

Diplomarbeit

Stabilität des nichtlinearen  
TAR-ARCH-Modells und die  
Piggyback-Methode

Timo Heinrich

September 2008

Fachbereich Mathematik  
der Westfälischen Wilhelms-Universität Münster  
Betreuer: Prof. Dr. G. Alsmeyer



# Inhaltsverzeichnis

<b>Einleitung</b>	<b>1</b>
<b>1 Markov-Ketten</b>	<b>3</b>
1.1 Markov-Ketten . . . . .	3
1.2 Irreduzibilität und $\varphi$ -Irreduzibilität . . . . .	6
1.3 Periodizität . . . . .	9
1.4 Rekurrenz und Transienz . . . . .	12
1.5 Harris-Rekurrenz . . . . .	14
<b>2 Stabilität und geometrische Ergodizität</b>	<b>15</b>
2.1 Rekurrenz und der Drift $\Delta V$ einer Markov-Kette . . . . .	15
2.2 Ergodizität . . . . .	17
2.3 Geometrische Ergodizität . . . . .	20
<b>3 Die Piggyback-Methode</b>	<b>27</b>
3.1 ARCH-Prozesse . . . . .	27
3.1.1 Der ARCH(q)-Prozess . . . . .	27
3.1.2 Der AR(p)-Prozess . . . . .	28
3.1.3 Der AR(p)-ARCH(q)-Prozess . . . . .	28
3.1.4 Der TAR-ARCH-Prozess . . . . .	29
3.2 Annahmen an das TAR-ARCH-Modell . . . . .	30
3.3 Einleitung zur Piggyback-Methode . . . . .	31
3.4 Der homogene, normierte Prozess $(\theta_t^*)_{t \geq 0}$ . . . . .	33
3.4.1 Die gleichmäßige Ergodizität von $(\theta_t^*)_{t \geq 0}$ . . . . .	33
3.4.2 Die near-equilibrium-Gleichung . . . . .	41
3.4.3 Die Stetigkeit von $\nu$ . . . . .	43
3.4.4 Der Übergang vom homogenen Prozess $(\theta_t^*)_{t \geq 0}$ zu $(\theta_t)_{t \geq 0}$ . . . . .	48
3.5 Der Prozess $(X_t)_{t \geq 0}$ . . . . .	54
3.6 Der Lyapounov-Exponent und die Transienz von $(X_t)_{t \geq 0}$ . . . . .	62
3.7 Ein einfaches Beispiel . . . . .	66
<b>Symbolverzeichnis</b>	<b>69</b>
<b>Literaturverzeichnis</b>	<b>71</b>



# Einleitung

In dieser Arbeit wird ein stochastischer Prozess der Form

$$\xi_t = a(\xi_{t-1}, \dots, \xi_{t-p}) + b(\xi_{t-1}, \dots, \xi_{t-p})e_t$$

bzw. der Zustandsvektor  $X_t = (\xi_{t-1}, \dots, \xi_{t-p+1})$  untersucht.  $a$  und  $b$  seien stückweise stetige Funktionen und  $(e_t)_{t \geq 1}$  eine Folge unabhängig identisch verteilter Zufallsvariablen. Die Unstetigkeitsstellen von  $a$  und  $b$  nennen wir Schwellenwerte (Schwellenwert, engl.: threshold), was gleichzeitig die Namensgebung des Modells begründet. Das TAR-ARCH-Modell (threshold autoregressive - autoregressive conditionally heteroscedastic) stellt eine Verallgemeinerung des von Robert F. Engle 1982 eingeführten ARCH-Modells dar. Bei dieser Art von Modellen ist es möglich die bedingte Varianz abhängig von der Vergangenheit des Prozesses zu modellieren. Dies hat sich als sehr erfolgreich bei der Untersuchung und Modellierung von Finanzzeitreihen erwiesen, da spezifische Eigenschaften wie zum Beispiel das Auftreten von Volatilitätsclustern gut reproduziert werden können. Ziel der Untersuchung solcher Modelle sind meist Aussagen über das Langzeitverhalten bzw. die Stabilität derselbigen. Die Existenz von Schwellenwerten sowie die Nichtlinearität der Funktionen  $a$  und  $b$  erschwert allerdings die Übertragung der Ergebnisse und der Beweismethoden ähnlicher Modelle auf das TAR-ARCH-Modell. Aus diesem Grund wird hier ein neuer Ansatz, die *Piggyback-Methode* (Huckepack, engl.: piggyback), vorgestellt. Dieses Verfahren wurde von Daren B. H. Cline und Huay-min H. Pu entwickelt. Der Artikel "Stability and the Lyapounov exponent of threshold AR-ARCH models" derselben Autoren diente als Grundlage für die vorliegende Arbeit. Im Zuge dieser Methode wird die Stabilität eines "einfacheren" Hilfsprozess ausgenutzt, um Stabilität für den eigentlichen Prozess nachzuweisen. Dieses Vorgehen bringt zwei wesentliche Vorteile mit sich. Zum einen sind nicht solch starke Annahmen notwendig, wie es sonst für nichtlineare Modelle der Fall ist, zum anderen stellt sich heraus, dass das gefundene Stabilitätskriterium scharf ist, das heißt Stabilität liegt genau dann vor, wenn das Kriterium erfüllt ist. Um Stabilität nachzuweisen, benutzt man oft ein Drift-Kriterium, in unserem Fall

$$\Delta V(x) \leq -\beta V(x) + b \mathbf{1}_C(x), \quad x \in X,$$

für eine nichtnegative Funktion  $V$ , eine petite Menge  $C$  sowie zwei Konstanten  $\beta > 0$  und  $b < \infty$ . Die Schwierigkeit bei der Bedingung besteht darin eine geeignete Funktion  $V$  zu finden. Die Piggyback-Methode verwendet deshalb einen Hilfsprozess, um die Drift-Bedingung nachzuweisen. Im folgenden wird ein kurzer Überblick über das Vorgehen dieser Methode gegeben. Eine genauere Skizze des Beweisverfahrens findet sich in Kapitel 3 auf Seite 32.

- Als erstes werden der homogene Prozess  $(X_t^*)_{t \geq 0}$  und der normierte homogene Prozess  $(\theta_t^*)_{t \geq 0}$  eingeführt, wobei  $\theta_t^* = \frac{X_t^*}{\|X_t^*\|}$ .
- Wir zeigen, dass der homogene, normierte Prozess  $(\theta_t^*)_{t \geq 0}$  gleichmäßig ergodisch ist.
- Aufgrund der gleichmäßigen Ergodizität von  $(\theta_t^*)_{t \geq 0}$  lässt sich die near-equilibrium-Gleichung für eine Testfunktion  $\nu(\theta)$  beweisen. Die near-equilibrium-Gleichung hat bereits eine ähnliche Struktur wie die Gleichung, die für den Nachweis der Drift-Bedingung des Prozesses  $(X_t)_{t \geq 0}$  benötigt wird. Dieser Schritt bildet den Schlüssel der Piggyback-Methode.
- Mit Hilfe der near-equilibrium-Gleichung sowie weiterer Eigenschaften der Prozesse  $(\theta_t^*)_{t \geq 0}$  und  $(\theta_t)_{t \geq 0}$  konstruiert man schließlich eine Testfunktion  $V$ , welche die Drift-Bedingung für  $(X_t)_{t \geq 0}$  erfüllt.
- Als wesentlichen Vorteil der Piggyback-Methode zeigen wir, dass die Stabilitätsbedingung scharf ist.

Für den eben skizzierten Beweis ist Vorwissen notwendig, deswegen sieht der Aufbau der vorliegenden Arbeit folgendermaßen aus:

Im ersten Kapitel werden wichtige Begriffe und Eigenschaften von Markov-Ketten auf einem allgemeinen Zustandsraum vorgestellt. Den Ausgangspunkt hierfür bildet die Theorie der diskreten Markov-Ketten.

Im zweiten Kapitel werden verschiedene Stabilitätsbegriffe und -kriterien zusammengefasst, um letztlich den wichtigen Begriff V-gleichmäßige Ergodizität bereitzustellen.

Im dritten Kapitel betrachten wir das zu untersuchende TAR-ARCH-Modell. Als Motivation wird ein kurzer Blick auf einfachere ARCH-Modelle und deren Stabilität geworfen. Weiterhin werden die Annahmen vorgestellt, die notwendig für die Untersuchung des Modells sind. Der eigentliche Beweis gliedert sich in zwei Teile. Im ersteren werden die Eigenschaften des normierten Prozesses  $(\theta_t^*)_{t \geq 0}$  bzw.  $(\theta_t)_{t \geq 0}$  erarbeitet, die im zweiten Teil für die Stabilitätsuntersuchung des Prozesses  $(X_t)_{t \geq 0}$  benötigt werden. Das daraus resultierende Stabilitätskriterium stellt das Hauptergebnis dieser Arbeit dar. Der darauffolgende Abschnitt beschäftigt sich mit der Transienz des Prozesses, falls das Stabilitätskriterium verletzt ist. Abschließend wird ein einfaches Beispiel präsentiert, bei dem das Kriterium leicht zu berechnen ist. Eine kleine Simulation für dieses Beispiel veranschaulicht dabei die Rekurrenz bzw. Transienz des Prozesses.

Ich bedanke mich bei Herrn Prof. Dr. G. Alsmeyer für die interessante Themenstellung und die Betreuung während der Entstehungsphase. Mein Dank gilt außerdem den Menschen, die mich bei der Anfertigung dieser Arbeit unterstützt haben.

# 1 Markov-Ketten

In diesem Abschnitt werden Markov-Ketten definiert und wichtige Eigenschaften vorgestellt, um diese zu untersuchen. Eine Markov-Kette ist ein Prozess, dessen zukünftiger Zustand nur von dem gegenwärtigen Zustand abhängt. Eine wichtige Frage bei Markov-Ketten ist, ob eine bestimmte Menge des Zustandsraumes von dem Prozess erreicht wird. Ist jeder Zustand des Zustandsraumes erreichbar, für einen beliebigen Startpunkt, dann spricht man von Irreduzibilität. Eine weitere wichtige Frage ist, wie oft eine Menge besucht wird. Geschieht dies unendlich oft mit positiver Wahrscheinlichkeit, so spricht man von Rekurrenz, andernfalls von Transienz. Die Struktur des Zustandsraum auf dem die Markov-Kette definiert ist, spielt eine wichtige Rolle für die weitere Untersuchung des Prozesses. Der einfachste Fall ist ein *endlicher* bzw. *abzählbarer* Zustandsraum. Die eben genannten Begriffe werden jeweils kurz für den abzählbaren Zustandsraum vorgestellt und dann auf den schwierigeren *allgemeinen* Zustandsraum übertragen. Später betrachten wir Zustandsräume, die eine zusätzliche Struktur, insbesondere eine Metrik, besitzen. Dies ist für uns der interessante Fall, da der TAR-ARCH-Prozess auf dem Zustandsraum  $\mathbb{R}^p$  definiert ist. Dieses Kapitel stützt sich hauptsächlich auf die Quelle [12], wobei für die Theorie der diskreten Markov-Ketten zusätzlich die Quellen [11] und [1] herangezogen wurde.

## 1.1 Markov-Ketten

Der zugrundeliegende Zustandsraum einer Markov-Kette ist entscheidend für die Untersuchung des Prozesses. Für unsere Zwecke genügt es zunächst zwei Fälle zu unterscheiden.

- i) Der Zustandsraum  $S$  heißt *abzählbar*, wenn  $S$  abzählbar viele Elemente besitzt, z.B.  $S = \mathbb{Z}$ . Dann benutzt man die  $\sigma$ -Algebra aller Untermengen von  $S$ .
- ii) Der Zustandsraum  $S$  heißt *allgemein*, wenn er mit einer abzählbar erzeugten  $\sigma$ -Algebra ausgestattet ist, z.B.  $S = \mathbb{R}$  mit der Borel- $\sigma$ -Algebra  $\mathcal{B}(\mathbb{R})$ .

Die Definition einer Markov-Kette lässt sich unabhängig von der Wahl des Zustandsraumes formulieren. Hierfür setzen wir voraus, dass die Übergangskerne

$$\mathbb{P}_{n+1}((s_0, s_1, \dots, s_n), \cdot) = P(\Phi_{n+1} \in \cdot | (\Phi_0, \dots, \Phi_n) = (s_0, \dots, s_n)), \quad s_0, \dots, s_n \in S,$$

existieren, wobei  $\mathbb{P}_0 = P(\Phi_0 = \cdot)$  die Startverteilung sei.

**Definition 1.1.** *Ein stochastischer Prozess  $\Phi$  mit Werten in dem Zustandsraum  $(S, \mathcal{B}(S))$  heißt Markov-Kette, falls*

$$P(\Phi_{n+1} \in \cdot | \Phi_0 = s_0, \dots, \Phi_{s_{n-1}} = s_{n-1}, \Phi_n = s) = P(\Phi_{n+1} \in \cdot | \Phi_n = s). \quad (1.1)$$

## 1 Markov-Ketten

Hängen die bedingten Verteilungen nicht vom Zeitparameter  $n$  ab, existiert also ein Übergangskern  $\mathbb{P}$  mit

$$P(\Phi_{n+1} = \cdot | \Phi_n) = \mathbb{P}(\Phi_n, \cdot),$$

nennen wir die Markov-Kette zeitlich homogen.

Wir betrachten im folgenden nur homogene Markov-Ketten. Im abzählbaren Fall ist eine Markov-Kette durch ihre elementaren Übergangswahrscheinlichkeiten  $\mathbf{P} = (p_{ij})_{i,j \in S}$  mit  $p_{ij} := \mathbb{P}(i, \{j\})$  festgelegt. Offensichtlich erfüllt der Übergangskern die Eigenschaften

$$p_{ij} \geq 0, \forall i, j \in S \text{ und } \sum_{j \in S} p_{ij} = 1, \forall i \in S.$$

Legt man einen allgemeinen Zustandsraum zugrunde, ist es nicht sinnvoll die Übergangswahrscheinlichkeit in eine Punktmenge  $\{j\} \in \mathcal{B}(S)$  zu betrachten, da diese für alle  $i \in S$  identisch 0 sein könnte. Aus diesem Grund führt man stochastische Übergangskerne ein.

**Definition 1.2.**  $\mathbb{P} = \{\mathbb{P}(s, A), s \in S, A \in \mathcal{B}(S)\}$  heißt stochastischer Übergangskern, falls

- i)  $\mathbb{P}(\cdot, A)$  für alle  $A \in \mathcal{B}(S)$  eine nichtnegative Funktion ist und
- ii)  $\mathbb{P}(s, \cdot)$  ein Wahrscheinlichkeitsmaß für alle  $s \in S$  ist.

Erfüllt ein stochastischer Übergangskern die Einschränkung  $\mathbb{P}(s, S) \leq 1, \forall s \in S$ , so heißt er substochastischer Übergangskern.

Für zwei Kerne  $\mathbb{P}_1 : (\Omega_0, \mathcal{A}_0) \rightarrow (\Omega_1, \mathcal{A}_1)$  und  $\mathbb{P}_2 : (\Omega_1, \mathcal{A}_1) \rightarrow (\Omega_2, \mathcal{A}_2)$  definieren wir die Hintereinanderschaltung von  $\mathbb{P}_1$  und  $\mathbb{P}_2$  durch

$$\mathbb{P}_1 \circ \mathbb{P}_2(\omega_0, A_2) = \int_{\Omega_1} \mathbb{P}_2(\omega_1, A_2) \mathbb{P}_1(\omega_0, d\omega_1).$$

Offensichtlich bildet  $\mathbb{P}_1 \circ \mathbb{P}_2$  einen stochastischen Übergangskern von  $(\Omega_0, \mathcal{A}_0)$  nach  $(\Omega_2, \mathcal{A}_2)$ . Allgemein erhält man für die Hintereinanderschaltung von  $\mathbb{P}_k = P(\Phi_k \in \cdot | \Phi_{k-1})$ ,  $k = 1, \dots, n$ ,

$$\mathbb{P}_1 \circ \dots \circ \mathbb{P}_n = \int_S \dots \int_S \mathbb{P}_n(s_{n-1}, A_n) \mathbb{P}_{n-1}(s_{n-2}, ds_{n-1}) \dots \mathbb{P}_1(s_0, ds_1).$$

Man sieht sofort ein, dass  $P(\Phi_n \in \cdot) = \mathbb{P}_0 \circ \dots \circ \mathbb{P}_n$  und  $P(\Phi_{n+k} \in \cdot | \Phi_n) = \mathbb{P}_{n+1} \circ \dots \circ \mathbb{P}_{n+k}$ . Im zeitlich homogenen Fall ergibt sich für die n-Schritt Übergangswahrscheinlichkeit

$$\mathbb{P}^n := \mathbb{P} \circ \dots \circ \mathbb{P} = P(\Phi_{m+n} \in \cdot | \Phi_m), \tag{1.2}$$

wobei  $\mathbb{P}^0(s, A) = \delta_s(A)$ . Aus der Markov-Eigenschaft (1.1) und dem n-Schritt Übergangskern (1.2) lässt sich die Chapman-Kolmogorov-Gleichung herleiten.



## 1 Markov-Ketten

**Satz 1.3.** (Chapman-Kolmogorov-Gleichung) Für eine zeitlich homogene Markov-Kette  $\Phi$  mit Zustandsraum  $(S, \mathcal{B}(S))$  erfüllen die  $n$ -Schritt Übergangskerne  $\mathbb{P}^n$ ,  $n \geq 0$ , die Halbgruppeneigenschaft

$$\mathbb{P}^{m+n} = \mathbb{P}^m \mathbb{P}^n,$$

für alle  $m, n \geq 0$ , das heißt

$$P(\Phi_{k+m+n} \in A | \Phi_k = s) = \int_S P(\Phi_{k+m+n} \in A | \Phi_{k+m} = t) P(\Phi_{k+m} \in dt | \Phi_k = s)$$

gilt für alle  $s \in S$ ,  $A \in \mathcal{B}(S)$ . Schreiben wir  $P^k(s, A) := P(\Phi_k \in A | \Phi_0 = s)$ , so ergibt sich

$$P^{m+n}(s, A) = \int_S P^n(t, A) P^m(s, dt), \quad s \in S, A \in \mathcal{B}(S). \quad (1.3)$$

Beweis: Beachtet man, dass die Hintereinanderschaltung eine assoziative Verknüpfung ist, so folgt die Behauptung direkt aus der Definition von  $\mathbb{P}^n$ .  $\square$

**Bemerkung 1.4.** Ein allgemeinerer Prozess  $\xi = (\xi_0, \xi_1, \dots)$  könnte auch von einer festen Anzahl von Punkten in der Vergangenheit abhängen, also

$$P(\xi_{n+m} \in A | \xi_j, j \leq n) = P(\xi_{n+m} \in A | \xi_j, j = n, n-1, \dots, n-k+1).$$

Dieser Prozess ist offenbar keine Markov-Kette, da die Übergangswahrscheinlichkeit von  $k$  Werten aus der Vergangenheit abhängt. Definiert man dagegen den Prozess  $\Phi = (\Phi_n)_{n \geq 0}$  mit  $\Phi_n = (\xi_n, \dots, \xi_{n-k+1})$ , dann ist  $\Phi$  eine Markov-Kette.

Beweis: Anschaulich ist dies klar, denn der Prozess  $\Phi_{n-1}$  enthält alle Zustände  $\xi_{n-1}, \dots, \xi_{n-k}$  von denen der Zustand  $\xi_n$  bzw.  $\Phi_n$  abhängt. Formal muss gezeigt werden, dass

$$P(\Phi_n \in A | \Phi_{n-1} = s_{n-1}, \dots, \Phi_0 = s_0) = P(\Phi_n \in A | \Phi_{n-1} = s_{n-1}).$$

Der Zustandsvektor  $s_n$  sei definiert durch  $s_n = (t_n, t_{n-1}, \dots, t_{n-k+1}) \in S^k$ . Schreibt man die Bedingungen auf der linken Seite aus, erhält man

$$\begin{aligned} P(\Phi_n \in A | \Phi_{n-1} = s_{n-1}, \dots, \Phi_0 = s_0) &= P(\Phi_n \in A | \xi_{n-1} = t_{n-1}, \dots, \xi_0 = t_0) \\ &= P(\Phi_n \in A | \xi_{n-1} = t_{n-1}, \dots, \xi_{n-k} = t_{n-k}), \end{aligned}$$

da der Prozess  $\xi_n$  von den  $k$  Werten aus der Vergangenheit abhängt. Die Bedingung auf der rechten Seite entspricht genau der Bedingung  $\Phi_{n-1} = s_{n-1}$ . Damit folgt

$$\begin{aligned} P(\Phi_n \in A | \Phi_{n-1} = s_{n-1}, \dots, \Phi_0 = s_0) &= P(\Phi_n \in A | \xi_{n-1} = t_{n-1}, \dots, t_{n-k}) \\ &= P(\Phi_n \in A | \Phi_{n-1} = s_{n-1}). \end{aligned}$$

## 1 Markov-Ketten

Die bedingte Wahrscheinlichkeit ist natürlich nur dann positiv, wenn die vorgegebenen Komponenten der  $\Phi_k$  auf der rechten Seite konsistent gewählt wurden. Dies wurde stillschweigend vorausgesetzt.  $\square$

Um verschiedene Startverteilungen eines Prozesses berücksichtigen zu können, geht man oft von einer Markov-Kette in einem Standardmodell aus. Die Übergangswahrscheinlichkeit ist dann gegeben durch

$$P_\lambda(\Phi_0 \in A_0, \dots, \Phi_n \in A_n) = \int_{A_0} \dots \int_{A_n} P(s_{n-1}, ds_n) \dots P(s_0, ds_1) \lambda(ds_0),$$

wobei  $P_\lambda$  das eindeutige Maß  $\lambda \otimes \mathbb{P}^\infty$  bezeichne. Startet der Prozess P-f.s. in einem Zustand  $s \in S$  schreibt man auch  $P_s$  anstelle von  $P_{\delta_s}$ . Analoges gilt für den Erwartungswert.

Um die eingangs erwähnten Fragestellungen zu untersuchen, formalisieren wir als nächstes die Begriffe Besuchszeit, Rückkehrzeit und Erreichbarkeit.

**Definition 1.5.** Die Anzahl  $\eta_A$  der Besuche einer Menge  $A \in \mathcal{B}(S)$  sei definiert als

$$\eta_A := \sum_{n=1}^{\infty} \mathbf{1}_{\{\Phi_n \in A\}} \quad (1.4)$$

Die erste Rückkehrzeit  $\tau_A$  und die erste Eintrittszeit  $\sigma_A$  in eine Menge  $A \in \mathcal{B}(S)$  seien definiert als

$$\tau_A := \min \{n \geq 1, \Phi_n \in A\} \quad (1.5)$$

$$\sigma_A := \min \{n \geq 0, \Phi_n \in A\} \quad (1.6)$$

Die erwartete Anzahl der Besuche einer Menge sei gegeben durch

$$U(s, A) := E_s[\eta_A] \quad (1.7)$$

Die Eintrittswahrscheinlichkeit in eine Menge  $A$  sei definiert durch

$$L(s, A) := P_s(\sigma_A < \infty) \quad (1.8)$$

**Bemerkung 1.6.** Die Rückkehrzeit  $\tau_A$  und die Besuchszeit  $\sigma_A$  sind Stoppzeiten, das heißt  $\{\tau_A = n\} \in \mathcal{F}_n$  bzw.  $\{\sigma_A = n\} \in \mathcal{F}_n$ . Hierbei sei  $(\mathcal{F}_n)_{n \geq 0}$  eine aufsteigende Folge von  $\sigma$ -Algebren in  $\mathcal{B}(S)$ , genannt *Filtration*. Stoppzeiten werden meist durch den Eintritt in eine Menge definiert, anschaulich bedeutet dann die Eigenschaft einer Stoppzeit, dass man zu jedem Zeitpunkt  $n$  entscheiden kann, ob die Menge bereits erreicht wurde.

## 1.2 Irreduzibilität und $\varphi$ -Irreduzibilität

Die Wahrscheinlichkeit  $L(s, A)$ , dass sich der Prozess nach endlich vielen Schritten in der Menge  $A$  aufhält, bildet den Ausgangspunkt für den Begriff der Erreichbarkeit einer Menge. Betrachten wir zunächst einen Prozess auf einem abzählbaren Zustandsraum. In diesem Fall ist von Interesse, ob ein bestimmter Zustand besucht wird.

## 1 Markov-Ketten

**Definition 1.7.** Ein Zustand  $t \in S$  ist erreichbar von  $s \in S$ , falls

$$L(s, t) = P_s(\tau_t < \infty) > 0.$$

Wir schreiben  $s \rightarrow t$ . Gilt für zwei Zustände  $s \rightarrow t$  und  $t \rightarrow s$ , dann sagt man, dass die Zustände  $s$  und  $t$  miteinander kommunizieren. Dafür schreiben wir  $s \leftrightarrow t$ .

**Satz 1.8.** Die Relation " $\leftrightarrow$ " ist eine Äquivalenzrelation.

Beweis: (vgl. [12], Proposition 4.1.1)

Die Reflexivität und die Symmetrie sind trivial. Es gelten  $s \rightarrow t$  und  $t \rightarrow u$ , dann existieren  $n, m$  mit  $P^n(s, t) > 0$  und  $P^m(t, u) > 0$ . Aus der Chapman-Kolmogorov-Gleichung erhält man

$$P^{n+m}(s, u) \geq P^n(s, t)P^m(t, u) > 0.$$

Also gilt  $s \rightarrow u$ , analog erhält man  $u \rightarrow s$ . Damit ist die Transitivität gezeigt. □

Mit der Äquivalenzrelation " $\leftrightarrow$ ", lässt sich die Äquivalenzklasse  $C(s)$  zu einem Zustand  $s \in S$  definieren:

$$C(s) = \{t \in S, s \leftrightarrow t\}.$$

Die Äquivalenzklasse beinhaltet alle Zustände, die mit  $s$  kommunizieren. Das heißt insbesondere, dass jeder Zustand  $t \in C(s)$  von  $s$  erreichbar ist und umgekehrt  $s$  von jedem Zustand aus  $C(s)$  erreicht werden kann.

**Definition 1.9.** Falls  $C(s) = S$ , dann heißt  $S$  oder die Markov-Kette  $\Phi$  irreduzibel.

Gilt  $P(t, C(s)) = 1, \forall t \in C(s)$ , dann heißt  $C(s)$  absorbierend.

Irreduzibilität bedeutet, dass alle Zustände in  $S$  miteinander kommunizieren. Der Zustandsraum  $S$  ist demnach nicht in mehrere Äquivalenzklassen bezüglich " $\leftrightarrow$ " zerlegbar. Ist eine Markov-Kette nicht irreduzibel, heißt dies nicht, dass sich der Prozess immer in einer Äquivalenzklasse  $C(s)$  aufhält. Es könnte nämlich ein Zustand  $t \in C(s)^c$  existieren, welcher von einem Zustand aus  $C(s)$  erreichbar ist. Hat der Prozess einmal  $t$  erreicht, kehrt er nie wieder zu der Äquivalenzklasse  $C(s)$  zurück. Die Absorptionseigenschaft gibt die wichtige Zusatzinformation, dass sich der Prozess P-f.s. in einer Äquivalenzklasse aufhält.

Das Problem bei der Übertragung der Irreduzibilität auf den allgemeinen Zustandsraum besteht darin, dass nicht klar ist wie in diesem Fall die Äquivalenzrelation " $\leftrightarrow$ " definiert werden muss. Wählt man das Lebesgue-Maß auf  $\mathbb{R}$  und gilt  $L(x, A) > 0, A \in \mathcal{B}(\mathbb{R})$ , so kehrt der Prozess nicht mit positiver Wahrscheinlichkeit zu  $x$  zurück. Aus diesem Grund führen wir die  $\varphi$ -Irreduzibilität ein.

