{(n)}}|_{M_n} \quad \mathbb{P}\text{-f.s.}$$

Da $\mathbf{1}_{\{\tau=n\}} \mathbb{P}^{M^{(n)}}|_{M_n}(A)$ $\sigma(\tau, M_\tau)$ -messbar ist für jedes $n \in \bar{\mathbb{N}}_0$ und $A \in \mathfrak{S}_\Delta^\infty$, gilt dasselbe für $\mathbb{P}^{M^{(\tau)}}|_{\mathcal{F}_\tau}(A)$, was wegen $\sigma(\tau, M_\tau) \subset \mathcal{F}_\tau$ die erste Gleichung in (1.27) zeigt. Die Gültigkeit der zweiten Gleichung folgt unter Verwendung von Korollar 1.4 und $\mathbb{P}^{M^{(\infty)}}|_{\mathcal{F}_\infty} = \mathbb{P}^{M^{(\infty)}}|_{M_\infty} = \delta_{(\Delta, \Delta, \dots)}$.

Ist M zeitlich homogen mit Übergangskern $P = P_1 = P_2 = \dots$, so gilt offenkundig (1.28), denn $\mathbb{P}^{M^{(\tau)}}|_{\mathcal{F}_\tau}$ gemäß (1.27) ist dann offenkundig M_τ -messbar und folglich auch eine Version von $\mathbb{P}^{M^{(\tau)}}|_{M_\tau}$. \square

Aus (1.27) bzw. (1.28) folgt direkt

$$\mathbb{E}(f(M^{(\tau)})|_{\mathcal{F}_\tau}) = \mathbb{E}(f(M^{(\tau)})|\tau, M_\tau) \quad \text{bzw.} \quad = \mathbb{E}(f(M^{(\tau)})|M_\tau) \quad \mathbb{P}\text{-f.s.}$$

für jede $\mathbb{P}^{M^{(\tau)}}$ -quasi-integrierbare numerische Funktion $f : (\mathcal{S}_\Delta^\infty, \mathfrak{S}_\Delta^\infty) \rightarrow (\bar{\mathbb{R}}, \bar{\mathcal{B}})$.

Setzen wir wieder $M_{0,n} = (M_0, \dots, M_n)$ für $n \geq 0$ und $M_{0,\infty} = (M_n)_{n \geq 0}$, so ergibt sich als Pendant zu Satz 1.5:

Satz 1.28. *Eine stochastische Folge $(M_n)_{n \geq 0}$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$ besitzt genau dann die starke Markov-Eigenschaft bezüglich der Filtration $(\mathcal{F}_n)_{n \geq 0}$, wenn sie adaptiert ist und wenn für jede $(\mathcal{F}_n)_{n \geq 0}$ -Zeit τ die σ -Algebra \mathcal{F}_τ und $M^{(\tau)}$ bedingt unter (τ, M_τ) stochastisch unabhängig sind, d.h.*

$$\mathbb{P}(A \cap \{M^{(\tau)} \in C\} | \tau, M_\tau) = \mathbb{P}(A | \tau, M_\tau) \mathbb{P}(M^{(\tau)} \in C | \tau, M_\tau) \quad \mathbb{P}\text{-f.s.} \quad (1.29)$$

für alle $A \in \mathcal{F}_\tau$ und $C \in \mathfrak{S}_\Delta^\infty$. Für eine zeitlich homogene MK $(M_n)_{n \geq 0}$ bleibt (1.29) bei Bedingen unter M_τ anstelle von (τ, M_τ) gültig.