**Definition 1.10.** Eine Markov-Kette  $\Phi$  ist  $\varphi$ -irreduzibel, wenn ein Maß  $\varphi$  auf  $\mathcal{B}(S)$  existiert, so dass

$$\varphi(A) > 0 \Rightarrow L(s, A) > 0, \forall s \in S.$$

## 1 Markov-Ketten

Ist eine Markov-Kette  $\varphi$ -irreduzibel, kann demnach jede Menge  $A \in \mathcal{B}(S)$  mit positivem Maß  $\varphi(A)$  von jedem Zustand  $s \in S$  aus erreicht werden. Dies ist offensichtlich eine Analogie zur Irreduzibilität im abzählbaren Fall, bei dem jeder Zustand  $t \in S$  von jedem Zustand  $s \in S$  erreichbar ist. Der Nachteil liegt darin, dass das Maß  $\varphi$  willkürlich gewählt werden kann. Weiß man zum Beispiel, dass eine Markov-Kette auf einem abzählbaren Raum  $S$  der Zustand  $t \in S$  von jedem  $s \in S$  erreichbar ist, dann ist die Markov-Kette bereits  $\delta_t$ -irreduzibel. In diesem Fall ist die  $\varphi$ -Irreduzibilität im Gegensatz zur gewöhnlichen Irreduzibilität eine sehr schwache Aussage. Ziel ist es ein solches Maß  $\varphi$  zu wählen, so dass die  $\varphi$ -Irreduzibilität im abzählbaren Fall mit der gewöhnlichen Irreduzibilität übereinstimmt. Es hat sich herausgestellt, dass ein "maximales", irreduzibles Maß die Forderung erfüllt.

**Satz 1.11.** *Sei  $\Phi$   $\varphi$ -irreduzibel für ein Maß  $\varphi$ . Dann existiert ein Wahrscheinlichkeitsmaß  $\psi$  auf  $\mathcal{B}(S)$ , so dass*

- i)  $\Phi$  ist  $\psi$ -irreduzibel.*
- ii) für jedes Maß  $\varphi'$  ist  $\Phi$  genau dann  $\varphi'$ -irreduzibel, wenn  $\varphi'$  absolutstetig zu  $\psi$  ist.*
- iii) wenn  $\psi(A) = 0$ , dann folgt  $\psi(\{t, L(t, A) > 0\}) = 0$ .*
- iv) Das Wahrscheinlichkeitsmaß  $\psi$  ist äquivalent zu*

$$\psi'(A) := \int_S \varphi'(ds) K_{a_{\frac{1}{2}}}(s, A)$$

*für jedes endliche irreduzible Maß  $\varphi'$ . Hierbei ist  $K_{a_{\frac{1}{2}}}(s, A)$  ein spezieller stochastischer Kern, gegeben durch*

$$K_{a_{\frac{1}{2}}}(s, A) = \sum_{n=0}^{\infty} P^n(s, A) 2^{-(n+1)}.$$

*Erfüllt ein Maß  $\psi$  diese Eigenschaften, nennen wir  $\psi$  maximal.*

Beweis: (vgl. [12], Proposition 4.2.2) □

Eine Richtung in ii) ist immer erfüllt. Ist  $\varphi'$  absolutstetig zu  $\psi$ , folgt aus  $\varphi'(A) > 0$ , dass  $\psi(A) > 0$ . Da  $\Phi$   $\psi$ -irreduzibel ist folgt daraus wiederum  $L(s, A) > 0$ . Also ist  $\Phi$  auch  $\varphi'$ -irreduzibel.

**Definition 1.12.** *i) Eine Markov-Kette  $\Phi$  heißt  $\psi$ -irreduzibel, wenn sie für ein  $\varphi$   $\varphi$ -irreduzibel ist und  $\psi$  die Maximalitätseigenschaften erfüllt.*

- ii) Wir schreiben  $\mathcal{B}^+(S) = \{A \in \mathcal{B}(S), \psi(A) > 0\}$ .*
- iii) Eine Menge  $A \in \mathcal{B}(S)$  heißt vollständig, falls  $\psi(A^c) = 0$ .*
- iv) Eine Menge  $A \in \mathcal{B}(S)$  heißt absorbierend, falls  $P(s, A) = 1, \forall s \in A$ .*

$\mathcal{B}^+(S)$  enthält also alle Mengen, welche die Markov-Kette im Sinne der  $\psi$ -Irreduzibilität besuchen kann. Ist eine Menge  $A$  vollständig und absorbierend, hält sich der Prozess P-f.s. in dieser Menge auf.

### 1.3 Periodizität

Während bis jetzt betrachtet wurde, ob eine Menge besucht wird, gibt die Periodizität Information über die Zeitpunkte, zu denen eine Menge erreicht werden kann. Betrachtet man das Beispiel einer Irrfahrt auf  $\mathbb{Z}$ , mit der Übergangswahrscheinlichkeit  $P(\Phi_n = s + 1 | \Phi_{n-1} = s) = p$  und  $P(\Phi_n = s - 1 | \Phi_{n-1} = s) = 1 - p$ , dann ist die Wahrscheinlichkeit  $P(\Phi_n = s | \Phi_{n-1} = s) = 0$ . Der Abstand der möglichen Zeitpunkte wieder nach  $s$  zurückzukehren beträgt also mindestens 2. Man sagt die Markov-Kette besitzt die Periode 2.

**Definition 1.13.** Die Periode  $d$  eines Zustands  $s \in S$  ist definiert als

$$d(s) := \text{ggT} \{n \geq 1, P^n(s, s) > 0\}$$

Die Periode eines Zustands  $s$  ist eine Solidaritätseigenschaft, d.h. jeder Zustand, der mit  $s$  kommuniziert, hat dieselbe Periode  $d$  (vgl. [12], Proposition 5.4.1). Ist die Markov-Kette irreduzibel, besitzt jeder Zustand in  $S$  die gleiche Periode.

**Definition 1.14.** Eine irreduzible Markov-Kette  $\Phi$  auf einem abzählbaren Zustandsraum heißt aperiodisch, wenn  $d(s) = 1, \forall s \in S$ .

Bei der Übertragung der Periodizität auf den allgemeinen Zustandsraum liegt das Problem wieder darin, dass im Allgemeinen ein einzelner Zustand nicht mit positiver Wahrscheinlichkeit besucht wird. Wir führen deshalb eine wichtige Klasse von Mengen ein.

**Definition 1.15.** Eine Menge  $C \in \mathcal{B}(S)$  heißt  $\nu_m$ -small, wenn ein  $m > 0$  und ein nichttriviales Maß  $\nu_m$  auf  $\mathcal{B}(S)$  existieren, so dass für jedes  $s \in C, B \in \mathcal{B}(S)$ ,

$$P^m(s, B) \geq \nu_m(B).$$

Wenn diese Bedingung erfüllt ist, sagen wir  $C$  ist  $\nu_m$ -small. Oft wird eine Menge einfach nur small genannt, wenn ein solches Maß existiert.

Die Existenz einer small Menge ist durch den folgenden Satz gesichert.

**Satz 1.16.** Ist  $\Phi$   $\psi$ -irreduzibel, dann existiert für jede Menge  $A \in \mathcal{B}^+(S)$  ein  $m \geq 1$  und eine  $\nu_m$ -small Menge  $C \subset A$ , so dass  $C \in \mathcal{B}^+(S)$  und  $\nu_m(C) > 0$ .

Beweis: (vgl. [12], Theorem 5.2.2) □

Small Mengen verhalten sich ähnlich wie Zustände im abzählbaren Fall. Aus diesem Grund lässt sich mit Hilfe von small Mengen die Periode definieren. Sei  $C$  eine  $\nu_m$ -small Menge, so dass

$$P^m(s, \cdot) \geq \nu_m(\cdot) \text{ für } s \in C.$$

Ohne Einschränkung können wir annehmen, dass  $\nu_m(C) > 0$  (vgl. [12], Proposition 5.2.4). Dann ist die Wahrscheinlichkeit zu  $C$  in  $m$  Schritten zurückzukehren größer Null. Sei

$$E_C := \{n \geq 1, C \text{ ist } \nu_n\text{-small, mit } \nu_n = \delta_n \nu_m, \text{ für ein } \delta_n > 0\}$$

## 1 Markov-Ketten

die Menge der Zeitpunkte, für welche die Menge  $C$  small ist, bezüglich eines Maßes proportional zu  $\nu_m$ . Dann ist  $E_C$  insbesondere die Menge der Zeitpunkte, für welche die Wahrscheinlichkeit zur Menge  $C$  zurückzukehren größer Null ist. Die Periode  $d$  einer Markov-Kette  $\Phi$  definieren wir als

$$d(C) := \text{ggT}\{E_C\}. \quad (1.9)$$

Als nächstes führen wir petite Mengen ein. Diese sind eine Verallgemeinerung der small Mengen. Die Eigenschaften der petite Mengen kommen besonders dann zum Tragen, wenn der allgemeine Zustandsraum mit einer Zusatzstruktur ausgestattet ist. Wir nehmen deshalb an, dass die  $\sigma$ -Algebra  $\mathcal{B}(S)$  auf dem Zustandsraum  $S$  mit einer lokalkompakten, separablen, metrisierbaren Topologie auf  $S$  übereinstimmt. In diesem Fall nennen wir den Zustandsraum *topologisch*. Da der Prozess, den wir später betrachten wollen auf dem Zustandsraum  $\mathbb{R}^p$  definiert ist, macht es Sinn diesen Fall näher zu untersuchen. Bei der Definition einer petite Menge betrachten wir anstelle der Markov-Kette  $\Phi$ , die Markov-Kette  $\Phi_a$  mit dem Übergangskern

$$K_a(s, A) = \sum_{n=0}^{\infty} P^n(s, A) a(n), \quad s \in S, \quad A \in \mathcal{B}(S).$$

Hierbei ist  $a = a(n)$  eine Wahrscheinlichkeitsverteilung auf  $\mathbb{N}$ , auch sampling-Verteilung genannt.

**Definition 1.17.** *Eine Menge  $C \in \mathcal{B}(S)$  heißt  $\nu_a$ -petite, wenn die Kette  $\Phi_a$  die Grenze*

$$K_a(s, B) \geq \nu_a(B),$$

*für alle  $s \in C$  und  $B \in \mathcal{B}(S)$  erfüllt, wobei  $\nu_a$  ein nichttriviales Maß auf  $\mathcal{B}(S)$  ist. Oft nennen wir eine Menge einfach nur petite, falls ein solches Maß existiert.*

Folgende Bemerkung sieht man sofort ein.

**Bemerkung 1.18.** *Jede Teilmenge einer petite Menge ist petite. Die endliche Vereinigung  $\bigcup_{i=1}^n C_i$  von petite Mengen  $C_i$  ist petite.*

Im folgenden wollen wir untersuchen, wie sich petite Mengen zu small Mengen verhalten. Ist die Menge  $C$  small, das heißt  $P^m(s, B) \geq \nu_m(B)$ ,  $\forall s \in C, B \in \mathcal{B}(S)$  für ein  $m > 0$ , und legt man als sampling Verteilung  $\delta_m$  zugrunde, dann gilt für alle  $s \in C$  und  $B \in \mathcal{B}(S)$

$$K_{\delta_m}(s, B) = \sum_{n=0}^{\infty} P^n(s, B) \delta_m = P^m(s, B) \geq \nu_m(B) =: \nu_{\delta_m}(B).$$

Also ist  $C$  petite.

**Bemerkung 1.19.** *Dadurch, dass eine small Menge auch petite ist, erhält man eine analoge Aussage zu Satz 1.16 über die Existenz einer petite Menge.*

## 1 Markov-Ketten

Der nächste Satz zeigt, wann umgekehrt eine petite Menge small ist.

**Satz 1.20.** *Wenn  $\Phi$  eine  $\psi$ -irreduzible und aperiodische Markov-Kette ist, so ist jede petite Menge auch small.*

Beweis: (vgl. [12], Theorem 5.5.7) □

Der folgende Satz gibt an, wie sich die petite-Eigenschaft einer Menge auf eine andere Menge überträgt.

**Satz 1.21.** *Sei  $A \in \mathcal{B}(S)$   $\nu_a$ -petite für eine beliebige sampling Verteilung  $a$  und  $b = \delta_p$  das Dirac-Maß für ein  $p > 0$ . Existiert ein  $\delta > 0$ , so dass  $\inf_{s \in D} L(s, A) \geq \delta$ , dann ist  $D$   $\nu_{b*a}$ -petite, wobei  $\nu_{b*a}$  als Vielfaches von  $\nu_a$  gewählt werden kann.  $b*a$  bezeichnet hierbei die Faltung der sampling Verteilungen  $a$  und  $b$ .*

Beweis: (vgl. [12], Theorem 5.5.4 i)) □

**Bemerkung 1.22.** *Falls  $\delta > 0$  existiert, so dass  $\inf_{s \in D} L(s, A) \geq \delta$  erfüllt ist, so nennt man die Menge  $A$  auch gleichmäßig erreichbar von der Menge  $D$ . In diesem Sinne drückt der vorhergehende Satz aus, dass eine Menge petite ist, falls von ihr eine petite Menge gleichmäßig erreichbar ist.*

Wir wollen nun die Zusatzstruktur ausnutzen. Da die Topologie metrisierbar ist, kann man von Stetigkeit und von kompakten Mengen sprechen. Wir fordern, dass der stochastische Kern in gewisser Weise stetig ist.

**Definition 1.23.** *Eine Funktion  $h : S \rightarrow \mathbb{R}$  ist unterhalbstetig, wenn*

$$\liminf_{t \rightarrow s} h(t) \geq h(s).$$

Eine oft verwendete unterhalbstetige Funktion ist die Indikatorfunktion  $\mathbf{1}_O(s)$  für offene Mengen  $O \in \mathcal{B}(S)$ . Wir zeigen, dass die Indikatorfunktion unterhalbstetig ist. Sei  $(s_n)_n$  eine Folge, die gegen  $s$  konvergiert. Es sind zwei Fälle zu unterscheiden.

- i) Sei  $s \in O$ . Da  $O$  offen ist, existiert eine Umgebung  $U_\epsilon(s) \subset O$  und ein  $n_0$ , so dass  $s_n \in U_\epsilon(s), \forall n \geq n_0$ . Daraus folgt

$$\liminf_{s_n \rightarrow s} \mathbf{1}_O(s_n) = 1 = \mathbf{1}_O(s).$$

- ii) Sei  $s \notin O$ , dann gilt trivialerweise

$$\liminf_{s_n \rightarrow s} \mathbf{1}_O(s_n) \geq 0 = \mathbf{1}_O(s).$$

Damit ist die Unterhalbstetigkeit von  $\mathbf{1}_O(s)$  gezeigt.

**Definition 1.24.** *i) Wenn  $P(\cdot, O)$  eine unterhalbstetige Funktion für jede offene Menge  $O \in \mathcal{B}(S)$  ist, dann heißt  $P$  eine (schwache) Feller chain.*

- ii) Wenn  $a$  eine sampling Verteilung und  $T$  ein substochastischer Übergangskern ist, der

$$K_a(s, A) \geq T(s, A), \quad s \in S, \quad A \in \mathcal{B}(S)$$

erfüllt, wobei  $T(\cdot, A)$  eine unterhalbstetige Funktion für jedes  $A \in \mathcal{B}(S)$  ist, dann heißt  $T$  eine stetige Komponente von  $K_a$ .

- iii) Wenn  $\Phi$  eine Markov-Kette ist für die eine sampling Verteilung  $a$  existiert, so dass  $K_a$  eine stetige Komponente  $T$  besitzt mit  $T(s, S) \geq 0$  für alle  $s$ , dann heißt  $\Phi$  T-chain.

Ein Kriterium für den Nachweis einer T-chain bildet der folgende Satz.

**Satz 1.25.** Falls für jedes  $s \in S$  eine Wahrscheinlichkeitsverteilung  $a_s$  auf  $\mathbb{N}$  existiert, so dass  $K_{a_s}$  eine stetige Komponente  $T_s$  besitzt, die nichttrivial in  $s$  ist, dann ist  $\Phi$  eine T-chain.

Beweis: (vgl. [12], Proposition 6.2.4) □

Mit Hilfe der T-chain und der Existenz kompakter Mengen lässt sich ein einfaches Kriterium für eine petite Menge formulieren.

**Satz 1.26.** Wenn jede kompakte Menge  $C \in \mathcal{B}(S)$  petite ist, so ist  $\Phi$  eine T-chain. Umgekehrt, wenn  $\Phi$  eine  $\psi$ -irreduzible T-chain ist, dann ist jede kompakte Menge petite.

Beweis: (vgl. [12], Theorem 6.0.1 ii)) □

## 1.4 Rekurrenz und Transienz

Der vorhergehende Satz über eine T-chain ist für uns das entscheidende Hilfsmittel um eine petite Menge zu bestimmen. Mit Hilfe von petite Mengen bzw. small Mengen lassen sich Aussagen über die Stabilität einer Markov-Kette treffen. Hierauf wird aber erst im nächsten Kapitel eingegangen. Den ersten Schritt in diese Richtung bildet der Begriff Rekurrenz, der in diesem Abschnitt eingeführt wird. Wir beginnen wieder mit dem Fall eines abzählbaren Zustandsraumes.

**Definition 1.27.** Sei  $\Phi$  eine Markov-Kette auf einem abzählbaren Zustandsraum. Dann heißt ein Zustand  $s \in S$  transient, falls  $E_s[\eta_s] < \infty$  und rekurrent, falls  $E_s[\eta_s] = \infty$ .

Es gilt:

$$U(s, t) = \sum_{n=1}^{\infty} P^n(s, t) = E_s \left[ \sum_{n=1}^{\infty} \mathbf{1}_{\{\Phi_n=t\}} \right] = E_s[\eta_t].$$

Rekurrenz ist genau wie die Periode eine Solidaritätseigenschaft, das heißt sie überträgt sich auf alle Zustände, die mit einem rekurrenten Zustand kommunizieren.

**Satz 1.28.** Wenn der Zustandsraum  $S$  abzählbar ist und die Markov-Kette  $\Phi$  irreduzibel, dann ist entweder  $U(s, t) = \infty, \forall s, t \in S$  oder  $U(s, t) < \infty, \forall s, t \in S$ .



## 1 Markov-Ketten

Beweis: (vgl. [12], Proposition 8.1.1)

Es gelte  $U(s, t) = \sum_{n=1}^{\infty} P^n(s, t) = \infty$ , für ein  $s$  und ein  $t$  in  $S$ . Aufgrund der Irreduzibilität existieren  $s', t' \in S$  und  $k$  und  $l$ , so dass  $P^k(s', s) > 0$  und  $P^l(t, t') > 0$ . Dann folgt

$$\sum_{n=1}^{\infty} P^{k+l+n}(s', t') > P^k(s', s) \left( \sum_{n=1}^{\infty} P^n(s, t) \right) P^l(t, t') = \infty.$$

Existieren dagegen Zustände mit  $U(s, t) < \infty$ , so muss dies wiederum für alle Zustände in  $S$  gelten, da man sonst einen Widerspruch zur ersten Aussage hätte.  $\square$

Die Solidaritätseigenschaft macht es möglich eine irreduzible Markov-Kette als rekurrent bzw. transient zu klassifizieren.

**Definition 1.29.** Eine Markov-Kette  $\Phi$  heißt transient, wenn jeder Zustand transient ist und rekurrent, wenn jeder Zustand rekurrent ist.

Als nächstes stellen wir die Verbindung zu der Wahrscheinlichkeit  $L(s, s) = P_s(\tau_s < \infty)$  her.

**Satz 1.30.** Für alle  $s \in S$  gilt  $U(s, s) = \infty$  genau dann, wenn  $L(s, s) = 1$ .

Beweis: (vgl. [12], Proposition 8.1.3)  $\square$

Durch die Wahrscheinlichkeit  $L(s, s)$  in endlicher Zeit zum Zustand  $s$  zurückzukehren, haben wir also ein äquivalentes Rekurrenz Kriterium gefunden. Betrachten wir nun den allgemeinen Zustandsraum.

**Definition 1.31.** Eine Menge  $A \in \mathcal{B}(S)$  heißt gleichmäßig transient, wenn ein  $M < \infty$  existiert, so dass  $U(s, A) = E_s[\eta_A] \leq M, \forall s \in A$  und rekurrent, wenn  $E_s[\eta_A] = \infty, \forall s \in A$ . Weiterhin heißt eine Menge transient, wenn sie von einer abzählbaren Kollektion von gleichmäßig transienten Mengen überdeckt werden kann.

Analog zu der Solidaritätseigenschaft im abzählbaren Fall definieren wir eine Markov-Kette als rekurrent bzw. transient.

**Definition 1.32.** Eine  $\psi$ -irreduzible Markov-Kette  $\Phi$  heißt rekurrent, falls  $U(s, A) \equiv \infty$  für alle  $s \in S$  und alle  $A \in \mathcal{B}^+(S)$ . Eine  $\psi$ -irreduzible Markov-Kette  $\Phi$  heißt transient, falls  $S$  transient ist.

Dass diese Definition sinnvoll ist, liefert die Solidaritätseigenschaft im allgemeinen Fall.

**Satz 1.33.** Sei  $\Phi$  eine  $\psi$ -irreduzible Markov-Kette, dann ist  $\Phi$  entweder rekurrent oder transient.

Beweis: (vgl. [12], Theorem 8.2.5)  $\square$

Im transienten Fall lässt sich für eine petite Menge noch eine wichtige Eigenschaft formulieren.

**Satz 1.34.** Sei  $\Phi$   $\psi$ -irreduzibel und transient. Dann ist jede petite Menge gleichmäßig transient.

Beweis: (vgl. [12], Theorem 8.3.5)  $\square$

## 1.5 Harris-Rekurrenz

Obwohl die Dichotomie aus Satz 1.33, welche von vorneherein nicht auf der Hand liegt, die Analyse von Markov-Ketten wesentlich vereinfacht, wäre es wünschenswert, wenn die Rekurrenz im allgemeinen Fall mit der Rekurrenz, formuliert durch die Eintrittswahrscheinlichkeit  $L(s, \cdot)$ , übereinstimmt. Oft wird nämlich diese Formulierung für die Definition von Rekurrenz vorangestellt. Im allgemeinen gilt die Äquivalenz aus Satz 1.30 aber nicht. Eine Verbindung stellt dennoch der folgende Satz bereit.

**Satz 1.35.** *i)  $\Phi$  ist rekurrent, wenn eine petite Menge  $C \in \mathcal{B}(S)$  existiert, so dass  $L(s, C) \equiv 1$  für alle  $s \in C$ .*

*ii)  $\Phi$  ist genau dann transient, wenn zwei petite Menge  $C, D \in \mathcal{B}^+(S)$  existieren mit  $L(s, C) < 1$  für alle  $s \in D$ .*

Beweis: (vgl. [12], Theorem 8.3.6) □

Um über einen Rekurrenzbegriff zu verfügen, der auf der Eintrittswahrscheinlichkeit  $L(s, \cdot)$  basiert, führen wir die Harris-Rekurrenz ein. Hierzu betrachtet man die Wahrscheinlichkeit  $Q(s, A)$ , dass eine Menge  $A$  P-f.s. unendlich oft besucht wird. Wie man leicht einsieht gilt

$$Q(s, A) = P_s(\Phi \in A \text{ unendlich oft}) \leq L(s, A).$$

**Definition 1.36.** *Eine Menge  $A$  heißt Harris-rekurrent, wenn*

$$Q(s, A) = P(\eta_A = \infty) = 1, \forall s \in A.$$

*Eine Markov-Kette  $\Phi$  heißt Harris(-rekurrent), wenn sie  $\psi$ -irreduzibel und jede Menge aus  $\mathcal{B}^+(S)$  Harris-rekurrent ist.*

Die gewünschte Äquivalenz erhält man, wenn

$$L(s, A) = 1 \Rightarrow Q(s, A) = 1$$

gilt.

**Satz 1.37.** *i) Sei  $A \in \mathcal{B}(S)$  eine Menge mit  $L(s, A) \equiv 1$ ,  $s \in A$ . Dann gilt  $Q(s, A) = L(s, A)$ ,  $\forall s \in S$ ,  $A$  ist insbesondere Harris-rekurrent.*

*ii) Sei  $\Phi$   $\psi$ -irreduzibel. Wenn eine petite Menge  $C \in \mathcal{B}(S)$  existiert, so dass  $L(s, C) \equiv 1$ ,  $s \in S$ , dann ist  $\Phi$  Harris-rekurrent.*

Beweis: (vgl. [12], Proposition 9.1.1 und Proposition 9.1.7) □

Für spätere Stabilitätsbetrachtungen wird oft die Harris-Rekurrenz gefordert. Um diese nachzuweisen ist die eben beschriebene Äquivalenz entscheidend.

## 2 Stabilität und geometrische Ergodizität

Es gibt verschiedene Ansätze um die Stabilität von dynamischen Systemen zu beschreiben. Dies gilt sowohl für den deterministischen als auch für den stochastischen Fall. Oft betrachtet man zu einem stochastischen dynamischen System das deterministische Analogon, um eine Aussage über die Stabilität des Systems zu treffen. Wir stellen zunächst ein einfaches Modell vor, von dem das Verhalten des deterministischen Analogons bekannt ist, um den Stabilitätsbegriff für den stochastischen Fall zu motivieren. Unter Stabilität stelle man sich zunächst vor, dass sich der Prozess in einer beschränkten Menge aufhält oder sogar gegen einen Punkt konvergiert.

Für stochastische Prozesse werden drei verschiedene Begriffe eine wichtige Rolle spielen: Rekurrenz, Ergodizität und geometrische Ergodizität. Rekurrenz bedeutet, dass es eine Menge gibt, die der Prozess unendlich oft besucht. Ergodizität meint, dass die Übergangswahrscheinlichkeit  $P^n(x, \cdot)$  gegen eine stationäre Verteilung  $\pi$  konvergiert. Geometrische Ergodizität sagt aus, dass diese Konvergenz mit einer exponentiellen Rate stattfindet. Von dieser Stelle an werden meist die Buchstaben  $x$  und  $y$  für die Zustände aus dem Zustandsraum  $S$  verwendet, da wir in Kapitel den Zustandsraum  $\mathbb{R}^p$  und die Ergebnisse so direkt übertragen werden können.

### 2.1 Rekurrenz und der Drift $\Delta V$ einer Markov-Kette

Betrachten wir als erstes ein einfaches Beispiel und zwar ein autoregressives Modell, gegeben durch

$$X_{t+1} = \alpha X_t + Z_{t+1},$$

wobei der Störterm  $Z_t$  eine Folge von unabhängig identisch verteilten Zufallsvariablen sei mit Erwartungswert  $\mu$  und Varianz  $\sigma^2$ . Ohne Einschränkung sei  $\mu = 0$  gewählt. Um Aussagen über die Stabilität zu treffen, ist es oft hilfreich das zugehörige deterministische Modell, in diesem Fall  $x_{t+1} = \alpha x_t$ , zu betrachten. Wählt man  $|\alpha| < 1$  sieht man schnell ein, dass der deterministische Prozess gegen Null konvergiert. Bei dem stochastischen Modell ist dies nicht gegeben. Für große  $|X_t|$  überwiegt die Tendenz des Prozesses gegen den Ursprung zu streben den Einfluss des Störterms. Ist dagegen  $|X_t|$  schon nahe bei Null, wird der Prozess hauptsächlich durch den Fehlerterm bestimmt, der den Prozess immer wieder von Null wegbewegt. Konvergenz gegen Null ist nicht zu erreichen. Trotzdem ist der Prozess in dem Sinne stabil, dass er stationär wird (vgl. [13], Kapitel 4.1). Wählt man  $|\alpha| > 1$  ist klar, dass der Prozess in beiden Fällen als instabil zu bewerten ist. Bei stochastischen Modellen kann man demnach nicht davon ausgehen, dass sie gegen einen

## 2 Stabilität und geometrische Ergodizität

bestimmten Zustand konvergieren oder sich immer in einer bestimmten Menge aufhalten. Ausnahme ist natürlich die Existenz eines absorbierenden Zustands. Eine sinnvolle Forderung wäre zum Beispiel die Rekurrenz. Dann wäre ein stochastischer Prozess stabil, wenn eine Menge  $C \subset S$  existiert, so dass der Prozess unendlich oft zu  $C$  zurückkehrt. Insbesondere wäre der Prozess dann nicht instabil, da er demnach nicht gegen unendlich streben kann.

Ein wichtiges Kriterium um Rekurrenz zu zeigen, ist die Drift-Bedingung, welche in ähnlicher Form eine zentrale Rolle in dieser Arbeit spielt.

**Definition 2.1.** *Der Drift einer Markov-Kette ist definiert als*

$$\Delta V(x) := \int_S P(x, dy)V(y) - V(x) = E_x \left[ V(\Phi_1) - V(\Phi_0) \right], \quad x \in S,$$

wenn  $\Phi_0 = x$  der Startwert der Kette ist und  $V$  eine nichtnegative messbare Funktion.

Der Drift entspricht demnach der erwarteten Abweichung des Prozesses  $V(\Phi_1)$  zum Startwert  $V(\Phi_0)$ . Der nächste Satz benutzt den Drift um ein Kriterium für die Rekurrenz einer Markov-Kette zu formulieren.

**Satz 2.2.** *Sei  $\Phi$  eine  $\psi$ -irreduzible Markov-Kette. Wir nehmen an es existieren eine petite Menge  $C$  und eine Funktion  $V$ , die unbeschränkt außerhalb von petite Mengen ist, das heißt die Menge  $C_V(n) = \{y, V(y) \leq n\}$  ist für alle  $n$  petite. Sei weiterhin die Drift-Bedingung*

$$\Delta V(x) = \int_S P(x, dy)V(y) - V(x) \leq 0, \quad \text{für } x \in C^c \tag{2.1}$$

erfüllt, dann gilt  $L(x, C) \equiv 1$  und  $\Phi$  ist (Harris-)rekurrent.

Beweis: (vgl. [12], Theorem 8.4.3 und Theorem 9.1.8)

Zur Vereinfachung sei  $V$  schon dann beschränkt außerhalb von petite Mengen, wenn eine endliche Folge von petite Mengen  $C_j$  existiert, so dass  $C_V \subset \bigcup C_j$ . Da  $\Phi$  irreduzibel, ist  $C_V$  als Teilmenge endlicher Vereinigung von petite Mengen selbst petite, siehe Bemerkung 1.18. Ohne Einschränkung kann dann  $C \in \mathcal{B}(S)$  mit  $\psi(C) > 0$  angenommen werden. Denn gilt dies nicht, gehe über zu  $C \cup C_V(n)$  mit  $n$  so groß, dass  $\psi(C \cup C_V(n)) > 0$ .

$L(x, C) = 1$  zeigen wir durch einen Widerspruch. Sei  $\Phi$  transient, dann existiert nach Satz 1.35 ein  $x^* \in C^c$  mit  $L(x^*, C) < 1$ . Nach Voraussetzung ist die Menge  $C_V(n)$  petite, deswegen ist  $C_V(n)$  nach Satz 1.34 ii) gleichmäßig transient für alle  $n$ .

Man wähle  $M$  so groß, dass

$$M > \frac{V(x^*)}{1 - L(x^*, C)}. \tag{2.2}$$

## 2 Stabilität und geometrische Ergodizität

Definiert man außerdem die Kette  $\hat{P}$  durch  $\hat{P}(x, A) := P(x, A)$ ,  $x \in C^c$ , und  $\hat{P}(x, x) := 1$  für  $x \in C$ , dann folgt mit der Drift-Bedingung (2.1)

$$\begin{aligned}
 V(x) &\geq \int \hat{P}^n(x, dy) V(y) \\
 &\geq \int_{C^c \cap C_V(M)^c} \hat{P}^n(x, dy) V(y) \\
 &\geq M \cdot \int_{C^c \cap C_V(M)^c} \hat{P}^n(x, dy) \\
 &= M \cdot (1 - \hat{P}^n(x, C \cup C_V(M))). \tag{2.3}
 \end{aligned}$$

Da die Menge  $C$  absorbierend für  $\hat{P}$  ist und  $\hat{P}$  außerhalb von  $C$  mit  $P$  übereinstimmt, gilt

$$\hat{P}^n(x, C) = P_x(\tau_C \leq n) \nearrow L(x, C), \quad x \in C^c.$$

Die Menge  $C_V(M)$  ist gleichmäßig transient, deshalb existiert ein  $n$ , so dass

$$E_x \eta_{C_V(M)} \leq K.$$

Zusammen mit der trivialen Gleichung  $\hat{P}^n(x, A) \leq P^n(x, A)$  für  $x \in C^c$  erhält man

$$\hat{P}^n(x, C_V(M) \cap C^c) \leq P^n(x, C_V(M) \cap C^c) \rightarrow 0, \quad x \in C^c.$$

Daraus ergibt sich wiederum

$$1 - \hat{P}^n(x^*, C_V(M) \cup C) \rightarrow 1 - L(x^*, C) \quad \text{für } n \rightarrow \infty$$

Mit (2.3) erhält man

$$V(x) \geq M \cdot (1 - L(x^*, C))$$

Damit ergibt sich ein Widerspruch zur Wahl von  $M$ . Demnach ist  $L(x, C) = 1$ ,  $\forall x \in S$ . Mit Satz 1.35 folgt hieraus die Rekurrenz von  $\Phi$ . Satz 1.37 liefert dagegen die Harris-Rekurrenz.  $\square$

## 2.2 Ergodizität

Eine stärkere Stabilitätsbedingung ist die Konvergenz der Übergangswahrscheinlichkeit gegen eine Wahrscheinlichkeitsverteilung  $\pi$ . Zunächst überprüfen wir, welche Eigenschaft eine solche Grenzwahrscheinlichkeit notwendigerweise besitzt. Hierzu definieren wir stationäre bzw. invariante Maße.

**Definition 2.3.** *Ein stationäres oder invariantes Maß ist ein  $\sigma$ -endliches Maß auf  $\mathcal{B}(S)$  mit der Eigenschaft*

$$\pi(A) = \int P(x, A) \pi(dx). \tag{2.4}$$

## 2 Stabilität und geometrische Ergodizität

Es stellt sich heraus, dass ein solches Maß existiert, wenn die Kette rekurrent ist.

**Satz 2.4.** *Wenn die Markov-Kette  $\Phi$  rekurrent ist, dann besitzt sie ein eindeutiges (bis auf Vielfache), invariantes Maß  $\pi$ , welches*

$$\pi(B) = \int_A E_\omega \left[ \sum_{n=1}^{\tau_A} \mathbf{1}_{\{\Phi_n \in B\}} \right] \pi(d\omega), \quad B \in \mathcal{B}(S),$$

*erfüllt. Das invariante Maß  $\pi$  ist endlich, wenn eine petite Menge  $C$  existiert, so dass*

$$\sup_{x \in C} E_x \tau_C < \infty. \quad (2.5)$$

Beweis: (vgl. [12], Theorem 10.0.1 bzw. Theorem 10.4.9 und 10.4.10) □

Erfüllt die Markov-Kette also die zusätzliche Bedingung (2.5), ist das invariante Maß  $\pi$  endlich und lässt sich somit zu einem Wahrscheinlichkeitsmaß normieren.

**Definition 2.5.** *Sei  $\Phi$  eine  $\psi$ -irreduzible Markov-Kette mit einem invarianten Wahrscheinlichkeitsmaß  $\pi$ , dann nennen wir  $\Phi$  eine positive Markov-Kette. Andernfalls heißt die Kette null.*

Als erste Verbindung eines stationären Wahrscheinlichkeitsmaßes zu unseren Stabilitätsüberlegungen, können wir festhalten, dass aus der Existenz eines solchen Maßes bereits Rekurrenz folgt.

**Satz 2.6.** *Wenn die Kette  $\Phi$  positiv ist, dann ist sie rekurrent.*

Beweis: (vgl. [12], Theorem 10.1.1)

Wir zeigen die Aussage durch Widerspruch. Sei  $\Phi$  transient.  $(A_j)_j$  sei eine Folge gleichmäßig transienter Mengen mit  $\bigcup A_j = S$  und Konstanten  $M_j$ , so dass mit  $U(x, A_j) \leq M_j$  (vgl. [12], Beweis von Satz 8.3.4). Nutzt man die Eigenschaft des invarianten Wahrscheinlichkeitsmaßes  $\pi$  (vgl. (2.4)) aus, erhält man

$$\begin{aligned} \pi(A) &= \int_S P(w, A) \pi(dw) \\ &= \int_S \int_S P(w, A) P(x, dw) \pi(dx) \\ &= \int_S P^2(x, A) \pi(dx) = \dots = \int_S P^n(x, A) \pi(dx) \\ &= P_\pi(\Phi_n \in A). \end{aligned}$$

Wendet man dies auf die Mengen  $A_j$  an, ergibt sich

$$k\pi(A_j) = \sum_{n=1}^k \int \pi(dw) P^n(w, A_j) \leq M_j, \quad \forall j, k.$$

## 2 Stabilität und geometrische Ergodizität

Für  $k \rightarrow \infty$  bleibt die linke Seite endlich, deshalb folgt  $\pi \equiv 0$ . Dies ist ein Widerspruch dazu, dass  $\pi$  ein Wahrscheinlichkeitsmaß ist.  $\square$

Bis jetzt wurde die Rekurrenz in Beziehung zu einem invarianten Wahrscheinlichkeitsmaß gesetzt. Nun soll gezeigt werden, in welchem Zusammenhang ein invariantes Maß zu der Konvergenz der Übergangswahrscheinlichkeit  $P^n(x, \cdot)$  steht. Die folgende Gleichung gibt hierfür einen Hinweis. Wenn ein Wahrscheinlichkeitsmaß  $\gamma_\mu$  existiert, so dass  $P_\mu(\Phi_n \in A) \rightarrow \gamma_\mu(A)$ , dann gilt

$$\begin{aligned} \gamma_\mu(A) &= \lim_{n \rightarrow \infty} \int_S P^n(x, A) \mu(dx) \\ &= \lim_{n \rightarrow \infty} \int_S \int_S P(\omega, A) P^{n-1}(x, d\omega) \mu(dx) \\ &= \int_S P(\omega, A) \gamma_\mu(d\omega). \end{aligned}$$

Das heißt, wenn die Übergangswahrscheinlichkeit  $P_\mu(\Phi_n \in A)$  gegen ein Wahrscheinlichkeitsmaß  $\gamma_\mu(A)$  konvergiert, dann ist  $\gamma_\mu(A)$  ein invariantes Maß. Falls das zu  $\Phi$  gehörige invariante Maß  $\pi$  eindeutig ist, konvergiert  $P_\mu(\Phi_n \in A)$  gegen  $\pi$ .

Es wurde bisher nicht explizit auf die Konvergenzart eingegangen. Im folgenden wollen wir die totale Variation betrachten.

**Definition 2.7.** Die totale Variation Norm  $\|\mu\|$  eines Wahrscheinlichkeitsmaßes ist definiert als

$$\|\mu\| := \sup_{f: |f| \leq 1} |\mu(f)| = \sup_{A \in \mathcal{B}(S)} |\mu(A)| - \inf_{A \in \mathcal{B}(S)} |\mu(A)|,$$

wobei  $\mu(f) := \int_S f(x) \mu(dx)$  sei.

Zusammenfassend lässt sich das aperiodische Ergodentheorem formulieren:

**Satz 2.8.** Sei  $\Phi$  eine aperiodische Harris-rekurrente Markov-Kette mit invariantem Maß  $\pi$ . Dann ist äquivalent:

- i) Die Markov-Kette  $\Phi$  ist positiv Harris, das heißt, das Maß  $\pi$  ist endlich.
- ii) Es existiert eine  $\nu$ -small Menge  $C \in \mathcal{B}^+(S)$  und  $P^\infty(C) > 0$ , so dass

$$P^n(x, C) \rightarrow P^\infty(C).$$

- iii) Es existiert eine petite Menge  $C$ , so dass

$$\sup_{x \in C} E_x \tau_C < \infty.$$

## 2 Stabilität und geometrische Ergodizität

iv) Es existiert eine petite Menge  $C$ , ein  $b < \infty$  und eine nichtnegative Funktion  $V$ , die für ein  $x_0 \in S$  endlich ist, so dass

$$\Delta V(x) \leq -1 + b\mathbf{1}_C(x), \quad x \in S \quad (2.6)$$

Aus jeder dieser Bedingungen folgt, dass ein eindeutiges invariantes Wahrscheinlichkeitsmaß  $\pi$  existiert mit

$$\lim_{n \rightarrow \infty} \|P^n(x, A) - \pi\| = 0, \quad \text{für alle } x \in S.$$

In diesem Fall heißt die Markov-Kette  $\Phi$  ergodisch.

Beweis: (vgl. [12], Kapitel 13, insbesondere Theorem 13.3.3) □

Ist die Markov-Kette  $\Phi$  ergodisch, sieht man leicht ein, dass auch

$$\lim_{n \rightarrow \infty} E_x f(\Phi_n) = \int_S f(x) d\pi, \quad (2.7)$$

für jedes beschränkte  $f$  gilt. Wünschenswert wäre aber die Konvergenz für eine allgemeinere Funktion  $f$ . Wir betrachten deshalb die  $f$ -Norm

$$\|\nu\|_f := \sup_{g, |g| \leq f} |\nu(g)|.$$

Es stellt sich heraus (vgl. [12], Theorem 14.0.1): Falls  $\Phi$  ergodisch ist und zusätzlich  $\pi(f) = \int_S f(x) \pi(dx) < \infty$  erfüllt ist, gilt

$$\lim_{n \rightarrow \infty} \|P^n(x, \cdot) - \pi\|_f = 0, \quad \forall x \in S_V := \{y \in S, V(y) < \infty\}.$$

### 2.3 Geometrische Ergodizität

In diesem Abschnitt betrachten wir ergodische Markov-Ketten, bei denen die Konvergenz gegen die invariante Verteilung  $\pi$  mit einer exponentiellen Rate stattfindet, das heißt

$$\|P^n(x, \cdot) - \pi\|_f = o(\rho^n),$$

wobei die Rate  $\rho < 1$  unabhängig vom Startwert  $x$  gewählt werden kann. Genauer lautet die Definition:

**Definition 2.9.** Eine Markov-Kette  $\Phi$  heißt  $f$ -geometrisch ergodisch, wenn  $f \geq 1$ ,  $\Phi$  eine positive Harris-Kette ist mit  $\pi(f) < \infty$  und eine Konstante  $r_f > 1$  existiert, so dass

$$\sum r_f^n \|P^n(x, \cdot) - \pi\|_f < \infty, \quad \forall x \in S.$$

Falls  $f = 1$  heißt  $\Phi$  auch geometrisch ergodisch.



## 2 Stabilität und geometrische Ergodizität

Analog zu Satz 2.8 lässt sich ein Satz formulieren, der äquivalente Bedingungen für die  $f$ -geometrische Ergodizität zusammenfasst.

**Satz 2.10.** (*Geometrisches Ergodentheorem*) *Sei  $\Phi$  eine  $\psi$ -irreduzible, aperiodische Markov-Kette. Dann sind folgende Aussagen äquivalent:*

- i) *Die Kette  $\Phi$  ist positiv rekurrent mit invarianten Wahrscheinlichkeitsmaß  $\pi$  und es existieren eine  $\nu$ -petite Menge  $C \in \mathcal{B}^+(S)$ ,  $\rho_C < 1$  und  $M_C < \infty$  und  $P^\infty(C) > 0$ , so dass für alle  $x \in C$*

$$|P^n(x, C) - P^\infty(C)| \leq M_C \rho_C^n.$$

- ii) *Es existiert eine petite Menge  $C \in \mathcal{B}(S)$  und  $\kappa > 1$ , so dass*

$$\sup_{x \in C} E_x \kappa^{\tau_C} < \infty.$$

- iii) *Es existieren eine petite Menge  $C$ , Konstanten  $b < \infty$ ,  $\beta > 0$  und eine Funktion  $V \geq 1$ , endlich für ein  $x_0 \in S$ , so dass*

$$\Delta V(x) \leq -\beta V(x) + b \mathbf{1}_C(x), \quad x \in S \tag{2.8}$$

*Jede dieser Bedingungen impliziert, dass die Menge  $S_V = \{x, V(x) < \infty\}$  absorbierend und vollständig ist und es existieren Konstanten  $r > 1$  und  $R < \infty$ , so dass für alle  $x \in S_V$*

$$\sum_n r^n \|P^n(x, \cdot) - \pi\|_V \leq R V(x).$$

Beweis: (vgl. [12], Theorem 15.0.1)

□

Der Beweis des geometrischen Ergodentheorems wird hier nicht geführt. Es wird nur der entscheidende Satz vorgestellt, der die V-geometrische Ergodizität liefert. Die Absorbtions-Eigenschaft und die Vollständigkeit der Menge  $S_V$  erhält man leicht mit Aussage iii). Die Drift-Bedingung 2.8 impliziert die Gleichung

$$\int_S V(y) P(x, dy) \leq V(x) + b.$$

Startet der Prozess in  $x \in S_V$ , dann ist  $V(x) < \infty$  und somit  $\int_S V(y) P(x, dy) < \infty$ .  $V(y)$  ist also  $P_x$ -f.s. endlich und damit  $P(x, S_V) = 1$ ,  $x \in S_V$ . Wäre  $S_V$  nicht vollständig, also  $\psi(S_V^c) > 0$ , ist dies ein Widerspruch zur  $\psi$ -Irreduzibilität von  $\Phi$ . Erfüllt  $V$  die Drift-Bedingung für eine petite Menge  $C$ , so ist  $V$  außerdem unbeschränkt außerhalb von petite Mengen (vgl. [12], Lemma 15.2.2). Mit Satz 2.2 folgt deshalb die Harris-Rekurrenz von  $\Phi$ .

Der entscheidende Satz, aus dem V-geometrische Ergodizität folgt, ist:

## 2 Stabilität und geometrische Ergodizität

**Satz 2.11.** *Sei  $\Phi$   $\psi$ -irreduzibel und aperiodisch und es existiere eine  $f$ -Kendall petite Menge  $C \in \mathcal{B}(S)$ , d.h. es existiere ein  $\kappa = \kappa(f) > 1$  mit  $\sup_{x \in C} E_x \left[ \sum_{k=0}^{\tau_C-1} f(\Phi_k) \kappa^k \right] < \infty$ .*

*Dann ist*

$$E_x \left[ \sum_{k=0}^{\tau_C-1} f(\Phi_k) \kappa^k \right]$$

*auf der absorbierenden Menge  $S_f := \{x, f(x) < \infty\}$  endlich und für alle  $x \in S_f$  gilt*

$$\sum_n r^n \|P^n(x, \cdot) - \pi\|_f \leq R E_x \left[ \sum_{k=0}^{\tau_A-1} f(\Phi_k) \kappa^k \right].$$

Beweis: (vgl. [12], Theorem 15.4.1) □

Die Existenz einer  $V$ -Kendall petite Menge ist äquivalent zur Lösung der Drift-Bedingung (2.8) (vgl. [12], Theorem 15.2.1, Theorem 15.2.4 und Theorem 15.2.5). Die Ergodizität von  $\Phi$  und die Drift-Bedingung implizieren schließlich  $\pi(V) = \int V(x) \pi(dx) < \infty$ . Also ist  $\Phi$  in der Situation von Satz 2.10  $V$ -geometrisch ergodisch auf der absorbierenden und vollständigen Menge  $S_V$ .

Abschließend präsentieren wir eine andere Sichtweise der  $f$ -geometrischen Ergodizität, indem wir die Konvergenz bezüglich einer Operatornorm betrachten. Hierzu führen wir den  $V$ -Norm Abstand zweier Übergangskerne ein.

**Definition 2.12.** *Der  $V$ -Norm Abstand zweier Übergangskerne  $P_1, P_2$  ist definiert als*

$$\| \|P_1(x, \cdot) - P_2(x, \cdot)\|_V := \sup_{x \in S} \frac{\|P_1(x, \cdot) - P_2(x, \cdot)\|_V}{V(x)}.$$

Wir zeigen, dass die  $V$ -Norm  $\| \cdot \|_V$  eine Operatornorm ist.

**Lemma 2.13.** *Sei  $L_V^\infty$  der Vektorraum aller Funktionen  $f : S \rightarrow \mathbb{R}_+$ , mit*

$$\|f\|_V := \sup_{x \in S} \frac{|f(x)|}{V(x)} < \infty.$$

*Wenn  $\| \|P_1 - P_2\|_V$  endlich ist, dann ist  $P_1 - P_2$  ein beschränkter Operator von  $L_V^\infty$  nach  $L_V^\infty$  und  $\| \|P_1 - P_2\|_V$  ist die zugehörige Operatornorm.*

Beweis: (vgl. [12], Lemma 16.1.1)

Schreibt man die Definition von  $\| \cdot \|_V$  aus, ergibt sich

$$\begin{aligned} \| \|P_1 - P_2\|_V &= \sup_{x \in S} \frac{\sup_{g: g \leq V} |P_1(x, g) - P_2(x, g)|}{V(x)} \\ &= \sup_{g: g \leq V} \frac{\sup_{x \in S} |P_1(x, g) - P_2(x, g)|}{V(x)} \\ &= \sup_{g: g \leq V} |P_1(\cdot, g) - P_2(\cdot, g)|_V \\ &= \sup_{|g|_V \leq 1} |P_1(\cdot, g) - P_2(\cdot, g)|_V. \end{aligned}$$

## 2 Stabilität und geometrische Ergodizität

Dies ist per Definition die Operatornorm von  $P_1 - P_2$ . □

Man möchte natürlich den Abstand  $\|P^n(x, \cdot) - \pi\|_V$  betrachten, der aber erst durch die Definition  $\pi(x, A) := \pi(A)$ ,  $A \in \mathcal{B}(S)$  wohldefiniert ist.

**Definition 2.14.** *Eine ergodische Markov-Kette  $\Phi$  heißt  $V$ -gleichmäßig ergodisch, wenn*

$$\|P^n - \pi\|_V \longrightarrow 0, \quad n \rightarrow \infty.$$

Als Hauptresultat zeigen wir, dass aus der Drift-Bedingung (2.8)  $V$ -gleichmäßige Ergodizität folgt.

**Satz 2.15.** *Sei  $\Phi$  eine  $\psi$ -irreduzible, aperiodische Markov-Kette und sei die Drift-Bedingung (2.8) erfüllt, mit  $C$  petite und  $V$  endlich. Dann gilt für ein  $r > 1$*

$$\sum_{n=1}^{\infty} r^n \|P^n - \pi\|_V < \infty$$

und  $\Phi$  ist  $V$ -gleichmäßig ergodisch.

Beweis: (vgl [12], Theorem 16.1.2)

In Satz 2.11 wird gezeigt, dass für ein  $R < \infty$  und  $\rho > 1$

$$\|P^n(x, \cdot) - \pi\|_V \leq RV(x)\rho^n, \quad n \in \mathbb{N}$$

gilt. Aus der Definition von  $\|\cdot\|_V$  folgt direkt die  $V$ -gleichmäßige Ergodizität. □

Ein wichtiger Spezialfall ist der für  $V \equiv 1$ . Falls eine Markov-Kette  $\Phi$

$$\sup_{x \in S} \|P^n(x, \cdot) - \pi\| = 0, \quad n \rightarrow \infty$$

erfüllt, so heißt  $\Phi$  gleichmäßig ergodisch. Folgenden Satz werden wir später benutzen um die gleichmäßige Ergodizität der Hilfskette  $(\theta_t^*)_{t \geq 0}$  zu beweisen.

**Satz 2.16.** *Für jede  $\psi$ -irreduzible aperiodische Markov-Kette  $\Phi$  ist äquivalent:*

i)  $\Phi$  ist gleichmäßig ergodisch.

ii) Es existieren  $r > 1$  und  $R < \infty$ , so dass für alle  $x$

$$\|P^n(x, \cdot) - \pi(\cdot)\| \leq Rr^{-n}.$$

iii) Es existiert ein  $n$  mit

$$\sup_{x \in S} \|P^n(x, \cdot) - \pi(\cdot)\| < 1.$$

iv) Der Zustandsraum  $S$  ist  $\nu_m$ -petite für ein  $m$ .

## 2 Stabilität und geometrische Ergodizität

v) Die Kette ist aperiodisch und es existiert eine beschränkte Lösung  $V \geq 1$  von

$$\Delta V(x) \leq -\beta V(x) + b \mathbf{1}_C(x), \quad x \in S,$$

für  $\beta > 0$ ,  $b < \infty$  und eine petite Menge  $C$ .

Beweis: (vgl. [12], Theorem 16.2.1 und Theorem 16.2.2)

Wir zeigen nur die Äquivalenz zwischen i) und iv), denn die Implikation iv)  $\Rightarrow$  i) benötigen wir später für den Nachweis der gleichmäßigen Ergodizität von  $(\theta_t^*)_{t \geq 0}$ .

i)  $\Rightarrow$  iv) Da  $\Phi$  gleichmäßig ergodisch und  $\psi$ -irreduzibel ist, folgt die  $\pi$ -Irreduzibilität von  $\Phi$ . Demnach existiert eine small Menge  $A \in \mathcal{B}^+(S)$  (vgl. Satz 1.16). Weiterhin existiert ein  $n_0$ , so dass  $\inf_{x \in S} P^{n_0}(x, A) > 0$ . Mit Satz 1.21 erhält man, dass  $S$  small bzw. petite ist.

iv)  $\Rightarrow$  v) Setze  $V \equiv 1$ ,  $\beta = b = \frac{1}{2}$  und  $C = S$ .

v)  $\Rightarrow$  i) Mit Satz 2.15 folgt direkt die gleichmäßige Ergodizität von  $\Phi$ .  $\square$

Für den eigentlichen Prozess  $(X_t)_{t \geq 0}$  werden wir  $V$ -gleichmäßige Ergodizität nachweisen. Hierzu muss eine Testfunktion  $V$  konstruiert werden, welche der Drift-Bedingung genügt. Die Drift-Bedingung wird allerdings nicht direkt gezeigt, sondern mit Hilfe eines Lemmas gefolgert. Das Lemma wird schon an dieser Stelle eingebracht, um den Hauptbeweis der Arbeit zu entlasten. Beachte, dass an dieser Stelle bereits die Testfunktion  $V$  konstruiert wird! Der Hauptteil beschäftigt sich damit, die Voraussetzungen des Lemmas, insbesondere (2.9) zu verifizieren.

**Lemma 2.17.** *Sei  $(X_t)_{t \geq 0}$  eine  $\psi$ -irreduzible  $T$ -chain mit Werten in  $\mathbb{R}^p$ ,  $\nu$  eine beschränkte Funktion und  $\rho_1 < 1$  sowie  $n \geq 0$ , so dass*

$$\limsup_{\|x\| \rightarrow \infty} E \left[ \nu(X_{n+1}) - \nu(X_n) + \log \left( \frac{1 + \|X_{n+1}\|}{1 + \|X_n\|} \right) \middle| X_0 = x \right] < \log \rho_1 < 0. \quad (2.9)$$

Dann gilt für eine Funktion  $\lambda : \mathbb{R}^p \rightarrow (0, \infty)$  beschränkt und beschränkt weg von Null und ein  $s > 0$

$$\limsup_{\|x\| \rightarrow \infty} E \left[ \frac{\lambda(X_1) \|X_1\|^s}{\lambda(x) \|x\|^s} \middle| X_0 = x \right] < \rho_1^s < 1, \quad (2.10)$$

falls  $\sup_{x \in \mathbb{R}_0^p} E_x \left[ \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < \infty$  und  $E \left[ \lambda(X_1) \|X_1\|^s \right]$  lokal beschränkt ist. Insbesondere ist  $(X_t)_{t \geq 0}$   $V$ -gleichmäßig ergodisch mit  $V(x) = 1 + \lambda(x) \|x\|^s$ .