Beweis. Der Beweis bildet im wesentlichen eine Adaption des Beweises von Satz 1.5 und bleibt dem Leser als Übung überlassen. \square

Anmerkung 1.29. Bezeichnet $(\mathcal{F}_n)_{n \geq 0}$ im vorherigen Satz die kanonische Filtration von $(M_n)_{n \geq 0}$ und beachtet man, dass dann $\mathcal{F}_\tau = \sigma(\tau, M_{0:\tau})$ gemäß Satz 1.18(f) gilt, so besitzt $(M_n)_{n \geq 0}$ in vollkommener Analogie zum Fall fester Zeit die starke Markov-Eigenschaft genau dann, wenn $(\tau, M_{0:\tau})$ und $M^{(\tau)}$ für jede $(\mathcal{F}_n)_{n \geq 0}$ -Zeit τ bedingt unter (τ, M_τ) (im Fall zeitlicher Homogenität auch unter M_τ) stochastisch unabhängig sind, d.h.

$$\mathbb{P}((\tau, M_{0:\tau}), M^{(\tau)} | \tau, M_\tau) = \mathbb{P}(\tau, M_{0:\tau} | \tau, M_\tau) \otimes \mathbb{P}^{M^{(\tau)} | \tau, M_\tau} \quad \mathbb{P}\text{-f.s.} \quad (1.30)$$

Betrachten wir zum Abschluss noch einmal eine zeitlich homogene MK $M = (M_n)_{n \geq 0}$ in einem Standardmodell $(\Omega, \mathfrak{A}, M, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})})$. In diesem Fall erhalten wir mit (1.28), dass für alle $\lambda \in \mathcal{P}(\mathcal{S})$ und $(\mathcal{F}_n)_{n \geq 0}$ -Zeiten τ

$$\mathbb{P}_\lambda^{M^{(\tau)}|\mathcal{F}_\tau} = \mathbb{P}_\lambda^{M^{(\tau)}|M_\tau} = \mathbb{P}_{M_\tau}^M \quad \mathbb{P}_\lambda\text{-f.s.} \quad (1.31)$$

oder ausführlicher

$$\mathbb{P}_\lambda^{M^{(\tau)}|\mathcal{F}_\tau}(\omega, A) = \mathbb{P}_\lambda^{M^{(\tau)}|M_\tau}(\omega, A) = \mathbb{P}_{M_\tau(\omega)}^M(A)$$

für \mathbb{P}_λ -fast alle $\omega \in \Omega$ und $A \in \mathcal{G}_\Delta^\infty$ gilt.

Beispiel 1.30. Sei $M = (M_n)_{n \geq 0}$ eine zeitlich homogene DMK in einem Standardmodell $(\Omega, \mathfrak{A}, M, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})})$ und $\tau(i)$ die Ersteintrittszeit in einen Zustand $i \in \mathcal{S}$ (\mathfrak{E}^∞ (1.26)), folglich eine Stopzeit bezüglich der kanonischen Filtration $(\mathcal{G}_n)_{n \geq 0}$. Dann gilt aufgrund der starken Markov-Eigenschaft, genauer gemäß (1.31),

$$\mathbb{P}_\lambda^{M^{(\tau(i))}|\mathcal{G}_{\tau(i)}} = \mathbb{P}_i^M \quad \mathbb{P}_\lambda\text{-f.s. auf } \{\tau(i) < \infty\}$$

für jede Anfangsverteilung λ , d.h., die MK verhält sich in Verteilung nach erstmaligem Erreichen des Zustands i unabhängig vom vorherigen Verlauf anschließend genauso als wäre sie in i gestartet. Setzen wir $f_{ij}^{(n)} = P_i(\tau(j) = n)$ für $i, j \in \mathcal{S}$ und $n \geq 1$, so ergibt sich als Anwendung folgende nützliche Beziehung zwischen den $f_{ij}^{(n)}$ und den n -Schritt-Übergangswahrscheinlichkeiten $p_{ij}^{(n)}$:

$$p_{ij}^{(n)} = \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)} \quad (1.32)$$

für alle $n \geq 1$. Zum Beweis notieren wir:

$$\begin{aligned} p_{ij}^{(n)} &= P_i(M_n = j) \\ &= \sum_{k=1}^n P_i(\tau(j) = k, M_{\tau(j)+n-k} = j) \\ &= \sum_{k=1}^n \int_{\{\tau(j)=k\}} \mathbb{P}(M_{\tau(j)+n-k} = j | \mathcal{G}_{\tau(j)}) d\mathbb{P}_i \\ &= \sum_{k=1}^n \int_{\{\tau(j)=k\}} P_j(M_{n-k} = j) d\mathbb{P}_i \\ &= \sum_{k=1}^n P_i(\tau(j) = k) P_j(M_{n-k} = j) \\ &= \sum_{k=1}^n f_{ij}^{(k)} p_{jj}^{(n-k)}, \end{aligned}$$

wobei die starke Markov-Eigenschaft offenbar beim Übergang von der dritten zur vierten Zeile verwendet wurde.

1.5 Stationäre Maße und Verteilungen

Im folgenden sei $M = (M_n)_{n \geq 0}$ stets eine zeitlich homogene MK in einem Standardmodell $(\Omega, \mathfrak{A}, (M_n)_{n \geq 0}, (\mathbb{P}_\lambda)_{\lambda \in \mathcal{P}(\mathcal{S})})$ mit Zustandsraum $(\mathcal{S}, \mathfrak{S})$ und Übergangskern P . Eine Frage, die allgemein bei stochastischen Prozessen sehr häufig von zentralem Interesse ist, lautet:

Wie verhält sich die Verteilung der Prozessvariablen oder auch des Post- t -Prozesses bei gegen unendlich strebender Zeit t ?

Auf M bezogen bedeutet dies:

Welche Aussagen lassen sich über $\mathbb{P}_\lambda^{M_n}$ und $\mathbb{P}_\lambda^{M^{(n)}}$ für $n \rightarrow \infty$ machen?

In diesem Zusammenhang spielen die Begriffe “stationäres Maß” und “stationäre Verteilung” eine große Rolle und werden deshalb als nächstes präzisiert:

Definition 1.31. Ein σ -endliches Maß π auf $(\mathcal{S}, \mathfrak{S})$ heißt *stationäres* oder *invariantes Maß* der MK M , wenn $\pi \neq 0$ und

$$\mathbb{P}_\pi^{M_1} = \pi P = \pi \quad (1.33)$$

gilt, was ausführlicher geschrieben

$$\mathbb{P}_\pi(M_1 \in A) = \int_{\mathcal{S}} P(x, A) \pi(dx) = \pi(A) \quad \text{für alle } A \in \mathfrak{S}$$

bedeutet. Hat π Gesamtmasse 1, nennt man π auch *stationäre* oder *invariante Verteilung* von M .

Beachtet man, dass stets $\mathbb{P}_\pi^{M_0} = \pi$ gilt, so bedeutet (1.33) nichts anderes als, dass M_0 und M_1 unter \mathbb{P}_π dieselben Bildmaße/Verteilungen besitzen (Invarianz). Unter Benutzung von

$$\mathbb{P}_\pi^{M_n} = \pi P^{(n)} = \pi P^{\circ n} = (\pi P) P^{\circ(n-1)} = \pi P^{\circ(n-1)}$$

für alle $n \geq 1$ (☞ (1.33)) folgt sofort per Induktion über n

Lemma 1.32. π ist genau dann stationäres Maß der MK M , wenn $\pi \neq 0$ und für alle $n \geq 0$

$$\mathbb{P}_\pi^{M_n} = \pi P^{(n)} = \pi \quad (1.34)$$

gilt, was ausführlicher geschrieben

$$\mathbb{P}_\pi(M_n \in A) = \int_{\mathcal{S}} P^{(n)}(x, A) \pi(dx) = \pi(A) \quad \text{für alle } A \in \mathfrak{G}$$

bedeutet.

Unter \mathbb{P}_π stimmen die Bildmaße aller M_n folglich überein, was im Fall einer Verteilung π auch wie folgt formuliert werden kann:

Besitzt M eine stationäre Verteilung π , so hat M_n unter \mathbb{P}_π , d.h. bei Anfangsverteilung π , für jedes $n \geq 0$ genau diese Verteilung π .