Beweis: (vgl. [6], Lemma 5.10)

Wir weisen zunächst die Drift-Bedingung (2.8) nach. Dazu prüfen wir, wann

$$\sup_{x \in \mathbb{R}^p} E_x \left[ \frac{V(X_1)}{V(x)} \right] < 1$$

## 2 Stabilität und geometrische Ergodizität

gilt. Hierfür betrachten wir

$$\begin{aligned}
\limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{V(X_1)}{V(x)} \right] &= \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{1 + \lambda(X_1) \|X_1\|^s}{1 + \lambda(x) \|x\|^s} \right] \\
&= \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{1}{1 + \lambda(x) \|x\|^s} + \frac{\lambda(X_1) \|X_1\|^s}{1 + \lambda(x) \|x\|^s} \right] \\
&\leq \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{1}{1 + \lambda(x) \|x\|^s} \right] + \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{\lambda(X_1) \|X_1\|^s}{\lambda(x) \|x\|^s} \right] \\
&= \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{\lambda(X_1) \|X_1\|^s}{\lambda(x) \|x\|^s} \right] < 1.
\end{aligned}$$

Demnach existiert ein  $M > 0$ , so dass

$$\sup_{\|x\| > M} E_x \left[ \frac{1 + \lambda(X_1) \|X_1\|^s}{1 + \lambda(x) \|x\|^s} \right] < 1. \tag{2.11}$$

Definiere die kompakte Menge  $C = \{x, \|x\| \leq M\}$ , die nach Satz 1.26 petite ist. Für  $x \in C^c$  haben wir die Drift-Bedingung dann schon gezeigt, denn aus (2.11) folgt die Existenz einer Konstante  $\beta > 0$ , so dass

$$E_x \left[ \frac{1 + \lambda(X_1) \|X_1\|^s}{1 + \lambda(x) \|x\|^s} \right] \leq 1 - \beta, \quad \forall x \in C^c.$$

Bleibt

$$E_x \left[ V(X_1) - V(x) \right] < \infty \quad \text{für } x \in C$$

zu zeigen. Dies ist aber äquivalent zu

$$E_x V(X_1) < \infty \quad \text{für } \|x\| \leq M$$

und somit erfüllt, da  $E_x \left[ \lambda(X_1) \|X_1\|^s \right]$  nach Voraussetzung lokal beschränkt ist. Die V-gleichmäßige Ergodizität der Kette  $(\bar{X}_t)_{t \geq 0}$  ist damit gezeigt.

Es muss jetzt die Ungleichung

$$E_x \left[ \frac{\lambda(X_1) \|X_1\|^s}{\lambda(x) \|x\|^s} \right] < \rho_1^s < 1$$

hergeleitet werden. Hierzu benötigen wir einen Hilfssatz.

**Satz 2.18.** *Sei  $(Y_n)_n$  eine Folge nichtnegativer Zufallsvariablen. Dann ist  $\sup EY_n^r < 1$ , für ein  $r > 0$  genau dann wenn  $\sup_n EY_n^r < \infty$  für ein  $r > 0$  und  $\sup_n E \log(\delta + Y_n) < 0$  für ein  $\delta > 0$ .*

## 2 Stabilität und geometrische Ergodizität

Nach Voraussetzung gilt

$$\begin{aligned}
 & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \nu(X_1) - \nu(x) + \log \left( \frac{1 + \|X_1\|}{1 + \|x\|} \right) \right] < \log \rho_1 \\
 \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \nu(X_1) - \nu(x) + \log \left( \frac{\|X_1\|}{\|x\|} \right) \right] < \log \rho_1 \\
 \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{e^{\nu(X_1)}}{e^{\nu(x)} \rho_1} \cdot \frac{\|X_1\|}{\|x\|} \right) \right] < 0.
 \end{aligned}$$

Beachte für die zweite Zeile die Rechnung von (3.28) und (3.29). Die Rechnung wurde nicht an dieser Stelle durchgeführt, da man hierfür mehr Wissen über den Prozess  $(X_t)_{t \geq 0}$  benötigt. Mit Satz 2.18 folgt wegen  $\sup_{x \in \mathbb{R}_0^p} E_x \left[ \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < \infty$ ,

$$\begin{aligned}
 & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{e^{\nu(X_1)s}}{e^{\nu(x)s}} \frac{1}{\rho_1^s} \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < 1 \\
 \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{\lambda(X_1)}{\lambda(x)} \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < \rho_1^s \\
 \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \frac{\lambda(X_1)}{\lambda(x)} \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < \rho_1^s,
 \end{aligned}$$

wobei  $\lambda(x) := e^{\nu(x) \cdot s}$  gewählt wurde. Damit ist das Lemma bewiesen. □

Um die Piggyback-Methode zu verstehen, ist es wichtig die Konstruktion der Testfunktion  $V$  nachzuvollziehen.  $V$  wurde basierend auf der Funktion  $\nu$  definiert. Die Piggyback-Methode leitet  $\nu$  mit Hilfe des Hilfsprozesses  $(\theta_t^*)_{t \geq 0}$  her.

# 3 Die Piggyback-Methode

## 3.1 ARCH-Prozesse

Dieser Abschnitt hat das Ziel, den TAR-ARCH-Prozess als Verallgemeinerung der ARCH-Prozesse zu verstehen. Dadurch soll ein besseres Verständnis der einzelnen Terme erreicht werden. Die von Robert F. Engle eingeführten ARCH(q)-Modelle haben den Zweck die bedingte Varianz eines Prozesses in Abhängigkeit von dessen Vergangenheit zu modellieren. Dabei hat sich eine gute Übereinstimmung dieser Modelle mit Finanzzeitreihen herausgestellt. Ein wesentlicher Aspekt ist das Auftreten von Volatilitätsclustern, d.h. große Abweichungen treten in Gruppen auf. Die ARCH(p)-Prozesse wurden z.B. von Weiss (1984) (AR-ARCH-Prozess) und Bollerslev (1986) (GARCH-Prozess) erweitert, bis zu dem TAR-ARCH-Modell, welches in dieser Arbeit behandelt wird.

### 3.1.1 Der ARCH(q)-Prozess

Der ARCH(1)-Prozess ist ein Spezialfall des von Engle (1982) eingeführten ARCH(q)-Prozesses definiert durch

$$\begin{aligned}\xi_t &= \sigma_t e_t \text{ mit } \sigma_t^2 = \beta_0 + \beta_1 \xi_{t-1}^2 \\ \Rightarrow \xi_t &= \sqrt{\beta_0 + \beta_1 \xi_{t-1}^2} \cdot e_t.\end{aligned}$$

Hierbei seien die  $e_t$  eine Folge unabhängiger identisch verteilter Zufallsvariablen. Es wird also insbesondere die bedingte Varianz  $\sigma_t^2$  durch die Vergangenheit des Prozesses modelliert. Für die Stabilitätsbetrachtung des Modells ergibt sich:

- Für  $\beta_1 = 0$  stimmt der Prozess  $(\xi_t)_{t \geq 0}$  bis auf einen skalaren Faktor mit den Zufallsvariablen  $e_t$  überein.
- Für  $\beta_1 \in (0, 1)$  ist  $(\xi_t)_{t \geq 0}$  stationär mit endlicher Varianz.
- Für  $\beta_1 \in [1, 2e^\gamma)$ , wobei  $\gamma$  die Eulersche Konstante bezeichnet, ist  $(\xi_t)_{t \geq 0}$  stationär mit unendlicher Varianz.

Vergleiche hierzu [7], Abschnitt 8.4, insbesondere Seite 466. Während der ARCH(1)-Prozess nur von dem vorherigen Wert abhängt, ist der ARCH(q)-Prozess von q Zeit-

### 3 Die Piggyback-Methode

punkten in der Vergangenheit abhängig.

$$\begin{aligned}\xi_t &= \sigma_t e_t \text{ mit } \sigma_t^2 = \beta_0 + \sum_{i=1}^q \beta_i \xi_{t-i}^2, \\ \Rightarrow \xi_t &= \sqrt{\beta_0 + \sum_{i=1}^q \beta_i \xi_{t-i}^2} \cdot e_t.\end{aligned}$$

#### 3.1.2 Der AR(p)-Prozess

Der AR(1)-Prozess tauchte bereits am Anfang der Stabilitätsüberlegungen auf. Im Gegensatz zu den ARCH-Prozessen liegt hier der Fokus auf der Modellierung der bedingten Erwartung. Der AR(1)-Prozess hängt nur von dem vorherigen Wert des Prozesses ab:

$$\xi_t = \alpha \xi_{t-1} + e_t.$$

In diesem Fall lässt sich eine Rekursion sinnvoll ausnutzen:

$$\xi_t = \alpha \xi_{t-1} + e_t = \alpha^2 \xi_{t-2} + \alpha e_{t-1} + e_{t-2} = \dots = \alpha^m \xi_{t-m} + \sum_{j=0}^{m-1} \alpha^j e_{t-j}.$$

Für den deterministischen Teil sieht man direkt ein, dass der Prozess für  $\alpha < 1$  stabil ist. Wie schon erwähnt ist der stochastische Prozess für  $\alpha < 1$  stationär (vgl. [13], Kapitel 4.1). Der AR-Prozess lässt sich verallgemeinern, indem man zulässt, dass der Prozess von  $p$  Werten aus der Vergangenheit abhängig ist:

$$\xi_t = \sum_{i=1}^p \alpha_i \xi_{t-i} + e_t.$$

#### 3.1.3 Der AR(p)-ARCH(q)-Prozess

Die beiden Prozessarten lassen sich zum AR(p)-ARCH(q)-Prozess kombinieren:

$$\xi_t = \sum_{i=1}^p \alpha_i \xi_{t-i} + \sqrt{\beta_0 + \sum_{j=1}^q \beta_j \xi_{t-j}^2} \cdot e_t.$$

Dabei bestimmt der AR(p)-Anteil die bedingte Erwartung, während der ARCH(q)-Anteil die bedingte Varianz modelliert. Der Spezialfall  $\xi_t = \alpha \xi_{t-1} + \sqrt{\beta_0 + \beta_1 \xi_{t-1}^2} \cdot e_t$  mit  $p = 1$  und  $q = 1$  wurde von Borkovec, M. und Klüppelberg, C. untersucht. Es stellt sich heraus, dass der Prozess geometrisch ergodisch ist, falls  $E \left[ \log |\alpha + \sqrt{\beta_1} e_t| \right] < 0$  erfüllt ist (vgl. [3], Theorem 1).



### 3.1.4 Der TAR-ARCH-Prozess

Eine weitere Verallgemeinerung des AR(p)-ARCH(p)-Prozess ist das FCAR-ARCH-Modell (functional coefficients AR-ARCH-Modell). Wie der Name des Modells impliziert, soll es sich bei den  $\alpha_i$  und  $\beta_j$  nicht um Konstanten, sondern um Funktionen  $\alpha_i(\xi_{t-1}, \dots, \xi_{t-p})$  und  $\beta_j(\xi_{t-1}, \dots, \xi_{t-p})$  handeln. Der Prozess hat dann die Form

$$\xi_t = \alpha_0(\xi_{t-1}, \dots, \xi_{t-p}) + \sum_{i=1}^p \alpha_i(\xi_{t-1}, \dots, \xi_{t-p}) \xi_{t-i} + \sqrt{\beta_0^2(\xi_{t-1}, \dots, \xi_{t-p}) + \sum_{j=1}^q \beta_j^2(\xi_{t-1}, \dots, \xi_{t-p}) \xi_{t-j}^2} \cdot e_t.$$

Lässt man weiterhin Unstetigkeiten der Funktionen zu, so handelt es sich um das TAR-ARCH-Modell.

### 3.2 Annahmen an das TAR-ARCH-Modell

Das TAR-ARCH-Modell, welches wir untersuchen wollen, hat die Form

$$\xi_t = a(\xi_{t-1}, \dots, \xi_{t-p}) + b(\xi_{t-1}, \dots, \xi_{t-p}) \cdot e_t,$$

wobei  $a, b : \mathbb{R}^p \rightarrow \mathbb{R}$  stückweise stetige Funktionen seien und  $(e_t)_{t \geq 0}$  eine Folge unabhängig identisch verteilter Zufallsvariablen sei. Die Funktionen  $a$  und  $b$  seien stetig auf einzelnen, zusammenhängenden Teilmengen des  $\mathbb{R}^p$ . Den Rand dieser Regionen nennen wir Schwellenwerte. Um die Stabilität des Prozesses  $(X_t)_{t \geq 0} = ((\xi_t, \dots, \xi_{t-p+1}))_{t \geq 0}$  zu erhalten, müssen einige Annahmen über die Funktionen  $a, b$  sowie den Fehlerterm  $(e_t)_{t \geq 0}$  getroffen werden. Zu beachten ist, dass es sich um Standardannahmen für nichtlineare Modelle handelt. Dass keine stärkeren Annahmen notwendig sind, um die V-geometrische Ergodizität zu zeigen, darf als Stärke der Piggyback-Methode angesehen werden.

Die erste Annahme sorgt dafür, dass es sich bei  $(X_t)_{t \geq 0}$  um eine aperiodische,  $\lambda_p$ -irreduzible T-chain auf  $\mathbb{R}^p$  handelt.  $\lambda_p$  sei das Lebesgue-Maß auf  $\mathbb{R}^p$ .

**Annahme 1.** *Die Verteilung der  $e_t$  habe eine Lebesgue-Dichte  $f$ , die lokal weg von 0 beschränkt ist. Ausserdem sei  $b$  positiv, lokal beschränkt und lokal weg von Null beschränkt.*

Dabei heißt eine Funktion  $g : \mathbb{R}^p \rightarrow [0, \infty)$  lokal weg von Null beschränkt, wenn für jede kompakte Menge  $A \subset \mathbb{R}^p$  ein  $M_A > 0$  existiert, so dass  $g(x) \geq M_A, \forall x \in A$ . Die zweite Annahme setzt die Existenz von Momenten voraus.

**Annahme 2.** *Es gelte  $\sup_{u \in \mathbb{R}} (1 + |u|)f(u) < \infty$  und  $E[|e_1|^{r_0}] < \infty$  für ein  $r_0 > 0$ .*

Die folgende Annahme beinhaltet insbesondere, dass die Funktionen  $a(x)$  und  $b(x)$  im Limes  $x \rightarrow \infty$  nicht schneller als linear anwachsen.

**Annahme 3.** *Die beiden Funktionen  $\frac{a(x)}{1+\|x\|}$  und  $\frac{b(x)}{1+\|x\|}$  seien beschränkt.*

Des Weiteren werden  $a$  und  $b$  als asymptotisch homogen angenommen, d.h.  $a$  und  $b$  konvergieren gegen homogene Funktionen.

**Annahme 4.** *Es existieren ein  $x_* \in \mathbb{R}^p$  und beschränkte Funktionen  $a_*$  und  $b_*$  auf  $\Theta := \{\theta \in \mathbb{R}^p : \|\theta\| = 1\}$ , so dass*

$$\limsup_{\omega \rightarrow \infty} \sup_{\theta \in \Theta} \left| \frac{a(x_* + \omega\theta)}{\omega} - a_*(\theta) \right| = 0 \quad \text{und} \quad \limsup_{\omega \rightarrow \infty} \sup_{\theta \in \Theta} \left| \frac{b(x_* + \omega\theta)}{\omega} - b_*(\theta) \right| = 0.$$

Zur Vereinfachung nehmen wir  $x_* = 0$  an. Man kann leicht folgern, dass  $a$  durch  $a(x) = a_0(x) + a^*(x)$  mit homogenem  $a^*(x)$  und  $a_0 = o(\|x\|)$  darstellbar ist. Geht man nämlich von einer Zerlegung  $a(x) = a_0(x) + a^*(x)$  aus, wobei  $a^*(x)$  homogen sei, dann muss

$$\limsup_{\omega \rightarrow \infty} \sup_{\theta \in \Theta} \left| \frac{a_0(\omega\theta)}{\omega} + \frac{a^*(\theta)\omega}{\omega} - a_*(\theta) \right| = 0$$

### 3 Die Piggyback-Methode

erfüllt sein.  $a^*(\theta)$  und  $a_*(\theta)$  hängen nicht von  $\omega$  ab, deshalb folgt  $a^*(\theta) = a_*(\theta)$ . Dann gilt aber gerade  $a_0(x) = o(\|x\|)$ . Solch eine Zerlegung erhält man auch wenn  $x_* \neq 0$ , indem man einfach den homogenen Teil um  $x_*$  verschiebt. Der Einfachheit halber beschränken wir uns im folgenden auf den Fall  $x_* = 0$ . Eine analoge Aussage erhält man natürlich für  $b(x) = b_0(x) + b^*(x)$ .

Auch wenn  $b$  lokal weg von Null beschränkt ist, gilt dies nicht unbedingt für  $b^*$ . Oft hat  $b^*$  bei ARCH-Modellen die Form  $b^*(x) = \sqrt{b_1^2 x_1^2 + \dots + b_p^2 x_p^2}$ , wobei einige der  $b_i$  Null sein können. Sei zum Beispiel  $b_1 = 0$ , dann verschwindet  $b^*$  auf der gesamten  $x_1$ -Achse und ist somit nicht lokal weg von Null beschränkt. Bei der Analyse des homogenen Prozesses nehmen wir deshalb anfangs die Achsenflächen  $H_0 := \{x \in \mathbb{R}^p, \min_i |x_i| = 0\}$  aus dem Zustandsraum heraus.

**Annahme 5.** Für  $p > 1$  sei  $\max(|a^*(x)|, b^*(x))$  lokal weg von Null beschränkt auf  $\mathbb{R}_0^p$  und  $b^*(x)$  sei lokal weg von Null beschränkt auf  $\mathbb{R}_\times^p = \mathbb{R}^p \setminus H_0$ .

Die letzte Annahme bezieht sich auf die Schwellenwerte des Modells.

**Annahme 6.** Für  $p > 1$  existieren  $p-1$ -dimensionale homogene Hyperflächen  $H_1, \dots, H_m$ , so dass  $a^*$  und  $b^*$  stetig sind auf  $\mathbb{R}^p \setminus (\bigcup_{j=1}^m H_j)$ . Homogene Hyperfläche bedeutet, dass die Fläche durch den Ursprung verläuft.

### 3.3 Einleitung zur Piggyback-Methode

In diesem Abschnitt wird die Piggyback-Methode vorgestellt, die benutzt wird, um Vergleichmäßige Ergodizität des Prozesses  $(X_t)_{t \geq 0}$  nachzuweisen. Der Prozess  $(X_t)_{t \geq 0}$  ist gegeben durch

$$X_t = (\xi_t, \dots, \xi_{t-p+1}) \text{ mit} \\ \xi_t = a(\xi_{t-1}, \dots, \xi_{t-p}) + b(\xi_{t-1}, \dots, \xi_{t-p})e_t.$$

Der Prozess  $(\xi_t)_{t \geq 0}$  ist keine Markov-Kette, da er von  $p$  Werten aus der Vergangenheit abhängt und nicht nur von dem vorherigen Zustand. Aus diesem Grund betrachten wir den Prozess  $(X_t)_{t \geq 0}$ , bei dem es sich um eine Markov-Kette handelt (vgl. Bemerkung 1.4). Der Vektor  $X_t$  ist gerade so gewählt, dass er in der ersten Komponente den Wert  $\xi_t$  enthält und der Vektor  $X_{t-1}$  die  $p$  vergangenen Werte, von denen  $\xi_t$  abhängt. Demnach hängt  $X_t$  nur von  $X_{t-1}$  ab. Definiert man den homogenen Prozess

$$X_t^* = (\xi_t^*, \dots, \xi_{t-p+1}^*) \text{ mit} \\ \xi_t^* = a^*(\xi_{t-1}^*, \dots, \xi_{t-p}^*) + b^*(\xi_{t-1}^*, \dots, \xi_{t-p}^*)e_t$$

sieht man schnell ein, dass es sich um eine Markov-Kette auf  $\mathbb{R}_0^p$  handelt. Aufgrund der Homogenität von  $a^*$  und  $b^*$  ist der homogene, normierte Prozess  $(\theta_t^*)_{t \geq 0}$  gegeben durch

$$\theta_t^* = \frac{X_t^*}{\|X_t^*\|}$$

### 3 Die Piggyback-Methode

eine Markov-Kette auf  $\Theta = \{x \in \mathbb{R}^p, \|x\| = 1\}$ .

In Kapitel 2 "Stabilität und geometrische Ergodizität" wurde gezeigt, dass eine Markov-Kette geometrisch ergodisch ist, falls die Drift-Bedingung (2.8),

$$\Delta V(x) \leq -\beta V(x) + b \mathbf{1}_{\{x \notin C\}},$$

für eine petite Menge  $C$  erfüllt ist. Wir werden die Drift-Bedingung allerdings nicht direkt zeigen, sondern mit Hilfe der Piggyback-Methode (2.9),

$$\limsup_{\|x\| \rightarrow \infty} E \left[ \nu(X_{n+1}) - \nu(X_n) + \log \left( \frac{1 + \|X_{n+1}\|}{1 + \|X_n\|} \right) \middle| X_0 = x \right] < \log \rho_1 < 0,$$

aus Lemma 2.17 nachweisen. Die Piggyback-Methode geht nun folgendermaßen vor:

- Wir zeigen, dass es sich bei  $(X_t^*)_{t \geq 0}$  und  $(\theta_t^*)_{t \geq 0}$  um aperiodische,  $\psi$ -irreduzible T-chains handelt.
- Der homogene, normierte Prozess  $(\theta_t^*)_{t \geq 0}$  ist gleichmäßig ergodisch und

$$\log \rho = \int_{\Theta} \int_{\mathbb{R}} \log \left( \|(a^*(\theta) + b^*(\theta)u, \theta_1, \dots, \theta_{p-1})\| \right) f(u) \pi(d\theta)$$

ist wohldefiniert.

- Mit Hilfe der gleichmäßigen Ergodizität leiten wir die near-equilibrium Gleichung für eine Testfunktion  $\nu(\theta)$  her, die stetig auf  $\Theta_{\times \times} \subset \Theta$  ist. Dieser Schritt bildet den Schlüssel der Piggyback-Methode. Ausgehend von  $\nu$  lässt sich nämlich eine Testfunktion  $V$  für die gesuchte Drift-Bedingung konstruieren (vgl. Lemma 2.17).
- Falls eine Funktion  $\tilde{V}$  existiert, so dass  $\frac{\log(\tilde{V}(X_{n+1}))}{\log(\tilde{V}(X_n))}$  negativen Erwartungswert hat unter der Bedingung  $X_0 = x$  für große  $x$ , dann kann eine Testfunktion konstruiert werden, um  $V$ -gleichmäßige Ergodizität von  $(X_t)_{t \geq 0}$  zu erhalten.
- Die entscheidenden Eigenschaften, um dies zu zeigen sind die near-equilibrium Gleichung, die Stetigkeit von  $\nu(\theta)$  auf  $\Theta_{\times \times}$  und die Eigenschaft, dass sich  $\frac{\tilde{X}_t^*}{\|\tilde{X}_t^*\|}$  und  $\frac{X_t}{\|X_t\|}$  mit hoher Wahrscheinlichkeit in derselben kompakten Menge  $\tilde{D}$  aufhalten.
- Schließlich zeigen wir, dass die Stabilitätsbedingung scharf ist. Hierbei nutzen wir aus, dass  $\log \rho$  mit dem Lyapounov-Exponenten übereinstimmt.

Der Kern der Piggyback-Methode soll noch einmal hervorgehoben werden: Aufgrund der gleichmäßigen Ergodizität von  $(\theta_t^*)_{t \geq 0}$  lässt sich eine Testfunktion  $\nu$  finden, welche die near-equilibrium-Gleichung erfüllt. Die besagte Gleichung weist bereits eine deutliche Analogie zu (2.9) auf. Basierend auf  $\nu$  konstruiert man eine Testfunktion  $V$ , welche der Drift-Bedingung genügt.

### 3.4 Der homogene, normierte Prozess $(\theta_t^*)_{t \geq 0}$

Auf  $\mathbb{R}^p$  legen wir das Lebesgue-Maß  $\lambda_p$  zu Grunde. Nach Annahme 1 besitzen die  $e_t$  eine Lebesgue-Dichte. Auf dem Zustandsraum  $\Theta = \{x \in \mathbb{R}_0^p, \|x\| = 1\}$  von  $(\theta_t^*)_{t \geq 0}$  hingegen betrachten wir das Oberflächenmaß  $\lambda_\Theta$ . Dabei gilt die Beziehung

$$\lambda_\Theta(A) = p \cdot \lambda_p \left( \left\{ x \in \mathbb{R}_0^p, \frac{x}{\|x\|} \in A, \|x\| \leq 1 \right\} \right), A \in \mathcal{B}(\mathbb{R}^p).$$

Die Hauptergebnisse dieses Abschnitts sind die gleichmäßige Ergodizität von  $(\theta_t^*)_{t \geq 0}$  und die daraus resultierende near-equilibrium-Gleichung (3.11). Eine weitere wichtige Eigenschaft ist die Stetigkeit der Testfunktion  $\nu$  auf einer Teilmenge von  $\Theta$ . Die beiden letzteren Resultate werden explizit für die Abschätzung von (2.9) benötigt.

#### 3.4.1 Die gleichmäßige Ergodizität von $(\theta_t^*)_{t \geq 0}$

Um die gleichmäßige Ergodizität von  $(\theta_t^*)_{t \geq 0}$  nachzuweisen, zeigen wir, dass der Zustandsraum  $\Theta$  petite ist (vgl. Satz 2.16 iv)). Da es im allgemeinen nicht leicht ist eine petite Menge zu bestimmen, zeigen wir dass  $(\theta_t^*)_{t \geq 0}$  eine  $\lambda_\Theta$ -irreduzible T-chain ist. In diesem Fall ist jede kompakte Teilmenge des Zustandsraums petite (vgl. Satz 1.26).

Analog zu  $\mathbb{R}_\times^p = \mathbb{R}^p \setminus H_0$  definieren wir  $\Theta_\times = \Theta \setminus H_0$ , wobei  $H_0 = \{x, \min_i |x_i| = 0\}$  die Menge der Achsenflächen bezeichnet.

**Satz 3.1.** *Seien Annahmen 1, 4 und 5 vorausgesetzt. Dann gilt:*

- i)  $(X_t)_{t \geq 0}$  ist eine  $\lambda_p$ -irreduzible, aperiodische T-chain auf  $\mathbb{R}^p$ .
- ii)  $(X_t^*)_{t \geq 0}$  ist eine  $\lambda_p$ -irreduzible, aperiodische T-chain auf  $\mathbb{R}_\times^p$ .
- iii)  $(\theta_t^*)_{t \geq 0}$  ist eine  $\lambda_\Theta$ -irreduzible, aperiodische T-chain auf  $\Theta_\times$ .

Beweis: (vgl. [6], Lemma 5.1 und Lemma 5.2)

i) (vgl. [4], Theorem 2.2 ii) und Example 2.1)

ii) Sei  $g_\times^n(x, \cdot)$  die n-Schritt Übergangsdichte für den Prozess  $(X_t^*)_{t \geq 0}$  auf  $\mathbb{R}_\times^p$ . Nach Annahme 1 und Annahme 5 ist  $b^*$  und die Dichte der  $e_t$  weg von Null beschränkt. Demnach ist  $g_\times^n(x, \cdot)$ ,  $\forall x \in \mathbb{R}_\times^p$  weg von Null beschränkt und wohldefiniert für  $n \geq p$ . Falls  $n < p$ , ist mindestens eine Komponente von  $X_{t+n}^*$  durch  $X_t^*$  festgelegt. Es lässt sich dann eine Menge  $B \in \mathcal{B}^+(\mathbb{R}_\times^p)$  finden, welche diese Komponenten nicht enthält. In diesem Fall wäre die Übergangswahrscheinlichkeit  $P^n(x, B) = 0$  und damit die Übergangsdichte  $g_\times^n(x, A)$ ,  $\forall A \subset B$  gleich Null. Zur Veranschaulichung wählen wir  $p=2$ . Sei  $X_1 = (10, 1)$ , dann gilt für  $X_2 = (\xi_2, 10)$ , dass die zweite Komponente fest ist.  $X_2$  kann keine Menge erreichen, die als zweite Komponente nicht die 10 enthält. Für  $X_3$  dagegen ist jede Menge mit positiver Wahrscheinlichkeit erreichbar.