Wir haben es hier mit einer besonders starken Form von *Stabilität* oder *Gleichgewicht* zu tun, die sich aufgrund der zeitlichen Homogenität auch auf die Post- n -Prozesse überträgt:

Satz 1.33. π ist genau dann stationäres Maß der MK M , wenn $\pi \neq 0$ und

$$\mathbb{P}_\pi^{M^{(n)}} = \mathbb{P}_\pi^M = \pi \otimes P^\infty \quad (1.35)$$

für alle $n \geq 0$ gilt.

Beweis. Zu zeigen ist nur die erste Gleichung in (1.35). Gemäß (1.17) in Satz 1.9 gilt für alle $n \geq 0$ und $s \in \mathcal{S}$

$$\mathbb{P}_s^{M^{(n)}|M_n} = \mathbb{P}_{M_n}^M \quad \mathbb{P}_s\text{-f.s.},$$

so dass unter Benutzung von (1.34)

$$\begin{aligned} \mathbb{P}_\pi^{M^{(n)}}(\cdot) &= \int_{\mathcal{S}} \mathbb{P}_s^{M^{(n)}}(\cdot) \pi(ds) \\ &= \int_{\mathcal{S}} \int_{\mathcal{S}} \mathbb{P}_s^{M^{(n)}|M_n=x}(\cdot) \mathbb{P}_s^{M_n}(dx) \pi(ds) \\ &= \int_{\mathcal{S}} \mathbb{P}_x^M(\cdot) \mathbb{P}_\pi^{M_n}(dx) \\ &= \int_{\mathcal{S}} \mathbb{P}_x^M(\cdot) \pi(dx) \\ &= \mathbb{P}_\pi^M(\cdot) \end{aligned}$$

für alle $n \geq 0$, d.h. die Behauptung folgt. \square

Eine Folge M mit der Eigenschaft (1.35) für alle $n \geq 0$ heißt *stationär* unter \mathbb{P}_π , was die Namensgebung für π erklärt.

Stationäre Maße und Verteilungen spielen in der Theorie der MK eine wichtige Rolle, müssen aber weder existieren noch eindeutig bestimmt sein ($\mathbb{E}^{\mathfrak{S}}$ hierzu Unterabschnitt 2.5.4). Die Eindeutigkeit betreffend notieren wir, dass die Menge \mathfrak{E} aller stationären Maße einer MK M einen positiven Halbraum bildet, d.h. $\pi_1, \pi_2 \in \mathfrak{E}$ impliziert $c_1\pi_1 + c_2\pi_2 \in \mathfrak{E}$ für alle $c_1, c_2 > 0$. Stationäre Maße sind also im günstigsten Fall bis auf ein skalares Vielfaches eindeutig bestimmt, was bedeutet, dass \mathfrak{E} eindimensional ist. Die Menge \mathfrak{E}^* der stationären Verteilungen von M bildet eine konvexe, möglicherweise leere Teilmenge von \mathfrak{E} . Sie enthält offenbar genau ein Element, wenn \mathfrak{E} eindimensional ist und aus lauter endlichen Maßen besteht. Die Bedeutung stationärer Verteilungen im Zusammenhang mit dem asymptotischen Verhalten von MK verdeutlicht das folgende einfache Lemma.