Wir zeigen die  $\lambda_p$ -Irreduzibilität von  $(X_t^*)_{t \geq 0}$ . Sei  $A \in \mathcal{B}^+(\mathbb{R}_\times^p)$  und  $n > p$ , dann folgt

$$P^n(x, A) = \int_A g_\times^n(x, y) \lambda_p(dy) > 0, \forall x \in \mathbb{R}^p,$$

### 3 Die Piggyback-Methode

da  $g_{\times}^n(x, \cdot)$  weg von null beschränkt ist. Damit ist insbesondere  $L(x, A) > 0$  und somit  $(X_t^*)_{t \geq 0}$   $\lambda_p$ -irreduzibel. Weiterhin ist  $\mathbb{R}_{\times}^p$  absorbierend, denn es gilt

$$\begin{aligned} P(x, \mathbb{R}_{\times}^p) &= \int_{\mathbb{R}_{\times}^p} g_{\times}^n(x, y) \lambda_p(dy) \\ &= \int_{\mathbb{R}^p} g_{\times}^n(x, y) \lambda_p(dy) - \int_{H_0} g_{\times}^n(x, y) \lambda_p(dy) = 1, \end{aligned}$$

da  $H_0$  eine  $\lambda_p$  Nullmenge ist.

Als nächstes zeigen wir, dass es sich bei  $(X_t^*)_{t \geq 0}$  um eine aperiodische handelt. Sei  $C$  eine small Menge, d.h es existiert ein Maß  $\nu$ , so dass

$$P^m(x, B) \geq \nu(B) > 0, \quad \forall x \in C, B \in \mathcal{B}(\mathbb{R}_{\times}^p) \text{ und } m \geq p.$$

Die Existenz einer solchen Menge ist durch Satz 1.16 gesichert. Um die Periode des Prozesses zu bestimmen, müssen wir den größten gemeinsamen Teiler der Menge

$$E_C = \{n \geq 1, C \text{ ist } \nu_n\text{-small mit } \nu_n = \delta_n \nu, \text{ für } \delta_n > 0\}$$

ermitteln (vgl. (1.9)). Dazu betrachten wir die Wahrscheinlichkeit ausgehend von  $x \in C$  eine Menge  $B \in \mathcal{B}^+(\mathbb{R}_{\times}^p)$  zum Zeitpunkt  $m+k$  zu erreichen,  $k \geq m \geq p$ . Eine Anwendung der Chapman-Kolmogorov-Gleichung (1.3) liefert

$$\begin{aligned} P^{m+k}(x, B) &= \int_X P^m(x, dy) P^k(y, B) \\ &\geq \int_B P^m(x, dy) P^k(y, B) \\ &\geq P^m(x, B) \inf_{y \in B} P^k(y, B) \\ &\geq \nu(B) \delta_k. \end{aligned} \tag{3.1}$$

$\delta_k := \inf_{y \in B} P^k(y, B)$  ist dabei für alle  $k \geq m$  größer als Null, da  $(X_t^*)_{t \geq 0}$   $\lambda_p$ -irreduzibel ist. (3.1) ist deshalb zum Beispiel für die Zeitpunkte  $m+p+1$  und  $m+p+2$  erfüllt. Die Periode beträgt deshalb  $\text{ggT}(E_C) = 1$  und die Markov-Kette ist aperiodisch.

Bleibt noch zu zeigen, dass es sich bei  $(X_t^*)_{t \geq 0}$  um eine T-chain handelt. Nach Satz 1.25 ist dies der Fall, wenn  $K_{a_z}$ , eine in  $z$  nichttriviale, stetige Komponente  $T_z$  besitzt. Zur Erinnerung: T ist eine stetige Komponente von  $K_a$ , falls  $T$  ein substochastischer Kern,  $T(\cdot, A)$  unterhalbstetig ist und

$$K_a(x, A) \geq T(x, A), \quad x \in X, A \in \mathcal{B}(X) \tag{3.2}$$

gilt. Für jedes  $z \in \mathbb{R}_{\times}^p$  sei  $G_z$  eine offene Menge, die  $z$  enthält und dessen Abschluss kompakt und in  $\mathbb{R}_{\times}^p$  enthalten ist. Dann definiert

$$T_z(x, A) := \mathbf{1}_{G_z}(x) \int_A \inf_{u \in G_z} g_{\times}^p(u, y) \lambda_p(dy), \quad A \in \mathcal{B}(\mathbb{R}_{\times}^p),$$

### 3 Die Piggyback-Methode

einen stochastischen Übergangskern auf  $\mathbb{R}_x^p$ , denn  $T_z(x, \cdot)$  ist offensichtlich ein Maß und  $T_z = (\cdot, A)$  ist wegen  $A \in \mathcal{B}(\mathbb{R}_x^p)$  und  $G_z \in \mathcal{B}(\mathbb{R}_x^p)$  messbar.

Für  $x = z$  verschwindet  $T_z(x, \cdot)$  nicht identisch, denn für  $A \in \mathcal{B}^+(\mathbb{R}_x^p)$  folgt

$$T_z(z, A) = \mathbf{1}_{G_z}(z) \int_A \inf_{u \in G_z} g_x^p(u, y) \lambda_p(dy) = \int_A \inf_{u \in G_z} g_x^p(u, y) \lambda_p(dy) > 0,$$

da  $g_x^p(z, \cdot)$  weg von Null beschränkt ist.  $T_z(z, \cdot)$  ist demnach nichttrivial. Weiterhin ist  $T_z(x, A)$  für jedes  $A \in \mathcal{B}(\mathbb{R}_x^p)$  eine unterhalbstetige Funktion, da das Integral nicht von der Variablen  $x$  abhängt und  $\mathbf{1}_{G_z}(x)$  unterhalbstetig ist (vgl. Überlegung nach Definition 1.23). Als letztes muss noch (3.2) gezeigt werden. Es gilt:

$$\begin{aligned} T_z(x, A) &= \mathbf{1}_{G_z}(x) \int_A \inf_{u \in G_z} g_x^p(u, y) \lambda_p(dy) \\ &\leq \int_A g_x^p(x, y) \lambda_p(dy) \\ &= P^p(x, A). \end{aligned}$$

(3.2) ist also für  $K_{\delta_p}(x, A)$  erfüllt. Damit ist der Beweis von ii) vollständig.

iii) Für jede Menge  $A \subset \Theta$  definieren wir  $C_A = \{x \in \mathbb{R}_0^p, \frac{x}{\|x\|} \in A\}$ . Die bedingte Wahrscheinlichkeit eine Menge  $A \in \mathcal{B}(\Theta_x)$  zu erreichen ist gegeben durch

$$P(\theta_t^* \in A | \theta_0^* = \theta) = P(X_t^* \in C_A | X_0^* = c\theta), \text{ für alle } A, t, c, \theta.$$

Alle Aussagen für  $(\theta_t^*)_{t \geq 0}$  lassen sich somit auf die Überlegungen aus ii) zurückführen.  $\square$

Wie in Kapitel 1 "Markov-Ketten" erwähnt, reicht es nicht aus die  $\lambda_p$ -Irreduzibilität zu fordern. Es muss zusätzlich gezeigt werden, dass das Maß  $\lambda_p$  maximal ist. Dies kann man sich aber leicht überlegen, da die  $e_t$  eine Lebesgue-Dichte besitzen. Bis jetzt wurde gezeigt, dass  $(\theta_t^*)_{t \geq 0}$  eingeschränkt auf  $\Theta_x$  eine  $\lambda_{\Theta}$ -irreduzible T-chain ist. In diesem Fall ist jede kompakte Menge in  $\Theta_x$  petite. Im nächsten Lemma wird eine kompakte Menge konstruiert, die von jedem  $\theta \in \Theta$  erreicht werden kann. Daraus folgt, dass auch  $\Theta$  petite ist.

**Lemma 3.2.** *Seien Annahmen 1-5 vorausgesetzt. Für jedes  $\epsilon \in (0, 1)$  existiert eine kompakte Menge  $C \subset \Theta_x$ , so dass folgendes gilt:*

$$i) P(\theta_1^* \in C | \theta_0^* = \theta) > 1 - \epsilon, \text{ für alle } \theta \in C; \tag{3.3}$$

$$ii) P(\theta_p^* \in C | \theta_0^* = \theta) > 1 - p\epsilon, \text{ für alle } \theta \in \Theta. \tag{3.4}$$

Beweis: (vgl. [6], Lemma 5.3)

i) Sei  $\theta_0^* = \theta, \theta \in C$ . Die Wahrscheinlichkeit, dass sich  $\theta_1^*$  in  $C$  befindet, muss durch

$$P(\theta_1^* \in C | \theta_0^* = \theta) > 1 - \epsilon$$

### 3 Die Piggyback-Methode

abgeschätzt werden. Zunächst konstruieren wir die kompakte Menge  $C$  gegeben durch

$$C := \{\theta = (\theta_1, \dots, \theta_p), |\theta_i| \geq \epsilon_C \gamma^{-i}, i = 1, \dots, p\}.$$

Die Bedingungen von  $C$  lassen nur einen solchen Prozess zu, der einen bestimmten Abstand zu den Achsenflächen hält. Die Menge  $C$  ist kompakt, da sie abgeschlossen und als Teilmenge von  $\Theta$  beschränkt ist. Es müssen nun  $\epsilon_C$  und  $\gamma$  so bestimmt werden, dass (3.3) erfüllt ist. Die folgenden Abschätzungen sind dabei hilfreich. Sei  $|u| \leq M$ , dann gilt

$$\begin{aligned} |a^*(\theta) + b^*(\theta)u| &\leq \max(|a^*(\theta)|, b^*(\theta)) \cdot |1 + u| \\ &\leq \sup_{\theta \in \Theta} \max(|a^*(\theta)|, b^*(\theta)) \cdot (1 + |u|) \\ &\leq L_{\text{sup}}(1 + M), \end{aligned}$$

mit  $L_{\text{sup}} := \sup_{\theta \in \Theta} \max(|a^*(\theta)|, b^*(\theta))$ . Nach Annahme 3 sind  $a(\theta)$  und  $b(\theta)$ , wegen  $\|\theta\| = 1$  beschränkt,  $L_{\text{sup}}$  ist deshalb endlich. Für den Vektor  $(a^*(\theta) + b^*(\theta)u, \theta_1, \dots, \theta_{p-1})$  folgt dann

$$\begin{aligned} \|(a^*(\theta) + b^*(\theta)u, \theta_1, \dots, \theta_{p-1})\| &\leq |a^*(\theta) + b^*(\theta)u| + \|\theta\| \\ &\leq L_{\text{sup}}(1 + M) + 1 := \gamma. \end{aligned} \tag{3.5}$$

Dies sei das  $\gamma$ , welches die kompakte Menge  $C$  bestimmt. Als nächstes untersuchen wir, wann sich

$$\theta_1^* = \frac{(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})}{\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\|}$$

in der Menge  $C$  befindet, falls schon  $\theta_0^* = \theta$  aus  $C$  stammt. Wendet man (3.5) auf die erste Komponente an, folgt

$$\frac{|a^*(\theta) + b^*(\theta)e_1|}{\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\|} \geq |a^*(\theta) + b^*(\theta)e_1| \gamma^{-1} \stackrel{!}{\geq} \epsilon_C \gamma^{-1}.$$

Die erste Komponente von  $\theta_1^*$  erfüllt also die Bedingung der Menge  $C$ , falls  $|a^*(\theta) + b^*(\theta)e_1| \geq \epsilon_C$  und  $\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma$ . Für die weiteren Komponenten,  $i = 2, \dots, p$ , erhält man wegen  $\theta \in C$

$$(\theta_1^*)_i = \frac{|\theta_{i-1}|}{\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\|} \geq \frac{\epsilon_C \gamma^{-(i-1)}}{\gamma} = \epsilon_C \gamma^{-i+1-1} = \epsilon_C \gamma^{-i}.$$

$\theta_1^*$  liegt in  $C$ , falls  $\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma$  und  $|a^*(\theta) + b^*(\theta)e_1| \geq \epsilon_C$ . Die erste Ungleichung folgte unter der Annahme  $|e_1| \leq M$ . Für die gesuchte Wahrscheinlichkeit ergibt sich demnach

$$\begin{aligned} P(\theta_1^* \in C | \theta_0^* = \theta) &\geq P(|a^*(\theta) + b^*(\theta)e_1| > \epsilon_C, \|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma) \\ &\geq P(|a^*(\theta) + b^*(\theta)e_1| > \epsilon_C, |e_1| \leq M). \end{aligned} \tag{3.6}$$



### 3 Die Piggyback-Methode

Bleibt zu zeigen: Zu jedem  $\epsilon \in (0, 1)$  können  $\epsilon_C$  und  $M$  so gewählt werden, dass

$$P(|a^*(\theta) + b^*(\theta)e_1| > \epsilon_C, |e_1| \leq M) > 1 - \epsilon.$$

Die notwendige Abschätzung führen wir im Detail durch, da sie im späteren Verlauf der Arbeit in ähnlicher Form noch einmal auftritt. Sei  $\epsilon_C \in (0, \frac{\max(|\alpha|, \beta)}{2})$ , wobei  $\beta \geq 0$  und  $\max(|\alpha|, \beta) > 0$ . Für  $|\alpha + \beta e_t| \leq \epsilon_C$  folgern wir

$$\begin{aligned} |\alpha + \beta e_t| &\leq \epsilon_C \\ \Leftrightarrow -\epsilon_C &\leq \alpha + \beta e_t \leq \epsilon_C \\ \Leftrightarrow -\alpha - \epsilon_C &\leq \beta e_t \leq -\alpha + \epsilon_C \\ \Leftrightarrow \frac{-\alpha - \epsilon_C}{\beta} &\leq e_t \leq \frac{-\alpha + \epsilon_C}{\beta}. \end{aligned}$$

Die Wahrscheinlichkeit desselben Ausdrucks lässt sich dann mit einem Integral über die eben ermittelten Grenzen berechnen:

$$\begin{aligned} P(|\alpha + \beta e_t| \leq \epsilon_C) &= \int_{\frac{-\alpha - \epsilon_C}{\beta}}^{\frac{-\alpha + \epsilon_C}{\beta}} f(u) du \\ &\leq \sup(1 + |u|)f(u) \int_{\frac{-\alpha - \epsilon_C}{\beta}}^{\frac{-\alpha + \epsilon_C}{\beta}} \frac{1}{1 + |u|} du \\ &\leq L \int_{-\alpha - \epsilon_C}^{-\alpha + \epsilon_C} \frac{1}{\beta + |z|} dz. \end{aligned}$$

Die letzte Zeile erhält man mit der Definition  $L := \sup(1 + |u|)f(u)$  und der Substitution  $z = \beta u$ . Für das Integral benutzen wir die Standardabschätzung

$$\int_{-\alpha - \epsilon_C}^{-\alpha + \epsilon_C} \frac{1}{\beta + |z|} dz \leq \max_{z \in [-\alpha - \epsilon_C, -\alpha + \epsilon_C]} \frac{1}{\beta + |z|} 2\epsilon_C.$$

Das Maximum ermitteln wir, indem wir das Minimum von  $|z|$  suchen.

- Falls  $0 \in [-\alpha - \epsilon_C, -\alpha + \epsilon_C]$  liegt das Minimum von  $|z|$  natürlich bei  $|z| = 0$ . In diesem Fall sieht man leicht ein, dass  $|\alpha| - \epsilon_C < 0$  ist.
- Falls  $0 \notin [-\alpha - \epsilon_C, -\alpha + \epsilon_C]$ , wird  $|z|$  minimal für  $|z| = |\alpha| - \epsilon_C$ . Außerdem folgt  $|\alpha| - \epsilon_C > 0$ .

### 3 Die Piggyback-Methode

Das Minimum von  $|z|$  wird also angenommen durch  $\max(|\alpha| - \epsilon_C, 0)$ . Für das Integral ergibt sich

$$\int_{-\alpha - \epsilon_C}^{-\alpha + \epsilon_C} \frac{1}{\beta + |z|} dz \leq 2\epsilon_C \frac{1}{\beta + \max(|\alpha| - \epsilon_C, 0)}.$$

Wählt man  $\epsilon_C \in (0, \frac{\max(|\alpha|, \beta)}{2})$  gilt für den Nenner

$$\beta + \max(|\alpha| - \epsilon_C, 0) > \frac{\max(|\alpha|, \beta)}{2},$$

denn:

- Für  $\beta > |\alpha|$  ist dies klar, da  $\max(|\alpha| - \epsilon_C, 0) \geq 0$  und  $\beta > \frac{\beta}{2}$ .
- Falls  $\beta < |\alpha|$ , folgt die Aussage wegen  $\max(|\alpha| - \epsilon_C, 0) > \frac{|\alpha|}{2}$  und  $\beta \geq 0$ .
- Wenn  $\beta = |\alpha|$ , gilt  $\max(|\alpha| - \epsilon_C, 0) > 0$  und wiederum  $\beta > \frac{\beta}{2}$ .

Wir haben also gezeigt

$$P(|\alpha + \beta e_t| \leq \epsilon_C) \leq \frac{4\epsilon_C L}{\max(|\alpha|, \beta)}.$$

Mit der Definition  $L_{\inf} := \inf_{\theta \in \Theta} \max(|a^*(\theta)|, b^*(\theta))$ , wobei  $L_{\inf}$  nach Annahme A5 größer Null ist, erhält man abschließend

$$P(|\alpha + \beta e_t| \leq \epsilon_C) \leq \frac{4\epsilon_C L}{L_{\inf}}. \quad (3.7)$$

Falls  $\beta = 0$  ist (3.7) trivialerweise ebenfalls erfüllt, da  $|\alpha + \beta e_t| = |\alpha| > \epsilon_C$ . Wählt man jetzt  $\epsilon \in (0, \frac{L_{\inf}}{2})$ , setzt  $\epsilon_C = \frac{\epsilon}{1 + \frac{4L}{L_{\inf}}}$  und definiert  $M$  außerdem durch

$$P(|e_t| > M) < \epsilon_C,$$

dann ergibt sich

$$\begin{aligned} P(|a^*(\theta) + b^*(\theta)e_t| > \epsilon_C, |e_t| \leq M) &= 1 - P(\{|a^*(\theta) + b^*(\theta)e_t| \leq \epsilon_C\} \cup \{|e_t| > M\}) \\ &\geq 1 - \left( \frac{4L}{L_{\inf}} \cdot \frac{\epsilon}{1 + \frac{4L}{L_{\inf}}} + \frac{\epsilon}{1 + \frac{4L}{L_{\inf}}} \right) \\ &= 1 - \frac{(1 + \frac{4L}{L_{\inf}})\epsilon}{1 + \frac{4L}{L_{\inf}}} = 1 - \epsilon. \end{aligned} \quad (3.8)$$

### 3 Die Piggyback-Methode

Die gesuchte Abschätzung wurde damit gefunden:

$$\begin{aligned}
 P(\theta_1^* \in C | \theta_0^* = \theta) &\geq P(|a^*(\theta) + b^*(\theta)e_1| \geq \epsilon_C, \|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma) \\
 &\geq P(|a^*(\theta) + b^*(\theta)e_1| \geq \epsilon_C, |e_1| \leq M) \\
 &> 1 - \epsilon.
 \end{aligned} \tag{3.9}$$

In der ersten Zeile wurde dabei (3.6) verwendet.

ii) Die zweite Behauptung zeigen wir mit Induktion über  $p$ .

Induktionsanfang: Wenn  $p=1$ , dann ist der Zustandsraum von  $(X_t^*)_{t \geq 0}$  gleich  $\mathbb{R}$  und der Zustandsraum von  $(\theta_t^*)_{t \geq 0}$  gleich  $\Theta = \{-1, +1\}$ . Der Vektor

$$\theta_1^* = \left( \frac{a^*(\theta) + b^*(\theta)e_1}{|a^*(\theta) + b^*(\theta)e_1|} \right)$$

besteht dann nur aus einer Komponente. Für  $\theta_0^* = \theta$  mit  $\theta \in \Theta$  ist der Induktionsanfang deshalb durch (3.9) erfüllt.

Induktionsvoraussetzung: Aussage ii) gelte für alle  $k = 1, \dots, p$ .

Induktionsschritt:  $p \mapsto p + 1$ .

Zu zeigen:  $P(\theta_{p+1}^* \in C | \theta_0^* = \theta) > 1 - (p + 1)\epsilon, \forall \theta \in \Theta$ .

Um den Induktionsschritt durchführen zu können, muss man sich klar machen, dass bei Aussage ii) die Anzahl der Schritte  $p$  mit der Dimension des Zustandsraumes  $\mathbb{R}^p$  übereinstimmt. Zwecks Hervorhebung der Dimension, schreiben wir in diesem Teil des Beweises

$$\Theta_p := \{x \in \mathbb{R}_0^p, \|x\| = 1\} \text{ und } C_p := \{\theta = (\theta_1, \dots, \theta_p), |\theta_i| \geq \epsilon_C \gamma^{-i}, i = 1, \dots, p\}.$$

Nach diesem Beweis verzichten wir der Einfachheit halber wieder auf die Dimensionsindizes. Es sind nun zwei Prozesse zu unterscheiden.  $(\theta_{p+1,t}^*)_{t \geq 0}$  sei der homogene Prozess auf  $\Theta_{p+1}$  und  $(\theta_{p,s}^*)_{s \geq 0}$  sei der homogene Prozess auf  $\Theta_p$ . Wir wenden also die Induktionsvoraussetzung auf den  $p$ -dimensionalen Prozess  $\theta_{p,s}^*$  anwenden, um die Aussage für den  $p+1$ -dimensionalen Prozess  $\theta_{p+1,t}^*$  zu beweisen. Mit der Chapman-Kolmogorov-Gleichung erhält man

$$\begin{aligned}
 P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,0}^* = \theta) &\geq \inf_{\theta' \in A} P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,p}^* = \theta') \\
 &\quad \cdot P(\theta_{p+1,p}^* \in A | \theta_{p+1,0}^* = \theta), \quad A \in \mathcal{B}(\mathbb{R}^p).
 \end{aligned}$$

Wir fassen jetzt  $C_p$  als Teilmenge von  $\Theta_{p+1}$  auf, bei der die  $(p + 1)$ -te Komponente beliebig ist. Wenn wir  $|a^*(\theta) + b^*(\theta)e_t| \geq \epsilon_C$  unabhängig von der Dimension von  $\theta$  fordern, genügt es zu untersuchen, wann sich  $\theta_{p,s}^*$  in  $C_p$  befindet, da die  $(p + 1)$ -te Komponente von  $\theta_{p+1,t}^*$  für  $C_p$  unerheblich ist. Deshalb folgt mit  $A = C_p$  unter Verwendung der Induktionsvoraussetzung

$$\begin{aligned}
 P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,0}^* = \theta) &\geq \inf_{\theta' \in C_p} P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,p}^* = \theta') \cdot P(\theta_{p,p}^* \in C_p | \theta_{p,0}^* = \theta) \\
 &> \inf_{\theta' \in C_p} P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,p}^* = \theta') \cdot (1 - p\epsilon).
 \end{aligned}$$

### 3 Die Piggyback-Methode

Bleibt noch der linke Term abzuschätzen. Die abkürzende Schreibweise

$$w^*(\theta, u) := (a^*(\theta) + b^*(\theta)u, \theta_1, \dots, \theta_{p-1})$$

führen wir ein, um die weitere Analyse der verschiedenen Prozesse zu vereinfachen. Analog zu Teil i) erhält man dann

$$\begin{aligned} |(\theta_{p+1,p+1}^*)_1| &= \frac{|a^*(\theta_{p+1,p}^*) + b^*(\theta_{p+1,p}^*)e_{p+1}|}{\|w^*(\theta_{p+1,p}^*, e_{p+1})\|} \geq \epsilon_C \gamma^{-1} \quad \text{und} \\ |(\theta_{p+1,p+1}^*)_i| &= \frac{|(\theta_{p+1,p}^*)_{i-1}|}{\|w^*(\theta_{p+1,p}^*, e_{p+1})\|} \geq \frac{\epsilon_C \gamma^{-i+1}}{\gamma} = \epsilon_C \gamma^{-i}, \quad \text{für } i = 2, \dots, p+1, \end{aligned}$$

falls  $|a^*(\theta_p^*) + b^*(\theta_p^*)e_{p+1}| \geq \epsilon_C$  und  $e_{p+1} \leq M$  (vgl. (3.6)). Die Induktion ist an dieser Stelle vollständig, denn

$$\begin{aligned} P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,0}^* = \theta) & \\ &> \inf_{\theta' \in C_p} P(\theta_{p+1,p+1}^* \in C_{p+1} | \theta_{p+1,p}^* = \theta') \cdot (1 - p\epsilon) \\ &\geq P(|a^*(\theta_{p+1,p}^*) + b^*(\theta_{p+1,p}^*)e_{p+1}| \geq \epsilon_C, e_{p+1} \leq M) \cdot (1 - p\epsilon) \\ &> (1 - \epsilon) \cdot (1 - p\epsilon) > 1 - (p+1)\epsilon. \end{aligned}$$

In der letzten Zeile wurde dabei (3.8) verwendet. □

Das letzte Lemma hatte den Sinn eine kompakte Menge zu definieren, die von dem  $(\theta_t^*)_{t \geq 0}$  erreicht wird und die von jedem Zustand aus  $\Theta$  gleichmäßig erreichbar ist. Mit diesen Gegebenheiten können wir uns dem Nachweis der gleichmäßigen Ergodizität der Kette  $(\theta_t^*)_{t \geq 0}$  zuwenden.

**Satz 3.3.** *Seien Annahmen 1-5 vorausgesetzt. Dann ist  $(\theta_t^*)_{t \geq 0}$  eine gleichmäßig ergodische T-chain auf  $\Theta$ .*

Beweis: (vgl. [6], Theorem 3.1)

Nach Satz 3.1 iii) ist die Markov-Kette  $(\theta_t^*)_{t \geq 0}$  eingeschränkt auf  $\Theta_\times$  eine  $\lambda_\Theta$ -irreduzible, aperiodische T-chain.  $(\theta_t^*)_{t \geq 0}$  besitzt diese Eigenschaften aber auch auf ganz  $\Theta$ . Nach Start in  $\Theta$  hält sich  $(\theta_t^*)_{t \geq 0}$  für  $t \geq p$  wegen  $\max(|a^*(\theta)|, b^*(\theta)) > 0$  (vgl. Annahme 5)  $\lambda_\Theta$ -fast sicher in  $\Theta_\times$  auf. Aufgrund der  $\lambda_\Theta$ -Irreduzibilität des Prozesses auf  $\Theta_\times$  impliziert dies die  $\lambda_\Theta$ -Irreduzibilität auf  $\Theta$ .

$(\theta_t^*)_{t \geq 0}$  ist gleichmäßig ergodisch, wenn der Zustandsraum  $\Theta$  petite ist (vgl. 2.16 iv)). Satz 1.26 liefert, dass für eine  $\psi$ -irreduzible T-chain jede kompakte Menge petite ist. Demnach ist die kompakte Menge  $C$  aus Lemma 3.2 für  $(\theta_t^*)_{t \geq 0}$  eingeschränkt auf  $\Theta_\times$  petite (vgl. Satz 3.1). Man sieht leicht ein, dass  $C$  auch für den Prozess auf  $\Theta$  petite ist, da sich  $\Theta$  und  $\Theta_\times$  nur um eine  $\lambda_\Theta$ -Nullmenge unterscheiden. Lemma 3.2 ii) zeigt, dass  $C$  von  $\Theta$  aus gleichmäßig erreicht wird und somit selbst petite ist (vgl. Satz 1.21). Der Satz ist damit bewiesen. □

Die wichtigste Eigenschaft, die wir in ähnlicher Form für den Ausgangsprozess  $(X_t)_{t \geq 0}$  zeigen wollen, haben wir jetzt bereits für die Hilfskette  $(\theta_t^*)_{t \geq 0}$  bewiesen.

### 3.4.2 Die near-equilibrium-Gleichung

Die gleichmäßige Ergodizität spielt die entscheidende Rolle bei der Herleitung der near-equilibrium-Gleichung, der wir uns als nächstes widmen. Die near-equilibrium-Gleichung liefert später einen wichtigen Beitrag zum Beweis der V-geometrischen Ergodizität von  $(X_t)_{t \geq 0}$ . Es sei noch einmal an die Schreibweise  $w^*(x, u) := (a^*(x) + b^*(x)u, x_1, \dots, x_{p-1})$  erinnert.

**Satz 3.4.** *Seien Annahmen 1-6 vorausgesetzt.  $w^*(\theta, e_1)$  stelle den homogenen Prozess  $X_1^*$  zum Zeitpunkt 1 mit Start in  $\theta$  dar.  $\rho$  sei gegeben durch*

$$\rho := \exp \left( \int_{\Theta} \int_{\mathbb{R}} \log(\|w^*(\theta, u)\|) f(u) du \pi(d\theta) \right), \quad (3.10)$$

wobei  $\pi$  die stationäre Verteilung von  $(\theta_t^*)_{t \geq 0}$  bezeichne. Dann existiert für jedes  $\epsilon > 0$  eine beschränkte Funktion  $\nu : \Theta \rightarrow \mathbb{R}$ , so dass

$$\sup_{\theta \in \Theta} \left| E \left[ \nu(\theta_1^*) - \nu(\theta) + \log(\|w^*(\theta, e_1)\|) \mid \theta_0^* = \theta \right] - \log \rho \right| < \epsilon. \quad (3.11)$$

Beweis: (vgl. [6], Lemma 5.6)

Bevor wir die near-equilibrium-Gleichung beweisen, muss die Wohldefiniertheit von  $\rho$  sichergestellt werden. Wir zeigen, dass die Menge  $\{|\log(\|w^*(\theta, e_1)\|)|\}_{\theta \in \Theta}$  gleichmäßig integrierbar ist. Es ist nachzuweisen, dass

$$\lim_{c \nearrow \infty} \sup_{\theta \in \Theta} \int |\log(\|w^*(\theta, u)\|)| \mathbf{1}_{\{|\log(\|w^*(\theta, e_1)\|)| \geq c\}} f(u) du = 0.$$

Betrachten wir also das Integral

$$\begin{aligned} & \sup_{\theta \in \Theta} \int |\log(\|w^*(\theta, u)\|)| f(u) \mathbf{1}_{\{|\log(\|w^*(\theta, e_1)\|)| \geq c\}} du \\ & \leq \sup_{\theta \in \Theta} \int |\log(|a^*(\theta) + b^*(\theta)u| + 1)| f(u) \mathbf{1}_{\{|a^*(\theta) + b^*(\theta)u| \geq e^c\}} du \\ & \leq \sup_{\theta \in \Theta} \int |a^*(\theta) + b^*(\theta)u| f(u) \mathbf{1}_{\{|a^*(\theta) + b^*(\theta)u| \geq e^c\}} du \\ & \leq L \cdot L_{\text{sup}} \int \mathbf{1}_{\{|a^*(\theta) + b^*(\theta)u| \geq e^c\}} du. \end{aligned}$$

Die Wahrscheinlichkeit, dass  $|a^*(\theta) + b^*(\theta)e_1| \geq e^c$ , strebt aufgrund der Beschränktheit von  $\max(|a^*(\theta)|, b^*(\theta))$  für  $c \rightarrow \infty$  gegen Null.  $\{|\log(\|w^*(\theta, e_1)\|)|\}_{\theta \in \Theta}$  ist somit gleichmäßig integrierbar und der Erwartungswert von  $|\log(\|w^*(\theta, e_1)\|)|$  deshalb endlich.  $\rho$  ist daher als Integral über den Erwartungswert von  $\log(\|w^*(\theta, e_1)\|)$  bezüglich  $\pi$  wohldefiniert.

Um die near-equilibrium-Gleichung zu beweisen, definiert man die beschränkte Funktion

$$q(\theta) := E \log(\|w^*(\theta, e_1)\|).$$

### 3 Die Piggyback-Methode

An dieser Stelle führen wir die wichtige Testfunktion

$$\nu(\theta) = \sum_{t=0}^{T-1} E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right] \quad (3.12)$$

ein. Zu beachten ist, dass  $\nu$  als Summe von  $q(\theta_t^*)$  ebenfalls beschränkt ist. Betrachtet man  $\nu(\theta_1^*)$ , wobei  $\theta_1^* = \frac{w^*(\theta, e_1)}{\|w^*(\theta, e_1)\|}$ , erhält man

$$\nu(\theta_1^*) = \sum_{t=0}^{T-1} E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta_1^* \right] = \sum_{t=1}^T E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right].$$

$\nu(\theta_1^*)$  unterscheidet sich von  $\nu(\theta)$  nur um den ersten und letzten Summanden:

$$\begin{aligned} \nu(\theta_1^*) - \nu(\theta) &= \sum_{t=1}^T E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right] - \sum_{t=0}^{T-1} E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right] \\ &= E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] - E \left[ q(\theta_0^*) \middle| \theta_0^* = \theta \right]. \end{aligned}$$

Umstellen und Ausnutzen der Projektivität der bedingten Erwartung sowie der Unabhängigkeit von  $e_1$  und  $\theta$  ergibt

$$\begin{aligned} &E \left[ \nu(\theta_1^*) - \nu(\theta) + E \left[ q(\theta_0^*) \middle| \theta_0^* = \theta \right] \middle| \theta_0^* = \theta \right] = E \left[ E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] \middle| \theta_0^* = \theta \right] \\ \Leftrightarrow &E \left[ \nu(\theta_1^*) - \nu(\theta) + q(\theta_0^*) \middle| \theta_0^* = \theta \right] = E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] \\ \Leftrightarrow &E \left[ \nu(\theta_1^*) - \nu(\theta) + \log(\|w^*(\theta, e_1)\|) \middle| \theta_0^* = \theta \right] = E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right]. \end{aligned}$$

Es muss jetzt folgende Differenz abgeschätzt werden:

$$\left| E \left[ \nu(\theta_1^*) - \nu(\theta) + \log(\|w^*(\theta, e_1)\|) \middle| \theta_0^* = \theta \right] - \log \rho \right| = \left| E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] - \log \rho \right|.$$

Ruft man sich in Erinnerung, dass es sich bei  $\log \rho$  um ein Integral über die Funktion  $q(\theta)$  handelt, erkennt man

$$\left| E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] - \log \rho \right| = \left| \int q(y) P(\theta_T^* \in dy \mid \theta_0^* = \theta) - \int q(\theta) \pi(d\theta) \right|.$$

Aufgrund der gleichmäßigen Ergodizität von  $(\theta_t^*)_{t \geq 0}$  und der Beschränktheit von  $q$  existieren  $\tilde{R} < \infty$  und  $r < 1$ , so dass

$$\begin{aligned} &\left| \int_A q(y) P(\theta_T^* \in dy \mid \theta_0^* = \theta) - \int_A q(\theta) \pi(d\theta) \right| \leq \tilde{R} r^T \\ \Rightarrow &\left| E \left[ q(\theta_T^*) \middle| \theta_0^* = \theta \right] - \log \rho \right| \leq \tilde{R} r^T, \quad \forall T \geq 1, \theta \in \Theta, A \in \mathcal{B}(\mathbb{R}^p) \end{aligned}$$

### 3 Die Piggyback-Methode

(vgl. Satz 2.16 ii)). Wählt man  $T$  so groß, dass  $\tilde{R}r^T < \epsilon$ , erhält man die gewünschte Abschätzung

$$\left| E \left[ \nu(\theta_1^*) - \nu(\theta) + \log(\|w^*(\theta, e_1)\|) \middle| \theta_0^* = \theta \right] - \log \rho \right| < \epsilon.$$

□

#### 3.4.3 Die Stetigkeit von $\nu$

Als nächstes weisen wir die die Stetigkeit von  $\nu(\theta) = \sum_{t=0}^T E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$  auf einer Teilmenge von  $\Theta$  nach. Die Funktionen  $a^*$  und  $b^*$  sind selbst nur auf der Menge  $\Theta \setminus \bigcup_{j=1}^m H_j$  stetig. Deshalb kommt auch nur diese Menge für die Stetigkeit von  $\nu$  in Frage. Die folgende Definition geht aber noch weiter. Die Abbildung  $F : \Theta \rightarrow \Theta$  permutiere einen Vektor  $\theta$  durch die Vorschrift

$$F\theta = (\theta_2, \dots, \theta_p, \theta_1).$$

$F^k$  sei die  $k$ -fache Hintereinanderausführung von  $F$  und die inverse Abbildung sei durch  $F^{-1}\theta = (\theta_p, \theta_1, \dots, \theta_{p-1})$  gegeben. Dann definiere

$$\Theta_{\times\times} := \{\theta \in \Theta_{\times}, F^k\theta \notin H_j, j = 1, \dots, m; k = 0, \dots, p-1\}.$$

Die Menge schließt zum einen aus, dass sich  $\theta$  an einer der Unstetigkeitsstellen befindet, zum anderen wird nicht zugelassen, dass sich der Prozess  $(\theta_t^*)_{t \geq 0}$ , der sich durch Permutation von  $p-1$  Komponenten des vorherigen Zustands ergibt, mit fortschreitender Zeit einer Unstetigkeitsstelle annähert. Für die gewünschte Stetigkeitsaussage benötigt man ein Lemma, welches analog zu Lemma 3.2 formuliert wird. Die Modifikation besteht darin, dass  $\Theta_{\times}$  durch  $\Theta_{\times\times}$  ersetzt wird.

**Lemma 3.5.** *Seien Annahmen 1-6 vorausgesetzt. Für jedes  $\epsilon \in (0, 1)$  existiert eine kompakte Menge  $D \in \Theta_{\times\times}$ , so dass*

$$P(\theta_1^* \in D | \theta_0^* = \theta) > 1 - \epsilon \text{ für alle } \theta \in D. \tag{3.13}$$

Beweis: (vgl.[6], Lemma 5.4)

Der Beweis verläuft analog zu dem Beweis von Lemma 3.2. Man prüft unter welchen Bedingungen sich  $\theta_1^*$  in der kompakten Menge  $D$  befindet, wenn schon  $\theta_0^*$  in  $D$  ist und schätzt diese Wahrscheinlichkeit ab. Die Menge  $D$  wird allerdings etwas anders gewählt. In Lemma 3.2 wurde  $C$  so konstruiert, dass jede Komponente von  $\theta$  einen bestimmten Abstand zu den Achsenflächen hält. An  $D$  stellen wir zusätzlich die Forderung, dass der Abstand zu den Unstetigkeitsstellen einen gewissen Betrag nicht unterschreitet. Wählt man zu jeder homogenen Hyperfläche  $H_j$  einen Vektor  $h_j \in \mathbb{R}_0^p$ , der orthogonal auf  $H_j$  steht, dann ist  $\theta$  genau dann in  $\Theta_{\times\times}$ , wenn das Standardskalarprodukt  $|\langle h_j, F^k\theta \rangle| \neq 0$  ist, für alle  $j = 1, \dots, m; k = 1, \dots, p$ . Der Abstand des Vektors  $F^k\theta$  zur Fläche  $H_j$  wird durch den Betrag des Skalarproduktes mit dem entsprechenden Vektor  $h_j$  gemessen.

### 3 Die Piggyback-Methode

Aufgrund der Abbildung  $F$  ist die Definition  $(h)_{p+1} := (h)_1$  bzw.  $\theta_{p+1} := \theta_1$  sinnvoll, um das Skalarprodukt zwischen  $h$  und  $F^k\theta$  zu berechnen. Es gilt:

$$\langle h, F^k\theta \rangle = \sum_{i=1}^{i=p} (h)_i \cdot \theta_{k+i} = \sum_{i=1}^{i=p} (h)_{p-k+i} \cdot \theta_i$$

Die kompakte Menge  $D$  definieren wir durch

$$D := \{\theta \in C, |\langle h_j, F^k\theta \rangle| \geq \epsilon_D \gamma^{-k}, \text{ für alle } j = 1, \dots, m; k = 0, \dots, p-1\},$$

wobei  $C \in \Theta_\times$  die kompakte Menge und  $\gamma$  die Konstante aus Lemma 3.2 sind.

Es muss geprüft werden unter welchen Bedingungen  $|\langle h_j, F^k\theta_1^* \rangle| \geq \epsilon_D \gamma^{-k}$  gilt, wenn schon  $|\langle h_j, F^k\theta \rangle| \geq \epsilon_D \gamma^{-k}$  erfüllt ist. Hierzu möchte man eine Beziehung zwischen den beiden Skalarprodukten herstellen. Betrachtet man die Differenz

$$\theta_1^* - \frac{F^{-1}\theta}{\|w^*(\theta, e_1)\|} = \frac{(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})}{\|w^*(\theta, e_1)\|} - \frac{(\theta_p, \theta_1, \dots, \theta_{p-1})}{\|w^*(\theta, e_1)\|},$$

erkennt man, dass sich die Vektoren nur um die erste Komponente unterscheiden. Wendet man auf beide Vektoren die Abbildung  $F^k$  an, unterscheiden sie sich um die  $(p-k+1)$ -te Komponente. Durch Umstellen der Gleichung ergibt sich für das Skalarprodukt

$$|\langle h_j, F^k\theta_1^* \rangle| = \frac{|(h_j)_{p-k+1} \cdot (a^*(\theta) + b^*(\theta)e_1 - \theta_p) + \langle h_j, F^{k-1}\theta \rangle|}{\|w^*(\theta, e_1)\|}. \quad (3.14)$$

Nimmt man  $\|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma$  an, sind zwei Fälle zu unterscheiden:

1. Gilt  $(h_j)_{p-k+1} = 0$ , dann folgt

$$|\langle h_j, F^k\theta_1^* \rangle| \geq \frac{|\langle h_j, F^{k-1}\theta \rangle|}{\gamma} \geq \epsilon_D \gamma^{-k+1} \geq \epsilon_D \gamma^{-k}.$$

In diesem Fall erfüllt  $\theta_1^*$  die Bedingung für die Menge  $D$ .

2. Falls  $(h_j)_{p-k+1} \neq 0$ , betrachten wir die Wahrscheinlichkeit, dass

$$\frac{|(h_j)_{p-k+1} \cdot (a^*(\theta) + b^*(\theta)e_1 - \theta_p) + \langle h_j, F^{k-1}\theta \rangle|}{\gamma} \geq \epsilon_D \gamma^{-k}$$

gilt. Dies ist äquivalent zu

$$\left| (a^*(\theta) + b^*(\theta)e_1 - \theta_p) + \frac{\langle h_j, F^{k-1}\theta \rangle}{(h_j)_{p-k+1}} \right| \geq \frac{\epsilon_D \gamma^{-k+1}}{(h_j)_{p-k+1}}. \quad (3.15)$$

Um die Übersicht zu bewahren, definieren wir

$$L_{\min, h} := \min_{j, k} \{(h_j)_k, (h_j)_k \neq 0\} \quad \text{und} \\ u_{j, k} = \begin{cases} \theta_p - \frac{\langle h_j, F^{k-1}\theta \rangle}{(h_j)_{p+1-k}} & , \text{ wenn } (h_j)_{p+1-k} \neq 0 \\ 0 & , \text{ wenn } (h_j)_{p+1-k} = 0. \end{cases}$$



### 3 Die Piggyback-Methode

Als Ergebnis für den 2. Fall halten wir fest

$$P(|\langle h_j, F^k \theta_1^* \rangle| \geq \epsilon_D \gamma^{-k}) \geq P(|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \geq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}}).$$

Nach diesen Vorüberlegungen lässt sich die Wahrscheinlichkeit  $P(\theta_1^* \in D | \theta_0^* = \theta)$  abschätzen. Beachte, dass  $j$  bzw.  $k$  die Werte  $1, \dots, m$  bzw. die Werte  $1, \dots, p$  annehmen kann.

$$\begin{aligned} P(\theta_1^* \in D | \theta_0^* = \theta) &\geq P(\theta_1^* \in C, \|\theta_1^*\| \leq \gamma, |\langle h_j, F^k \theta \rangle| \geq \epsilon_D \gamma^{-k}, \forall j, k | \theta_0^* = \theta) \\ &\geq P(\theta_1^* \in C, \|\theta_1^*\| \leq \gamma, |a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \geq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}}, \forall j, k | \theta_0^* = \theta) \\ &\geq 1 - P(\{\theta_1^* \in C, \|\theta_1^*\| \leq \gamma\}^c | \theta_0^* = \theta) \\ &\quad - P(\{|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \geq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}}, \forall j, k\}^c) \\ &= 1 - P(\{\theta_1^* \in C, \|\theta_1^*\| \leq \gamma\}^c | \theta_0^* = \theta) \\ &\quad - P(|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \leq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}} \text{ für irgendein } j, k) \quad (3.16) \end{aligned}$$

Nach Lemma 3.2 existiert ein  $\epsilon$ , so dass

$$\begin{aligned} P(\theta_1^* \in C, \|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma | \theta_0^* = \theta) &> 1 - \frac{\epsilon}{2} \\ \Rightarrow P(\{\theta_1^* \in C, \|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma\}^c | \theta_0^* = \theta) &< \frac{\epsilon}{2}. \end{aligned}$$

Der Term (3.16) wird analog zu (3.7) in Lemma 3.2 abgeschätzt:

$$P(|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \leq c) \leq \frac{L \cdot 2c}{\inf_{\theta \in C} b^*(\theta)}.$$

Für  $c = \frac{\gamma^{-k+1}}{L_{\min,h}} \epsilon_D$  ergibt sich damit

$$\begin{aligned} P(|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \leq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}} \text{ für irgendein } j, k) & \\ &= P(\bigcup_{j,k} \{|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \leq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}}\}) \\ &\leq \sum_{j=1}^p \sum_{k=1}^m P(|a^*(\theta) + b^*(\theta)e_1 - u_{j,k}| \leq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min,h}}) \\ &\leq mp \frac{L}{\inf_{\theta \in C} b^*(\theta)} \frac{2\epsilon_D \gamma^{1-k}}{L_{\min,h}} \\ &\leq mp \frac{L}{\inf_{\theta \in C} b^*(\theta)} \frac{2\epsilon_D \gamma^1}{L_{\min,h}}, \end{aligned}$$

### 3 Die Piggyback-Methode

da  $\gamma \geq 1$  (vgl. (3.5)). Wählen wir  $\epsilon_D = \frac{L_{\min, h} \cdot \inf_{\theta \in C} b^*(\theta)}{4mp\gamma L} \cdot \epsilon$ , folgt schließlich die Behauptung:

$$\begin{aligned} P(\theta_1^* \in D | \theta_0^* = \theta) &\geq 1 - P(\{\theta_1^* \in C, \|(a^*(\theta) + b^*(\theta)e_1, \theta_1, \dots, \theta_{p-1})\| \leq \gamma\}^c | \theta_0^* = \theta) \\ &\quad - P(\|(a^*(\theta) + b^*(\theta)e_1 - u_{j,k})\| \geq \frac{\epsilon_D \gamma^{-k+1}}{L_{\min}} \text{ für irgendein } j, k | \theta_0^* = \theta) \\ &\geq 1 - \frac{\epsilon}{2} - mp \frac{2L}{\inf_{\theta \in C} b^*(\theta)} \frac{\epsilon_D \gamma}{L_{\min, h}} \\ &= 1 - \epsilon. \end{aligned}$$

□

Die Aussage des letzten Lemmas benötigen wir, um die Stetigkeit von

$$\nu(\theta) = \sum_{t=0}^{T-1} E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$$

zu beweisen. Entscheidend für die Stetigkeit von  $\nu$  ist die Stetigkeit von  $E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$ .

**Satz 3.6.**  $\nu(\theta) = \sum_{t=0}^{T-1} E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$  ist stetig auf  $\Theta_{\times \times}$ .

Beweis: (vgl. [6], Lemma 5.6)

Zunächst bemerken wir, dass  $q(\theta) = E \log(\|w^*(\theta, e_1)\|)$  beschränkt ist (vgl. Satz 3.4). Außerdem ist  $q(\theta)$  auf  $\Theta_{\times \times}$  stetig, da  $a^*(\theta)$  und  $b^*(\theta)$  auf  $\Theta_{\times \times}$  stetig sind. Wir zeigen:

$E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$  ist für alle  $t \geq 1$  stetig auf  $\Theta_{\times \times}$ , falls  $q(\theta)$  beschränkt auf  $\Theta$  und stetig auf  $\Theta_{\times \times}$  ist.  $E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta \right]$  ist stetig, wenn

$$\begin{aligned} &\forall \epsilon > 0, \exists \delta > 0, \text{ so dass für } \|\theta' - \theta''\| < \delta \\ &\Rightarrow \left| E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta' \right] - E \left[ q(\theta_t^*) \middle| \theta_0^* = \theta'' \right] \right| < \epsilon. \end{aligned}$$

Abkürzend schreiben wir  $\theta_{t, \theta'}^*$  für den in  $\theta'$  gestarteten Prozess. Es gilt:

$$\begin{aligned} \left| E \left[ q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*) \right] \right| &\leq \left| E \left[ (q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*)) \mathbf{1}_{\{|q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*)| < \epsilon\}} \right] \right| \\ &\quad + \left| E \left[ (q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*)) \mathbf{1}_{\{|q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*)| \geq \epsilon\}} \right] \right| \\ &\leq \epsilon + 2 \sup_{\theta \in \Theta} |q(\theta_{t, \theta})| \cdot P(|q(\theta_{t, \theta'}^*) - q(\theta_{t, \theta''}^*)| \geq \epsilon). \end{aligned}$$

### 3 Die Piggyback-Methode

Bleibt der Term  $P(|q(\theta_{t,\theta'}^*) - q(\theta_{t,\theta''}^*)| \geq \epsilon)$  zu bearbeiten. Hierfür benutzen wir, dass  $q$  stetig ist. Die weitere Abschätzung erfolgt mit Hilfe von Lemma 3.5.  $q$  ist auf der kompakten Menge  $D \subset \Theta_{\times \times}$  sogar gleichmäßig stetig, deshalb gilt

$$\exists \delta_0, \text{ so dass } \|\theta' - \theta''\| < \delta_0, \theta', \theta'' \in D \Rightarrow |q(\theta') - q(\theta'')| < \epsilon. \quad (3.17)$$

Die Funktionenmenge  $\left\{ \frac{w^*(\theta', u)}{\|w^*(\theta', u)\|} = (a^*(\theta') + b^*(\theta')u, \theta'_1, \dots, \theta'_{p-1}), u \leq M \right\}$  ist gleichmäßig gleichgradig stetig auf  $D$ , das heißt für  $u \leq M$  folgt

$$\|\theta' - \theta''\| < \delta_t, \theta', \theta'' \in D \Rightarrow \left\| \frac{w^*(\theta', u)}{\|w^*(\theta', u)\|} - \frac{w^*(\theta'', u)}{\|w^*(\theta'', u)\|} \right\| < \delta_{t-1}.$$

Beachte, dass es sich bei  $\frac{w^*(\theta', e_1)}{\|w^*(\theta', e_1)\|}$  um den normierten, homogenen Prozess zum Zeitpunkt 1, gestartet in  $\theta$ , handelt. Für die Prozesse zum Zeitpunkt  $k$  erhält man deshalb

$$\|\theta_{k,\theta'}^* - \theta_{k,\theta''}^*\| < \delta_{t-k},$$

falls  $\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}$ ,  $\theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D$  und  $|e_k| \leq M$ . Damit ergibt sich

$$\begin{aligned} & P(\|\theta_{k,\theta'}^* - \theta_{k,\theta''}^*\| < \delta_{t-k}, \theta_{k,\theta'}^*, \theta_{k,\theta''}^* \in D) \\ & \geq P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| \leq M, \theta_{k,\theta'}^*, \theta_{k,\theta''}^* \in D) \\ & = P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| \leq M) \\ & \quad - P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| \leq M, \theta_{k,\theta'}^*, \theta_{k,\theta''}^* \notin D) \\ & \geq P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| \leq M) \\ & \quad - P(\theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, \theta_{k,\theta'}^*, \theta_{k,\theta''}^* \notin D) \end{aligned}$$

Die Wahrscheinlichkeit, dass sich der Prozess nicht in  $D$  befindet, wenn er sich einen Schritt vorher in  $D$  aufgehalten hat, ist durch (3.13) in Lemma 3.5 abgeschätzt. Anwendung von (3.13) auf die obige Rechnung liefert

$$\begin{aligned} & P(\|\theta_{k,\theta'}^* - \theta_{k,\theta''}^*\| < \delta_{t-k}, \theta_{k,\theta'}^*, \theta_{k,\theta''}^* \in D) \\ & > P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| \leq M) - \epsilon \\ & = P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D) \\ & \quad - P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D, |e_k| > M) - \epsilon \\ & \geq P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D) \\ & \quad - P(|e_k| > M) - \epsilon \\ & > P(\|\theta_{k-1,\theta'}^* - \theta_{k-1,\theta''}^*\| < \delta_{t-k+1}, \theta_{k-1,\theta'}^*, \theta_{k-1,\theta''}^* \in D) - 2\epsilon, \end{aligned}$$

wobei  $M$  so gewählt wurde, dass  $P(|e_k| > M) < \epsilon$  erfüllt ist (vgl. Lemma 3.2). Die Ungleichung gilt für alle  $k = 1, \dots, t$ . Die Wahrscheinlichkeit für den Fall  $k=t$  lässt sich

### 3 Die Piggyback-Methode

nun rekursiv berechnen. Hierfür setzen wir  $\|\theta' - \theta''\| < \delta_t$ ,  $\theta', \theta'' \in D$  voraus.

$$\begin{aligned}
 & P(\|\theta_{t,\theta'}^* - \theta_{t,\theta''}^*\| < \delta_0, \theta_{t,\theta'}^*, \theta_{t,\theta''}^* \in D) \\
 & > P(\|\theta_{t-1,\theta'}^* - \theta_{t-1,\theta''}^*\| < \delta_1, \theta_{t-1,\theta'}^*, \theta_{t-1,\theta''}^* \in D) - 2\epsilon \\
 & > P(\|\theta_{t-2,\theta'}^* - \theta_{t-2,\theta''}^*\| < \delta_2, \theta_{t-2,\theta'}^*, \theta_{t-2,\theta''}^* \in D) - 2 \cdot 2\epsilon \\
 & > \dots > P(\|\theta' - \theta''\| < \delta_t, \theta', \theta'' \in D) - t \cdot 2\epsilon \\
 & = 1 - t \cdot 2\epsilon.
 \end{aligned}$$

Aufgrund der gleichmäßigen Stetigkeit von  $q$  ergibt sich mit (3.17)

$$\begin{aligned}
 P(|q(\theta_{t,\theta'}^*) - q(\theta_{t,\theta''}^*)| < \epsilon) & \geq P(\|\theta_{t,\theta'}^* - \theta_{t,\theta''}^*\| < \delta_0, \theta_{t,\theta'}^*, \theta_{t,\theta''}^* \in D) \\
 & > 1 - t \cdot 2\epsilon.
 \end{aligned}$$

Die Abschätzung ist damit vollständig, es folgt nämlich

$$\begin{aligned}
 \left| E \left[ q(\theta_{t,\theta'}^*) - q(\theta_{t,\theta''}^*) \right] \right| & \leq \epsilon + 2 \sup |q(\theta_t^*)| P(|q(\theta_{t,\theta'}^*) - q(\theta_{t,\theta''}^*)| \geq \epsilon) \\
 & < \epsilon + 2 \sup |q(\theta_t^*)| \cdot t \cdot 2\epsilon.
 \end{aligned}$$

Wählt man jetzt  $\epsilon > 0$ , so dass  $\epsilon \cdot (1 + 4t \sup q(\theta_t^*)) < \tilde{\epsilon}$  und  $\delta = \delta_t$ . Die Stetigkeit der bedingten Erwartung ist dann wegen

$$\|\theta' - \theta''\| < \delta \quad \Rightarrow \quad \left| E \left[ q(\theta_{t,\theta'}^*) - q(\theta_{t,\theta''}^*) \right] \right| < \tilde{\epsilon}$$

bewiesen.  $\nu$  ist als Summe dieser bedingten Erwartungen stetig. □

#### 3.4.4 Der Übergang vom homogenen Prozess $(\theta_t^*)_{t \geq 0}$ zu $(\theta_t)_{t \geq 0}$

Als nächstes untersuchen wir, wie sich die Prozesse unterscheiden, wenn man einmal die homogenen Funktionen  $a^*$ ,  $b^*$  und einmal die Funktionen  $a$ ,  $b$  auf den Zustand  $X_{t-1}$  anwendet. Genauer definieren wir

$$\tilde{\theta}_t := \frac{w(X_{t-1}, e_t)}{\|w(X_{t-1}, e_t)\|} \quad \text{und} \quad \tilde{\theta}_t^* := \frac{w^*(X_{t-1}, e_t)}{\|w^*(X_{t-1}, e_t)\|}, \tag{3.18}$$

wobei  $w(x, e_1) = (a(x) + b(x)e_1, x_1, \dots, x_{p-1})$  analog zu  $w^*(x, e_1)$  gegeben sei.  $(\tilde{\theta}_t)_{t \geq 0}$  ist dann der gewöhnliche, normierte Prozess, angewendet auf den Zustand  $X_{t-1}$ , während  $(\tilde{\theta}_t^*)_{t \geq 0}$  der homogene, normierte Prozess, angewendet auf den Zustand  $X_{t-1}$ , bezeichnet. Wir zeigen in diesem Abschnitt, dass sich beide Prozesse mit hoher Wahrscheinlichkeit in derselben kompakten Menge  $\tilde{D} \subset \Theta_{\times \times}$  aufhalten. Später wird bewiesen, dass sich die Prozesse für große "Anfangszustände"  $X_{t-1}$  annähern. Dies ist auf Annahme 4 zurückzuführen, da sich die Funktionen  $a$  und  $b$  von den homogenen Funktionen  $a^*$  und  $b^*$  nur um eine Funktion  $a_0$  bzw.  $b_0$  unterscheiden, die weniger als linear anwächst. Beachte hierfür

$$w(X_{t-1}, e_t) = w^*(X_{t-1}, e_t) + (a_0(X_{t-1}) + b_0(X_{t-1})e_t, 0, \dots, 0).$$

### 3 Die Piggyback-Methode

Um zu beweisen, dass sich  $(\tilde{\theta}_t)_{t \geq 0}$  und  $(\tilde{\theta}_t^*)_{t \geq 0}$  in einer kompakten Teilmenge von  $\Theta_{\times \times}$  aufhalten, benötigen wir ein Hilfslemma, welches besagt, dass sich beide Prozesse zunächst in einer kompakten Teilmenge von  $\Theta_{\times}$  aufhalten.