Lemma 1.34. *Sei $M = (M_n)_{n \geq 0}$ eine MK, für die $\lambda, \nu \in \mathcal{P}(\mathcal{S})$ existieren, so dass*

$$\lim_{n \rightarrow \infty} \mathbb{P}_\lambda(M_n \in A) = \nu(A) \quad (1.36)$$

für alle $A \in \mathfrak{S}$. Dann ist ν eine stationäre Verteilung der Kette, d.h. $\nu \in \mathfrak{E}^$.*

Beweis. Per Funktions-Erweiterungsargument folgt aus (1.36)

$$\lim_{n \rightarrow \infty} \mathbb{E}_\lambda f(M_n) = \int_{\mathcal{S}} f(s) \nu(ds)$$

für alle $f \in b\mathfrak{S}$, dem Raum der beschränkten \mathfrak{S} -messbaren reellen Funktionen. Damit erhalten wir aber unter Benutzung der Markov-Eigenschaft

$$\begin{aligned} \nu(A) &= \lim_{n \rightarrow \infty} \mathbb{P}_\lambda(M_{n+1} \in A) = \lim_{n \rightarrow \infty} \int_{\mathcal{S}} P(s, A) \mathbb{P}_\lambda^{M_n}(ds) \\ &= \lim_{n \rightarrow \infty} \mathbb{E}_\lambda P(M_n, A) = \int_{\mathcal{S}} P(s, A) \nu(ds), \end{aligned}$$

d.h. $\nu \in \mathfrak{E}^*$. □

Der Verteilungslimes von M_n unter irgendeinem \mathbb{P}_λ definiert also stets eine stationäre Verteilung. Diese ist außerdem eindeutig bestimmt, wenn (1.36) für jede Anfangsverteilung λ gilt, da dann speziell für $\mu \in \mathfrak{E}^*$ und alle $A \in \mathfrak{S}$

$$\mu(A) = \mathbb{P}_\mu(M_n \in A) \xrightarrow{n \rightarrow \infty} \nu(A),$$

d.h. $\mu = \nu$ folgt.

Wenden wir uns abschließend noch kurz dem diskreten Fall zu, in dem \mathcal{S} abzählbar ist. Dann lässt sich jedes σ -endliche Maß λ mit dem Vektor $(\lambda_i)_{i \in \mathcal{S}}$ mit den endlichen, nichtnegativen Komponenten $\lambda_i = \lambda(\{i\})$ identifizieren und der Übergangskern P mit der Übergangsmatrix $P = (p_{ij})_{i, j \in \mathcal{S}}$, $p_{ij} = P(i, \{j\})$.

Satz 1.35. Ein σ -endliches Maß $\pi = (\pi_i)_{i \in \mathcal{S}}$ definiert genau dann ein stationäres Maß der DMK M mit Übergangsmatrix P , wenn $\pi \neq 0$ und

$$\pi_j = \sum_{i \in \mathcal{S}} \pi_i p_{ij} = \mathbb{P}_\pi(M_1 = j) \quad (1.37)$$

für alle $j \in \mathcal{S}$, d.h. $\pi = \pi P$ gilt.

Beweis. Es ist nichts zu zeigen, da π durch seine Werte $\pi_j = \pi(\{j\})$ eindeutig festgelegt wird. \square

Ein stationäres Maß erhält man demnach im diskreten Fall durch Lösen des durch (1.37) für alle $j \in \mathcal{S}$ gegebenen linearen Gleichungssystems. Aus (1.37) folgt natürlich weiter $\pi = \pi P^n$ (vgl. Lemma 1.32), d.h.

$$\pi_j = \sum_{i \in \mathcal{S}} \pi_i p_{ij}^{(n)} = \mathbb{P}_\pi(M_n = j)$$

für alle $j \in \mathcal{S}$ und $n \geq 0$.

Ist der Zustandsraum \mathcal{S} sogar endlich, d.h. M eine EMK, so ist jedes stationäre Maß π notwendigerweise endlich mit Gesamtmasse $\|\pi\| = \sum_{i \in \mathcal{S}} \pi_i$ und dessen Normierung $\pi^* = \pi/\|\pi\|$ eine stationäre Verteilung. Man beachte ferner, dass π in diesem Fall genau dann ein stationäres Maß definiert, wenn π einen nichtnegativen linken Eigenvektor zum Eigenwert 1 der endlichen Matrix P bildet. Wir werden im nächsten Kapitel zeigen, dass eine EMK stets mindestens eine stationäre Verteilung besitzt.