**Lemma 3.7.** *Seien Annahmen 1-5 vorausgesetzt und seien  $\tilde{\theta}_t$  und  $\tilde{\theta}_t^*$  wie in (3.18) definiert. Für jedes  $\epsilon \in (0, 1)$  existieren  $\gamma > 1$ ,  $\eta \in (0, 1)$ ,  $\tilde{M} < \infty$  und eine kompakte Menge  $\tilde{C} \subset \Theta_{\times}$ , so dass*

$$P(\tilde{\theta}_t \in \tilde{C}, \tilde{\theta}_t^* \in \tilde{C}, \eta^t \leq \frac{\|X_t\|}{\|x\|} \leq (\gamma + \eta)^t |X_0 = x) > (1 - \epsilon)^t,$$

falls  $\|x\| > \tilde{M}\eta^{1-t}$  und zusätzlich  $\frac{x}{\|x\|} \in \tilde{C}$  oder  $t \geq p$  gilt.

Beweis: (vgl. [6], Lemma 5.7)

Zunächst konstruieren wir die kompakte Menge  $\tilde{C} \subset \Theta_{\times}$ . Dazu definieren wir eine Menge, die noch von  $x$  und  $t$  abhängt

$$\tilde{C}_{x,t} = \{\theta : |\theta_i| \geq \eta(\gamma + \eta)^{-i}, i \leq \min(t, p); |\theta_i| \geq \frac{|x_{i-t}|}{\|x\|}(\gamma + \eta)^{-1}, t < i \leq p\}.$$

Falls  $t = p$ , hängt  $\tilde{C} := \tilde{C}_{x,p}$  nicht von  $x$  und  $p$  ab, da die zweite Bedingung wegfällt. In diesem Fall setzen wir  $\tilde{C} := \tilde{C}_{x,p}$ . Für  $\frac{x}{\|x\|} \in \tilde{C}$  oder  $t \geq p$  gilt ebenfalls  $\tilde{C}_{x,t} = \tilde{C}$ . Aus diesem Grund genügt es für  $t \geq 1$  und  $\|x\| > \tilde{M}\eta^{1-t}$

$$P(\tilde{\theta}_t \in \tilde{C}_{x,t}, \tilde{\theta}_t^* \in \tilde{C}_{x,t}, \eta^t \leq \frac{\|X_t\|}{\|x\|} \leq (\gamma + \eta)^t |X_0 = x) > (1 - \epsilon)^t \quad (3.19)$$

zu zeigen. Gilt nämlich die Abschätzung für alle  $t$ , dann gilt sie ebenso für  $t \geq p$ . In dem Fall ist aber  $\tilde{C}_{x,t} = \tilde{C}$ . Für den Induktionsanfang zeigen wir

$$P(\tilde{\theta}_1 \in \tilde{C}_{x,1}, \tilde{\theta}_1^* \in \tilde{C}_{x,1}, \eta \leq \frac{\|X_1\|}{\|x\|} \leq (\gamma + \eta) |X_0 = x) > 1 - \epsilon, \quad (3.20)$$

falls  $\|x\| > \tilde{M}$ . Um zu sehen, wann  $\tilde{\theta}_1$  und  $\tilde{\theta}_1^*$  in  $C_{x,1}$  liegen, bestimmen wir für jede Komponente das Minimum der beiden Prozesse. Behauptung:

$$\min(|(\tilde{\theta}_1)_i|, |(\tilde{\theta}_1^*)_i|) \geq \begin{cases} \eta(\gamma + \eta)^{-1}, & \text{wenn } i = 1 \\ \frac{|x_{i-1}|}{\|x\|}(\gamma + \eta)^{-1}, & \text{wenn } i = 2, \dots, p. \end{cases}$$

Für die erste Komponente gilt

$$\min(|(\tilde{\theta}_1)_1|, |(\tilde{\theta}_1^*)_1|) = \min\left(\frac{a(x) + b(x)e_1}{\|w(x, e_1)\|}, \frac{a^*(x) + b^*(x)e_1}{\|w^*(x, e_1)\|}\right).$$

Wegen  $a_0(x) = o(\|x\|)$  und  $b_0(x) = o(\|x\|)$  existiert ein  $\tilde{M} < \infty$  mit

$$\frac{|a_0(x) + b_0(x)u|}{\|x\|} < \eta := \frac{\epsilon_C}{2}, \quad \forall \|x\| > \tilde{M}, |u| \leq M, \quad (3.21)$$

### 3 Die Piggyback-Methode

wobei  $M$  und  $\epsilon_C$  wie in Lemma 3.2 definiert seien. Setzt man  $|u| < M$ ,  $\|x\| > \tilde{M}$  und  $|a^*(x) + b^*(x)u| \geq \epsilon_C \|x\| = 2\eta \|x\|$  voraus, folgt deshalb

$$\begin{aligned}
 \eta \|x\| &\leq |a^*(x) + b^*(x)u + a_0(x) + b_0(x)u| \\
 &= |a(x) + b(x)u| \\
 &< \|(a(x) + b(x)u, x_1, \dots, x_{p-1})\| \\
 &\leq |a(x) + b(x)u| + \|x\| \\
 &\leq (\gamma + \eta) \|x\|.
 \end{aligned} \tag{3.22}$$

Dabei wurde  $|a^*(x) + b^*(x)u| + \|x\| \leq \gamma \|x\|$  und  $|a_0(x) + b_0(x)u| \leq \eta \|x\|$  in der letzten Zeile angewendet (vgl. Lemma 3.2 und (3.21)). Für den homogenen Prozess gilt die Abschätzung

$$2\eta \|x\| \leq \|(a^*(x) + b^*(x)u, x_1, \dots, x_p)\| \leq \gamma \|x\|. \tag{3.23}$$

Wendet man (3.22) und (3.23) auf die ersten Komponenten an, ergibt sich

$$\begin{aligned}
 \min(|(\tilde{\theta}_1)_1|, |(\tilde{\theta}_1)_1^*|) &= \min\left(\frac{|a(x) + b(x)u|}{\|w(x, u)\|}, \frac{|a^*(x) + b^*(x)u|}{\|w^*(x, u)\|}\right) \\
 &\geq \min\left(\frac{\eta \|x\|}{(\gamma + \eta) \|x\|}, \frac{2\eta \|x\|}{\gamma \|x\|}\right) = \frac{\eta}{\gamma + \eta}.
 \end{aligned}$$

Um das Minimum der übrigen Komponenten zu berechnen, benutzen wir ebenfalls die Resultate aus (3.22) und (3.23). Man erhält

$$\begin{aligned}
 \min(|(\tilde{\theta}_1)_i|, |(\tilde{\theta}_1)_i^*|) &= \min\left(\frac{|x_{i-1}|}{\|w(x, u)\|}, \frac{|x_{i-1}|}{\|w^*(x, u)\|}\right) \\
 &\geq \min\left(\frac{|x_{i-1}|}{(\gamma + \eta) \|x\|}, \frac{|x_{i-1}|}{\gamma \|x\|}\right) = \frac{|x_{i-1}|}{\|x\|} (\gamma + \eta)^{-1}.
 \end{aligned}$$

Damit wurde die Behauptung für das Minimum gezeigt. Der Induktionsanfang (3.20) ist jetzt wegen

$$\begin{aligned}
 P(\tilde{\theta}_1 \in \tilde{C}_{x,1}, \tilde{\theta}_1^* \in \tilde{C}_{x,1}, \eta \leq \frac{\|X_1\|}{\|x\|} \leq (\gamma + \eta) |X_0 = x|) \\
 \geq P(|e_1| \leq M, |a^*\left(\frac{x}{\|x\|}\right) + b^*\left(\frac{x}{\|x\|}\right)e_1| > \epsilon_C) \\
 > 1 - \epsilon
 \end{aligned}$$

erfüllt (vgl. (3.9) aus Lemma 3.2). Für den Nachweis des Induktionsschrittes nehme man an, dass (3.19) für alle  $1 \leq k \leq t$  erfüllt sei. Schreibt man sich die einzelnen Komponenten der Prozesse von  $\tilde{\theta}_k$  und  $\tilde{\theta}_k^*$  auf, um die Bedingungen der Menge  $\tilde{C}_{x,t}$  prüfen, erkennt man:

$$\tilde{\theta}_{k+1}, \tilde{\theta}_{k+1}^* \in \tilde{C}_{X_k,1} \text{ und } \tilde{\theta}_k \in \tilde{C}_{x,k} \Rightarrow \tilde{\theta}_{k+1}, \tilde{\theta}_{k+1}^* \in \tilde{C}_{x,k+1}.$$

### 3 Die Piggyback-Methode

Außerdem gilt

$$\begin{aligned} \eta &\leq \frac{\|X_{k+1}\|}{\|X_k\|} \leq (\gamma + \eta) \quad \text{und} \quad \eta^k \leq \frac{\|X_k\|}{\|x\|} \leq (\gamma + \eta)^k \\ \Rightarrow \quad \eta^{k+1} &\leq \frac{\|X_{k+1}\|}{\|x\|} \leq (\gamma + \eta)^{k+1}, \end{aligned}$$

wegen  $\frac{\|X_{k+1}\|}{\|x\|} = \frac{\|X_{k+1}\|}{\|X_k\|} \frac{\|X_k\|}{\|x\|}$ . Nimmt man weiterhin  $\|x\| > \frac{\tilde{M}}{\eta^t}$  an, lässt sich der Induktionsschritt leicht folgern:

$$\begin{aligned} P(\tilde{\theta}_{k+1}, \tilde{\theta}_{k+1}^* \in \tilde{C}_{x,k+1}, \eta^{k+1} \leq \frac{\|X_{k+1}\|}{\|x\|} \leq (\gamma + \eta)^{k+1} | X_0 = x) \\ \geq P(\tilde{\theta}_{k+1}, \tilde{\theta}_{k+1}^* \in \tilde{C}_{X_k,1}, \eta \leq \frac{\|X_{k+1}\|}{\|X_k\|} \leq (\gamma + \eta), \\ \tilde{\theta}_k \in \tilde{C}_{x,k}, \eta^k \leq \frac{\|X_k\|}{\|x\|} \leq (\gamma + \eta)^k | X_0 = x) \\ \geq \inf_{\substack{y \in \tilde{C}_{x,k} \\ \|y\| > \frac{\tilde{M}}{\eta^t}}} P(\tilde{\theta}_{k+1}, \tilde{\theta}_{k+1}^* \in \tilde{C}_{X_k,1}, \eta \leq \frac{\|X_{k+1}\|}{\|X_k\|} \leq (\gamma + \eta) | X_k = y) \\ \cdot P(\tilde{\theta}_k \in \tilde{C}_{x,k}, \eta^k \leq \frac{\|X_k\|}{\|x\|} \leq (\gamma + \eta)^k | X_0 = x) \\ > (1 - \epsilon)(1 - \epsilon)^k = (1 - \epsilon)^{k+1}. \end{aligned}$$

In der letzten Zeile wurde die Induktionsvoraussetzung ausgenutzt. □

**Lemma 3.8.** *Seien Annahmen 1-6 vorausgesetzt. Für jedes  $\epsilon \in (0, 1)$  existieren dann  $\gamma > 1$ ,  $\eta \in (0, 1)$ ,  $\tilde{M} < \infty$  sowie eine kompakte Menge  $\tilde{D} \subset \Theta_{\times \times}$ , so dass*

$$P(\tilde{\theta}_t \in \tilde{D}, \tilde{\theta}_t^* \in \tilde{D}, \eta^t \leq \frac{\|X_t\|}{\|x\|} \leq (\gamma + \eta)^t | X_0 = x) > (1 - \epsilon)^t,$$

*falls  $\|x\| > \tilde{M}\eta^{1-t}$  und zusätzlich  $\frac{x}{\|x\|} \in \tilde{D}$  oder  $t \geq 2p$  gilt.*

Beweis: (vgl. [6], Lemma 5.8)

Dieses Lemma verhält sich zu Lemma 3.7 genauso wie Lemma 3.5 zu Lemma 3.2. Deswegen wird hier nur kurz vorgestellt, wie der Beweis aus 3.7 modifiziert werden muss. Die kompakte Menge  $\tilde{D}_{t,p}$  sei definiert durch

$$\begin{aligned} \tilde{D}_{t,p} = \{ \theta \in \tilde{C}, |\langle h_j, F^k(\tilde{\theta}_t) \rangle| \geq \epsilon_D (\eta + \gamma)^{-k}, k \leq \min(t, p), \\ |\langle h_j, F^k(\tilde{\theta}_t) \rangle| \geq |\langle h_j, F^k(\frac{x_{i-t}}{\|x\|}) \rangle| \cdot \epsilon_D (\eta + \gamma)^{-1}, t < k \leq p \}. \end{aligned}$$

Dann gilt wiederum  $\tilde{D}_t = \tilde{D}$  für  $t \geq p$  bzw.  $\frac{x}{\|x\|} \in \tilde{D}$ . Es muss gezeigt werden, dass man bei Start in  $\theta_0 \in \tilde{C}$  in  $p$  Schritten die Menge  $\tilde{D}$  erreichen kann. Der Induktionsanfang

$$P(\tilde{\theta}_1, \tilde{\theta}_1^* \in \tilde{D}_1, \eta \leq \frac{\|X_1\|}{\|x\|} \leq (\gamma + \eta) | X_0 = x) > (1 - \epsilon)$$

### 3 Die Piggyback-Methode

ist durch (3.20) und einer analogen vorausgehenden Rechnung gewährleistet. Der Induktionsschritt ist möglich, da

$$\tilde{\theta}_t, \tilde{\theta}_t^* \in \tilde{D}_t \text{ und } \tilde{\theta}_1, \tilde{\theta}_1^* \in \tilde{D}_1 \Rightarrow \tilde{\theta}_{t+1}, \tilde{\theta}_{t+1}^* \in \tilde{D}_{t+1}.$$

□

Das letzte Lemma stellt bereits einen Bezug zwischen dem normierten  $(\theta_t)_{t \geq 0}$  und dem homogenen, normierten Prozess  $(\theta_t)_{t \geq 0}^*$  her, indem es zeigt, dass sich beide Prozesse für große Startwerte  $x$  und  $t \geq 2p$  mit hoher Wahrscheinlichkeit in derselben kompakten Menge aufhalten. Wie schon erwähnt lässt sich mit Annahme 4 folgern, dass sich beide Prozesse zum Zeitpunkt 1 für große Startwerte  $x$  annähern. Dafür benötigen wir eine asymptotische Aussage für  $\|X_1\| = \|w(x, e_1)\|$ . Für  $\|x\| \rightarrow \infty$  mit  $\frac{x}{\|x\|} \in \tilde{D}$  streben  $\|X_1\| = \|w(x, e_1)\|$  und  $\|X_1^*\| = \|w^*(x, e_1)\|$  P-f.s. gegen unendlich. Es gilt nämlich

$$\begin{aligned} \liminf_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{\|w(x, e_1)\|}{\|x\|} &= \liminf_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{\|w^*(x, e_1) + (a_0(x) + b_0(x)e_1, 0, \dots, 0)\|}{\|x\|} \\ &= \liminf_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{\|w^*(x, e_1)\|}{\|x\|} \\ &= \inf_{\theta \in \tilde{D}} \|w^*(\theta, e_1)\| > 0 \text{ P-f.s. .} \end{aligned} \tag{3.24}$$

Also ist  $\liminf_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{\|w(x, e_1)\|}{\|x\|} > 0$  P-f.s. und  $\|X_1\| = \|w(x, e_1)\|$  strebt P-f.s. gegen unendlich.

Eigentlich ist diese Aussage im nächsten Kapitel anzusiedeln, da erst dort der Prozess  $(X_t)_{t \geq 0}$  betrachtet wird. Sie wird aber in der nächsten Bemerkung benötigt.

**Bemerkung 3.9.** *Es gilt:*

$$\lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left\| \tilde{\theta}_1 - \tilde{\theta}_1^* \right\| = 0 \text{ P-f.s. .}$$



### 3 Die Piggyback-Methode

Beweis:

$$\begin{aligned}
\lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left\| \tilde{\theta}_1 - \tilde{\theta}_1^* \right\| &= \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left\| \frac{w(x, e_1)}{\|w(x, e_1)\|} - \frac{w^*(x, e_1)}{\|w^*(x, e_1)\|} \right\| \\
&= \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{1}{\|w^*(x, e_1)\|} \left\| \frac{\|w^*(x, e_1)\|}{\|w(x, e_1)\|} w(x, e_1) - w^*(x, e_1) \right\| \\
&\leq \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{1}{\|w^*(x, e_1)\|} \left\| \left( \frac{\|w^*(x, e_1)\|}{\|w(x, e_1)\|} - 1 \right) w^*(x, e_1) \right\| \\
&\quad + \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \frac{1}{\|w^*(y, e_1)\|} \left\| \frac{\|w^*(x, e_1)\|}{\|w(x, e_1)\|} (a_0(x) + b_0(x)e_1) \right\| \\
&= \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left\| \left( \frac{\|w^*(x, e_1)\|}{\|w(x, e_1)\|} - 1 \right) \right\| + \lim_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left\| \frac{|a_0(x) + b_0(x)e_1|}{\|w(x, e_1)\|} \right\| \\
&= 0,
\end{aligned}$$

da  $\|w(x, e_1)\| \xrightarrow{\|x\| \rightarrow \infty} \infty$  wegen (3.24) und  $\frac{\|w^*(x, e_1)\|}{\|w(x, e_1)\|} \xrightarrow{\|x\| \rightarrow \infty} 1$ . □

### 3.5 Der Prozess $(X_t)_{t \geq 0}$

In diesem Abschnitt wird der ursprüngliche Prozess  $(X_t)_{t \geq 0}$  untersucht. Bevor wir allerdings die Stabilität dieses Prozesses nachweisen, benötigen wir ein Lemma, um die Voraussetzungen von Lemma 2.17 zu verifizieren.

**Lemma 3.10.** *Seien Annahmen 1-5 vorausgesetzt. Dann ist*

$$\sup_{x \in \mathbb{R}^p} E \left[ \left( \frac{1 + \|w(x, e_1)\|}{1 + \|x\|} \right)^s \right] < \infty, \quad \forall s \in (-1, r_0]. \quad (3.25)$$

Beweis: (vgl. [6], Lemma 5.9)

Wir zeigen die Aussage nur für  $s \in [0, r_0]$ . Für negative Exponenten folgt die Aussage mit einer ähnlichen Abschätzung, die schon für (3.7) aus Lemma 3.2 durchgeführt wurde. Sei also  $s \in [0, r_0]$ . Eine Abschätzung des Terms im Erwartungswert ergibt

$$\begin{aligned} \frac{1 + \|w(x, e_1)\|}{1 + \|x\|} &\leq \frac{1 + |(a(x) + b(x)e_1)| + \|x\|}{1 + \|x\|} \\ &\leq 1 + \frac{|(a(x) + b(x)e_1)|}{1 + \|x\|} \\ &\leq 1 + \sup_{x \in \mathbb{R}^p} \left( \frac{\max(|a(x)|, b(x))}{1 + \|x\|} \right) \cdot (1 + |e_1|), \end{aligned}$$

wobei  $\sup_{x \in \mathbb{R}^p} \left( \frac{\max(|a(x)|, b(x))}{1 + \|x\|} \right)$  nach Annahme 3 endlich ist. Für den Erwartungswert gilt dann

$$\begin{aligned} E \left[ \left( \frac{1 + \|w(x, e_1)\|}{1 + \|x\|} \right)^s \right] &\leq E \left[ \left( 1 + \sup_{x \in \mathbb{R}^p} \left( \frac{\max(|a(x)|, b(x))}{1 + \|x\|} \right) \cdot (1 + |e_1|) \right)^s \right] \\ &\leq E \left[ \left( 1 + \sup_{x \in \mathbb{R}^p} \left( \frac{\max(|a(x)|, b(x))}{1 + \|x\|} \right) \cdot (1 + |e_1|) \right)^{r_0} \right]. \quad (3.26) \end{aligned}$$

Die zweite Zeile folgt dabei aus  $1 \leq 1 + \sup_{x \in \mathbb{R}^p} \left( \frac{\max(|a(x)|, b(x))}{1 + \|x\|} \right) (1 + |e_1|)$  und  $s \leq r_0$ . Es bleibt die Endlichkeit dieses Erwartungswertes zu zeigen. Betrachtet man

$$(1 + |z|)^{r_0} = (1 + |z|)^{\lfloor r_0 \rfloor} \cdot (1 + |z|)^{r_0 - \lfloor r_0 \rfloor}, \quad z \in \mathbb{R},$$

erkennt man, dass der erste Faktor aus Summanden besteht (Pascalsches Dreieck), deren Exponenten kleiner bzw. gleich  $r_0$  sind. Für den zweiten Faktor zeigen wir

$$(1 + |z|)^{r'} \leq 1 + |z|^{r'} \quad \text{für } 0 < r' < 1.$$

Sei  $|z| \leq 1$ , dann folgt

$$(1 + |z|)^{r'} \leq 1 + |z|,$$

### 3 Die Piggyback-Methode

da  $r' < 1$ . Andererseits gilt

$$1 + |z|^{r'} \geq 1 + |z|,$$

da  $|z| \leq 1$ . Also erhält man insgesamt

$$(1 + |z|)^{r'} \leq 1 + |z|^{r'} \text{ für } |z| \leq 1.$$

Falls  $|z| > 1$ , ergibt sich

$$\begin{aligned} (1 + |z|)^{r'} &= |z|^{r'} \cdot \left(1 + \frac{1}{|z|}\right)^{r'} \\ &\leq |z|^{r'} \cdot \left(1 + \frac{1}{|z|^{r'}}\right) \\ &= 1 + |z|^{r'}, \end{aligned}$$

wobei die zweite Zeile aus dem vorherigen Ergebnis folgt, da  $\frac{1}{|z|} < 1$ . Die Behauptung ist damit gezeigt und wir erhalten, dass  $(1 + |z|)^{r_0}$  in Summanden zerfällt, deren Exponenten kleiner bzw. gleich  $r_0$  sind. Anwendung dieses Ergebnisses auf (3.26) liefert Summanden der Form

$$E \left[ K \cdot |e_1|^{r'} \right] \text{ mit } r' \leq r_0, K < \infty.$$

Die Endlichkeit von (3.26) folgt aus  $E[|e_1|^{r_0}] < \infty$  (vgl. Annahme 2). □

Beachte, dass die Behauptung aus Lemma 3.10 für  $s \in [0, r_0]$  aufgrund der Abschätzung

$$\begin{aligned} \sup_{x \in \mathbb{R}^p} \frac{\max(|a^*(x)|, b^*(x))}{1 + \|x\|} &\leq \sup_{x \in \mathbb{R}^p} \frac{\max(|a^*(x)|, b^*(x))}{\|x\|} \\ &\leq \sup_{\frac{x}{\|x\|} \in \Theta} \max \left( \left| a^* \left( \frac{x}{\|x\|} \right) \right|, b^* \left( \frac{x}{\|x\|} \right) \right) < \infty \end{aligned}$$

ebenfalls für den homogenen Prozess  $X_1^* = w^*(x, e_1)$  erfüllt ist. Nach diesen Vorüberlegungen können wir uns der V-gleichmäßigen Ergodizität von  $(X_t)_{t \geq 0}$  zuwenden. Diese wird oft mit Hilfe der Drift-Bedingung (2.8) nachgewiesen. In unserem Fall nutzen wir Lemma 2.17 aus, welches die Drift-Bedingung impliziert. Wir konzentrieren uns auf den Nachweis von (2.9) und setzen hierfür die bisherigen Ergebnisse für die Prozesse  $(\theta_t^*)_{t \geq 0}$  und  $(\theta_t)_{t \geq 0}$  gewinnbringend ein.

**Satz 3.11.** *Seien Annahmen 1-6 vorausgesetzt und sei  $\rho$  wie in Satz 3.4 definiert. Wenn  $\rho < 1$ , dann existieren für jedes  $\rho_1 \in (\rho, 1)$  ein  $K < \infty$ ,  $s > 0$  und  $V : \mathbb{R}^p \rightarrow [1, \infty)$ , so dass*

$$E \left[ V(X_1) \middle| X_0 = x \right] \leq K + \rho_1^s V(x) \text{ für alle } x \in \mathbb{R}^p, \tag{3.27}$$

wobei die Testfunktion die Form  $V(x) = 1 + \lambda(x)\|x\|^s$  hat, mit einer beschränkten und von 0 weg beschränkten Funktion  $\lambda$ . Insbesondere ist  $(X_t)_{t \geq 0}$  V-gleichmäßig ergodisch.

### 3 Die Piggyback-Methode

Beweis: (vgl. [6], Theorem 3.2)

Es genügt die Drift-Bedingung (2.8) zu zeigen. Hierfür weisen wir die Voraussetzungen für Lemma 2.17 nach. Nach Lemma 3.1 i) handelt es sich bei  $(X_t)_{t \geq 0}$  um eine aperiodische,  $\lambda_p$ -irreduzible T-chain und  $\rho_1 < 1$  gilt nach Voraussetzung. Mit Hilfe des Beweises von Lemma 3.10 folgert man, dass  $\sup_{x \in \mathbb{R}_0^p} E_x \left[ \left( \frac{\|X_1\|}{\|x\|} \right)^s \right] < \infty$ .  $E[\|X_1\|^s]$  ist lokal beschränkt nach Annahme 2 und Annahme 4. Es bleibt (2.9) zu zeigen, das heißt

$$\limsup_{\|x\| \rightarrow \infty} E \left[ \nu(X_{n+1}) - \nu(X_n) + \log \left( \frac{1 + \|X_{n+1}\|}{1 + \|X_n\|} \right) \middle| X_0 = x \right] < \log \rho_1 < 0,$$

wobei  $\nu$  eine beschränkte Funktion ist. An dieser Stelle kommt die Piggyback-Methode zum Einsatz, welche die Anwendung der Ergebnisse für die Prozesse  $(\theta_t)_{t \geq 0}$  und  $(\theta_t^*)_{t \geq 0}$  beinhaltet. Zunächst teilen wir die bedingte Erwartung mit Hilfe einer geeigneten Menge in vier Terme auf. Die Ergebnisse des letzten Kapitels werden ausgenutzt, um jeden dieser Terme abzuschätzen. Eine entscheidende Eigenschaft ist die near-equilibrium-Gleichung (3.11), die offensichtlich eine ähnliche Struktur wie die zu untersuchende bedingte Erwartung aufweist. Weiterhin spielt die Stetigkeit von  $\nu$  auf  $\Theta_{\times \times}$  (vgl. Satz 3.6) in Verbindung mit Bemerkung 3.9 eine wichtige Rolle. Die Beziehung zwischen den beiden Prozessen  $(\theta_t^*)_{t \geq 0}$  und  $(\theta_t)_{t \geq 0}$ , die in Lemma 3.8 zum Ausdruck kommt, liefert zum einen die Menge, welche die vier Terme definiert und trägt bei allen vier Termen zur Abschätzung bei. Als erstes modifizieren wir die near-equilibrium-Gleichung, um sie auf unser Problem anwenden zu können.

$$\sup_{\theta \in \Theta} \left| E \left[ \nu(\theta_1^*) - \nu(\theta) + \log(\|w^*(\theta, e_1)\|) \middle| \theta_0^* = \theta \right] - \log \rho \right| < \frac{\epsilon}{4}$$

ist aufgrund der Definition von  $\nu$  nur für  $\theta \in \Theta$  definiert. Damit sie für den Prozess  $(X_t)_{t \geq 0}$  ausgenutzt werden kann, muss  $\nu$  auf  $x \in \mathbb{R}^p$  erweitert werden, daher definiert man

$$\nu(x) := \nu\left(\frac{x}{\|x\|}\right) \text{ für } x \in \mathbb{R}_0^p \text{ und } \nu(0) := 0.$$

Wegen  $\|w^*(x, e_1)\| \xrightarrow[\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}]{\infty} \infty$  P-f.s. gilt weiterhin

$$\limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{1 + \|w^*(x, e_1)\|}{1 + \|x\|} \right) \right] = \limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{\|w^*(x, e_1)\|}{1 + \|x\|} \right) \right] \quad (3.28)$$

sowie

$$\begin{aligned} \limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{\|w^*(x, e_1)\|}{\|x\|} \right) \right] &= \limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{\|w^*(x, e_1)\|}{1 + \|x\|} \right) + \log \left( \frac{1 + \|x\|}{\|x\|} \right) \right] \\ &= \limsup_{\|x\| \rightarrow \infty} E_x \left[ \log \left( \frac{\|w^*(x, e_1)\|}{1 + \|x\|} \right) \right]. \end{aligned} \quad (3.29)$$

### 3 Die Piggyback-Methode

Für die near-equilibrium-Gleichung ergibt sich nun

$$\begin{aligned}
& \limsup_{\|y\| \rightarrow \infty} \left| E \left[ \nu(w^*(y, e_1)) - \nu(y) + \log \left( \frac{1 + \|w^*(y, e_1)\|}{1 + \|y\|} \right) \right] - \log \rho \right| & (3.30) \\
&= \limsup_{\|y\| \rightarrow \infty} \left| E \left[ \nu(w^*(y, e_1)) - \nu(y) + \log \left( \frac{\|w^*(y, e_1)\|}{\|y\|} \right) \right] - \log \rho \right| \\
&= \sup_{\frac{y}{\|y\|} \in \Theta} \left| E \left[ \nu \left( \frac{w^*(\frac{y}{\|y\|}, e_1)}{\|w^*(\frac{y}{\|y\|}, e_1)\|} \right) - \nu \left( \frac{y}{\|y\|} \right) + \log \left( \|w^*(\frac{y}{\|y\|}, e_1)\| \right) \right] - \log \rho \right| < \frac{\epsilon}{4}.
\end{aligned}$$

Wenden wir uns der Abschätzung von (2.9) zu. Es muss gezeigt werden:

$$\limsup_{\|x\| \rightarrow \infty} E \left[ \nu(X_{n+1}) - \nu(X_n) + \log \left( \frac{1 + \|X_{n+1}\|}{1 + \|X_n\|} \right) \middle| X_0 = x \right] < \log \rho_1.$$

Zur bereits erwähnten Aufteilung dieses Ausdrucks definiert man die Menge

$$A_t := \{\tilde{\theta}_t, \tilde{\theta}_t^* \in \tilde{D}, \|X_t\| \geq \eta^t \|X_0\|\},$$

wobei  $\tilde{D} \subset \Theta_{\times \times}$  die kompakte Menge aus Lemma 3.8 ist. Zur Erinnerung:  $\tilde{\theta}_t = w(X_{t-1}, e_t)$  und  $\tilde{\theta}_t^* = w^*(X_{t-1}, e_t)$  unterscheiden sich darin, dass einmal die Funktionen  $a$  und  $b$  auf  $X_{t-1}$  und einmal die homogenen Funktionen  $a^*$  und  $b^*$  auf  $X_{t-1}$  angewendet werden. (2.9) zerlegen wir dann mit der Menge  $A_{2p}$  zu

$$\begin{aligned}
& \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right] - \log \rho \right| \\
&\leq \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| & (3.31)
\end{aligned}$$

$$+ \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}^c} \right] \right|. \quad (3.32)$$

Betrachten wir zunächst (3.31) für die Menge  $A_{2p}$ . Sei  $(\tilde{X}_t^*)_{t \geq 0}$  der homogene Prozess, angewendet auf den Zustand  $X_{t-1}$ , also  $\tilde{X}_t^* = (a^*(X_{t-1}) + b^*(X_{t-1})e_t, \xi_{t-1}, \dots, \xi_{t-p+1})$ .

Durch das Einfügen der Terme  $\nu(\tilde{X}_{2p+1}^*) - \nu(\tilde{X}_{2p+1}^*)$  und  $\log \left( \frac{\tilde{X}_{2p+1}^*}{\tilde{X}_{2p+1}^*} \right)$  ergibt sich

$$\begin{aligned}
& \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| \\
&\leq \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(\tilde{X}_{2p+1}^*) \right) \mathbf{1}_{A_{2p}} \right] \right| & (i) \\
&+ \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(\tilde{X}_{2p+1}^*) - \nu(X_{2p}) + \log \left( \frac{1 + \|\tilde{X}_{2p+1}^*\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| & (ii) \\
&+ \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|\tilde{X}_{2p+1}^*\|} \right) \mathbf{1}_{A_{2p}} \right] \right|. & (iii)
\end{aligned}$$

### 3 Die Piggyback-Methode

Diese drei Terme schätzen wir mit den Ergebnissen für  $(\theta_t^*)_{t \geq 0}$  und  $(\theta_t)_{t \geq 0}$  ab. Bei Ausdruck (i) nutzen wir die Stetigkeit von  $\nu$  auf der kompakten Menge  $A_{2p} \subset \tilde{D} \subset \Theta_{\times \times}$  aus. Da  $|\nu(x) - \nu(x')| \leq 2 \sup_{x \in \mathbb{R}^p} |\nu(x)|$  erhält man, indem man die bedingte Erwartung ein weiteres Mal mit der Menge  $A_{2p+1}$  aufteilt,

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} E_x \left[ \left| \nu(X_{2p+1}) - \nu(\tilde{X}_{2p+1}^*) \right| \mathbf{1}_{A_{2p}} \right] \\ & \leq \limsup_{\|x\| \rightarrow \infty} E_x \left[ \left| \nu(X_{2p+1}) - \nu(\tilde{X}_{2p+1}^*) \right| \mathbf{1}_{A_{2p}} \mathbf{1}_{A_{2p+1}} \right] \\ & \quad + 2 \sup_{x \in \mathbb{R}^p} |\nu(x)| \cdot \limsup_{\|x\| \rightarrow \infty} E_x \left[ \mathbf{1}_{A_{2p+1}^c} \right] \\ & \leq \lim_{\substack{\|y\| \rightarrow \infty \\ \frac{y}{\|y\|} \in \tilde{D}}} E_y \left[ \left| \nu\left(\frac{X_1}{\|X_1\|}\right) - \nu\left(\frac{X_1^*}{\|X_1^*\|}\right) \right| \mathbf{1}_{A_1} \right] \\ & \quad + 2 \sup_{x \in \mathbb{R}^p} |\nu(x)| \cdot \limsup_{\|x\| \rightarrow \infty} P(A_{2p+1}^c | X_0 = x). \end{aligned}$$

Den letzten Schritt kann man sich folgendermaßen überlegen:

In der bedingten Erwartung tauchen zwei Indikatorfunktionen auf. Diese geben nur einen Beitrag, wenn aus  $X_{2p} \in A_{2p}$  folgt, dass  $X_{2p+1} \in A_{2p+1}$  ist. Setzt man  $y := X_{2p}$  gilt deshalb  $X_1 \in A_1$ , falls  $\frac{y}{\|y\|} = \frac{X_{2p}}{\|X_{2p}\|}$  in  $\tilde{D}$ , natürlich unter Verwendung der Markov-Eigenschaft.

Wir zeigen jetzt, dass der erste Summand für  $\|y\| \rightarrow \infty$  gegen Null strebt. Aus Bemerkung 3.9 erhalten wir

$$\lim_{\substack{\|y\| \rightarrow \infty \\ \frac{y}{\|y\|} \in \tilde{D}}} \|\tilde{\theta}_1 - \tilde{\theta}_1^*\| = 0 \text{ P-f.s. .}$$

Die Funktion  $\nu$  ist stetig auf  $\Theta_{\times \times}$  (vgl. Satz 3.6), deshalb folgt

$$\lim_{\substack{\|y\| \rightarrow \infty \\ \frac{y}{\|y\|} \in \tilde{D}}} |\nu(\tilde{\theta}_1) - \nu(\tilde{\theta}_1^*)| = 0 \text{ P-f.s. .}$$

$\nu$  ist außerdem beschränkt, also ergibt sich

$$\lim_{\substack{\|y\| \rightarrow \infty \\ \frac{y}{\|y\|} \in \tilde{D}}} E \left[ \left| \nu(\tilde{\theta}_1) - \nu(\tilde{\theta}_1^*) \right| \middle| X_0 = y \right] = 0.$$

Aus Lemma 3.8 wissen wir, dass

$$P(A_t | X_0 = x) > (1 - \tilde{\epsilon})^t \geq 1 - t\tilde{\epsilon}$$

gilt. Die zweite Ungleichung sieht man zum Beispiel leicht mit Induktion über  $t$  ein. Für Term (i) haben wir dann insgesamt

$$\limsup_{\|x\| \rightarrow \infty} E_x \left[ \left| \nu(X_{2p+1}) - \nu(\tilde{X}_{2p+1}^*) \right| \mathbf{1}_{A_{2p}} \right] \leq 2 \sup |\nu(x)| \cdot (2p + 1)\tilde{\epsilon}$$

### 3 Die Piggyback-Methode

herausbekommen. Als nächstes behandeln wir Term (ii). Hier wenden wir die modifizierte near-equilibrium-Gleichung (3.30) an, um die gewünschte Abschätzung zu erhalten. Schreibt man die bedingte Erwartung in Term (ii), aus erhält man

$$\begin{aligned}
& \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(\tilde{X}_{2p+1}^*) - \nu(X_{2p}) + \log \left( \frac{1 + \|\tilde{X}_{2p+1}^*\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| \\
&= \limsup_{\|x\| \rightarrow \infty} \left| \int_{\mathbb{R}^p} E_y \left[ \nu(\tilde{X}_1^*) - \nu(y) + \log \left( \frac{1 + \|\tilde{X}_1^*\|}{1 + \|y\|} \right) \right] \mathbf{1}_{A_{2p}} P_x(X_{2p} \in dy) - \log \rho \right| \\
&\leq \limsup_{\|x\| \rightarrow \infty} \int_{\mathbb{R}^p} \left| E_y \left[ \nu(\tilde{X}_1^*) - \nu(y) + \log \left( \frac{1 + \|\tilde{X}_1^*\|}{1 + \|y\|} \right) \right] \mathbf{1}_{A_{2p}} - \log \rho \right| P_x(X_{2p} \in dy) \\
&= \limsup_{\|x\| \rightarrow \infty} \int_{\mathbb{R}^p} \left| E_y \left[ \nu(\tilde{X}_1^*) - \nu(y) + \log \left( \frac{1 + \|\tilde{X}_1^*\|}{1 + \|y\|} \right) \right] - \log \rho \right| \mathbf{1}_{A_{2p}} P_x(X_{2p} \in dy) \\
&\quad + \limsup_{\|x\| \rightarrow \infty} \int_{\mathbb{R}^p} |\log \rho| \mathbf{1}_{A_{2p}^c} P_x(X_{2p} \in dy) \\
&\leq \limsup_{\|y\| \rightarrow \infty} \left| E_y \left[ \nu(\tilde{X}_1^*) - \nu(y) + \log \left( \frac{1 + \|\tilde{X}_1^*\|}{1 + \|y\|} \right) \right] - \log \rho \right| + |\log \rho| \limsup_{\|x\| \rightarrow \infty} P_x(A_{2p}^c) \\
&\leq \frac{\epsilon}{4} + |\log \rho| 2p\tilde{\epsilon}.
\end{aligned}$$

In der letzten Zeile wurde die modifizierte near-equilibrium-Gleichung (3.30) und wiederholt Lemma 3.8 angewendet. Bleibt noch Term (iii) zu behandeln. Da sich  $\|X_1\|$  und  $\|X_1^*\|$  höchstens um  $|a_0(x) + b_0(x)e_1|$  unterscheiden mit  $a_0(x) = o(\|x\|)$  und  $b_0(x) = o(\|x\|)$ , sieht man ein, dass

$$\limsup_{\substack{\|x\| \rightarrow \infty \\ \frac{x}{\|x\|} \in \tilde{D}}} \left| E_x \left[ \log \left( \frac{1 + \|X_1\|}{1 + \|X_1^*\|} \right) \right] \right| = 0.$$

Für  $x \in \mathbb{R}^p$ ,  $\frac{x}{\|x\|} \in \tilde{D}$  gilt nämlich

$$\begin{aligned}
\left| E_x \left[ \log \left( \frac{1 + \|X_1\|}{1 + \|X_1^*\|} \right) \right] \right| &\leq \left| E_x \left[ \log \left( \frac{1 + \|X_1^*\| + |a_0(x) + b_0(x)e_1|}{1 + \|X_1^*\|} \right) \right] \right| \\
&\leq \left| E_x \left[ \log \left( 1 + \frac{|a_0(x) + b_0(x)e_1|}{\|X_1^*\|} \right) \right] \right| \\
&\leq \left| E_x \left[ \log \left( 1 + \frac{|a_0(x) + b_0(x)e_1|}{\eta\|x\|} \right) \right] \right| < \infty \\
&\leq \left| E_x \left[ \log \left( 1 + \frac{\max(|a_0(x)|, |b_0(x)|) \cdot (1 + |e_1|)}{\eta\|x\|} \right) \right] \right| < \infty,
\end{aligned}$$

### 3 Die Piggyback-Methode

da  $E[|e_1|^{r_0}] < \infty$  nach Annahme 2. Für  $\|x\| \rightarrow \infty$  strebt dieser Erwartungswert gegen 0. Wendet man dies auf Term (iii) an, erhält man

$$\limsup_{\|x\| \rightarrow \infty} \left| E \left[ \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|\tilde{X}_{2p+1}^*\|} \right) \mathbf{1}_{A_{2p}} \middle| X_0 = x \right] \right| = 0,$$

da  $\|y\| = \|X_{2p}\| \xrightarrow{\|x\| \rightarrow \infty} \infty$  auf der Menge  $A_{2p}$ . Damit haben wir alle drei Terme für die Menge  $A_{2p}$  abgeschätzt:

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| \\ & \leq |(i)| + |(ii)| + |(iii)| \\ & \leq 2 \sup |\nu(x)|(2p+1)\tilde{\epsilon} + \frac{\epsilon}{4} + |\log \rho| 2p\tilde{\epsilon} + 0. \end{aligned}$$

Um (2.9) vollständig abzuschätzen, muss die Ungleichung noch für die Menge  $A_{2p}^c$  behandelt werden. Es gilt:

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}^c} \right] \right| \\ & \leq \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ (\nu(X_{2p+1}) - \nu(X_{2p})) \mathbf{1}_{A_{2p}^c} \right] \right| \\ & \quad + \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \mathbf{1}_{A_{2p}^c} \right] \right|. \end{aligned}$$

Mit den Resultaten, die wir schon für die vorherigen Terme benutzt haben, ergibt sich

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}^c} \right] \right| \\ & \leq 2 \sup_{x \in \mathbb{R}^p} |\nu(x)| 2p\tilde{\epsilon} + \sup_{y \in \mathbb{R}^p} E_y \left[ \left| \log \left( \frac{1 + \|w(y, e_1)\|}{1 + \|y\|} \right) \right| \right] 2p\tilde{\epsilon} \\ & \leq (2 \sup_{x \in \mathbb{R}^p} |\nu(x)| + \tilde{K}) \cdot 2p\tilde{\epsilon}, \end{aligned}$$

wobei  $\sup_{y \in \mathbb{R}^p} E \left[ \left| \log \left( \frac{1 + \|w(y, e_1)\|}{1 + \|y\|} \right) \right| \right] \leq \tilde{K}$ , wie man schnell mit Lemma 3.10 einsieht. Wir haben jetzt (2.9) sowohl für die Menge  $A_{2p}$  als auch für die Menge  $A_{2p}^c$  bearbeitet und



### 3 Die Piggyback-Methode

erhalten abschließend

$$\begin{aligned}
& \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right] - \log \rho \right| \\
& \leq \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}} \right] - \log \rho \right| \\
& \quad + \limsup_{\|x\| \rightarrow \infty} \left| E_x \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \mathbf{1}_{A_{2p}^c} \right] \right| \\
& < 2 \sup_{x \in \mathbb{R}^p} |\nu(x)|(2p+1)\tilde{\epsilon} + \frac{\epsilon}{4} + |\log \rho|2p\tilde{\epsilon} + (2 \sup_{x \in \mathbb{R}^p} |\nu(x)| + \tilde{K}) \cdot 2p\tilde{\epsilon}.
\end{aligned}$$

Wählt man  $\epsilon > 0$ , so dass  $\log \rho + \epsilon < 0$  und definiert

$$E_{\max} := \max \left( 2 \sup_{x \in \mathbb{R}^p} |\nu(x)|(2p+1), |\log \rho|2p, (2 \sup_{x \in \mathbb{R}^p} |\nu(x)| + \tilde{K})2p \right),$$

dann gilt für alle  $\tilde{\epsilon} < \frac{\epsilon}{4 \cdot E_{\max}}$

$$\begin{aligned}
& \limsup_{\|x\| \rightarrow \infty} \left| E \left[ \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \middle| X_0 = x \right] - \log \rho \right| < \epsilon \quad (3.33) \\
\Rightarrow & \limsup_{\|x\| \rightarrow \infty} E \left[ \left( \nu(X_{2p+1}) - \nu(X_{2p}) + \log \left( \frac{1 + \|X_{2p+1}\|}{1 + \|X_{2p}\|} \right) \right) \middle| X_0 = x \right] < \log \rho + \epsilon.
\end{aligned}$$

Mit der Definition  $\log \rho_1 := \log \rho + \epsilon < 0$  sind alle Voraussetzungen für Lemma 2.17 erfüllt und  $(X_t)_{t \geq 0}$  ist  $V$ -gleichmäßig ergodisch.  $\square$

### 3.6 Der Lyapounov-Exponent und die Transienz von $(X_t)_{t \geq 0}$

In diesem Abschnitt wird ein wesentlicher Vorteil der Piggyback-Methode bzw. des resultierenden Stabilitätskriterium herausgearbeitet. Bisher wissen wir, dass der Prozess  $(X_t)_{t \geq 0}$  ergodisch und damit rekurrent ist, falls  $\log \rho < 0$ . Wir zeigen in diesem Teil der Arbeit:  $(X_t)_{t \geq 0}$  ist transient, falls  $\log \rho \geq 0$ . Hierbei benutzen wir, dass  $\log \rho$  mit dem Lyapounov-Exponenten übereinstimmt. Diese Korrespondenz an sich ist schon bemerkenswert, da  $(X_t)_{t \geq 0}$  nach einem weiteren Stabilitätsbegriff stabil ist. Wir stellen zuerst die Verbindung zum Lyapounov-Exponenten her.

**Satz 3.12.** *Annahmen 1-6 seien vorausgesetzt und  $\rho$  sei definiert wie in Satz 3.4. Dann ist  $\log \rho$  der Lyapounov-Exponent von  $(X_t)_{t \geq 0}$ , das heißt*

$$\lim_{n \rightarrow \infty} \limsup_{\|x\| \rightarrow \infty} \left| \frac{1}{n} E_x \left[ \log \left( \frac{\|X_n\|}{\|X_0\|} \right) \right] - \log \rho \right| = 0. \quad (3.34)$$

Beweis: (vgl. [6], Theorem 3.3)

Es genügt wegen (3.28) und (3.29)

$$\limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E_x \left[ \log \left( \frac{\tilde{V}(X_n)}{\tilde{V}(X_0)} \right) \right] - \log \rho \right| < \epsilon, \forall n \geq n_0,$$

zu zeigen, wobei  $\tilde{V}(x) = 1 + \|x\|$ . Es gilt:

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E_x \left[ \log \left( \frac{\tilde{V}(X_n)}{\tilde{V}(X_0)} \right) \right] - \log \rho \right| \\ & \leq \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E_x \left[ \nu(X_{2p+n}) - \nu(X_{2p}) + \log \left( \frac{\tilde{V}(X_{2p+n})}{\tilde{V}(X_{2p})} \right) \right] - \log \rho \right| \\ & \quad + \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} E_x \left[ \left| \log \left( \frac{\tilde{V}(X_n)}{\tilde{V}(X_0)} \right) - \log \left( \frac{\tilde{V}(X_{2p+n})}{\tilde{V}(X_{2p})} \right) \right| \right] \\ & \quad + \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E_x \left[ \nu(X_{2p}) - \nu(X_{2p+n}) \right] \right|. \end{aligned}$$

Die beiden letzten Terme können wie in ähnlicher Form schon mehrmals gesehen durch

$$\frac{1}{n} 4p \sup_{x \in \mathbb{R}^p} E_x \left| \log \left( \frac{\tilde{V}(X_1)}{\tilde{V}(x)} \right) \right| + \frac{1}{n} 2 \sup_{x \in \mathbb{R}^p} \nu(x)$$

abgeschätzt werden.  $\sup_{x \in \mathbb{R}^p} E_x \left[ \left| \log \left( \frac{\tilde{V}(X_1)}{\tilde{V}(x)} \right) \right| \right]$  ist dabei nach Lemma 3.10 endlich. Der erste Term bedarf ein wenig mehr Arbeit. Definiert man

$$B_j := \nu(X_{2p+j}) - \nu(X_{2p+j-1}) + \log \left( \frac{\tilde{V}(X_{2p+j})}{\tilde{V}(X_{2p+j-1})} \right),$$

### 3 Die Piggyback-Methode

dann folgt

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E_x \left[ \nu(X_{2p+n}) - \nu(X_{2p}) + \log \left( \frac{\tilde{V}(X_{2p+n})}{\tilde{V}(X_{2p})} \right) \right] - \log \rho \right| \\ &= \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| \sum_{j=1}^n E_x [B_j] - \log \rho \right| = \limsup_{\|x\| \rightarrow \infty} \left| \frac{1}{n} \sum_{j=1}^n \left( E_x [B_j] - \log \rho \right) \right|. \end{aligned}$$

Die Terme  $\limsup_{\|x\| \rightarrow \infty} \left| E_x [B_j] - \log \rho \right|$  schätzen wir per Induktion ab. Der Induktionsanfang

$\limsup_{\|x\| \rightarrow \infty} \left| E_x [B_1] - \log \rho \right| < \epsilon$  ist durch (3.33) gegeben, während der Induktionsschritt durch

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \left| E [B_{j+1} | X_0 = x] - \log \rho \right| \\ & \leq \limsup_{\|x\| \rightarrow \infty} E \left[ \left| E [B_{j+1} | X_1] - \log \rho \right| \middle| X_0 = x \right] \\ & \leq \limsup_{\|x\| \rightarrow \infty} \left| E [B_j | X_0 = x] - \log \rho \right| \end{aligned}$$

gezeigt ist. Insgesamt ergibt sich somit

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} \left| E \left[ \log \left( \frac{\tilde{V}(X_n)}{\tilde{V}(X_0)} \right) \middle| X_0 = x \right] - \log \rho \right| \\ & \leq \epsilon + \frac{1}{n} 4p \sup_{x \in \mathbb{R}^p} E \left[ \left| \log \left( \frac{\tilde{V}(X_1)}{\tilde{V}(x)} \right) \right| \right] + \frac{1}{n} 2 \sup_{x \in \mathbb{R}^p} \nu(x) < 2\epsilon, \quad \forall n \geq n_0, \end{aligned}$$

wobei  $n_0$  so gewählt wurde, dass  $\frac{1}{n_0} 4p \sup_{x \in \mathbb{R}^p} E \left[ \left| \log \left( \frac{\tilde{V}(X_1)}{\tilde{V}(x)} \right) \right| \right] + \frac{1}{n_0} 2 \sup_{x \in \mathbb{R}^p} \nu(x) < \epsilon$ . Damit ist die Behauptung gezeigt.  $\square$

Zum Abschluss der Ergebnisse zeigen wir, dass die Bedingung aus Satz 3.11 scharf ist, das heißt, dass der Prozess transient, sogar geometrisch transient, ist, falls  $(X_t)_{t \geq 0}$  (3.27) nicht erfüllt.  $(X_t)_{t \geq 0}$  heißt geometrisch transient, falls

$$P(q^t \|X_t\| \rightarrow \infty | X_0 = x) > 0, \quad \forall x \in \mathbb{R}^p$$

für ein  $q < 1$  gegeben ist.

**Satz 3.13.** *Gilt in unserem Modell  $\log \rho > 0$ , dann ist  $(X_t)_{t \geq 0}$  geometrisch transient.*

Beweis: (vgl. [5], Theorem 2.1 und Theorem 2.2)

Wir weisen nur die entscheidende dritte Eigenschaft aus [5], Theorem 2.1 mit Hilfe von

### 3 Die Piggyback-Methode

Theorem 2.2 nach. Es ist zu zeigen, dass eine Menge  $D \subset \mathbb{R}^p$  und ein  $r > 0$  existieren, so dass

$$\sup_x E_x \left[ \left( \frac{\tilde{V}(x)}{\tilde{V}(X_1)} \right)^r \mathbf{1}_{\{X_1 \in D\}} \right] < \infty.$$

Außerdem muss für ein  $q \in (0, 1)$

$$\limsup_{\tilde{V}(x) \rightarrow \infty} E_x \left[ \sum_{j=1}^m \log \left( \frac{\tilde{V}(X_{j-1})}{q \tilde{V}(X_j)} \right) \mathbf{1}_{\{X_j \in D\}} \right] < 0$$

erfüllt sein. Wir zeigen zuerst die zweite Bedingung. Wegen (3.34) und  $\rho > 1$  existiert ein  $\rho'$ ,  $1 < \rho' < \rho$ , mit

$$\lim_{n \rightarrow \infty} \limsup_{\|x\| \rightarrow \infty} \frac{1}{n} E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\|} \right) \middle| X_0 = x \right] - \log \rho' > 0.$$

Also existiert ein  $n_0$ , so dass

$$\limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\|} \right) \middle| X_0 = x \right] - \log \rho' > 0.$$

Umformen der Logarithmen ergibt

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\|} \right) \middle| X_0 = x \right] - \log \rho' > 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} \left( E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\|} \right) \middle| X_0 = x \right] - n_0 \cdot \log \rho' \right) > 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} \left( E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\|} \right) \middle| X_0 = x \right] - \log \rho'^{n_0} \right) > 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\| \cdot \rho'^{n_0}} \right) \middle| X_0 = x \right] > 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} -\frac{1}{n_0} E \left[ \log \left( \frac{\|X_{n_0}\|}{\|X_0\| \cdot \rho'^{n_0}} \right) \middle| X_0 = x \right] < 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_0\| \cdot \rho'^{n_0}}{\|X_{n_0}\|} \right) \middle| X_0 = x \right] < 0. \end{aligned}$$

Definiert man  $q := \frac{1}{\rho'}$ , erhält man

$$\begin{aligned} & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_0\| \cdot \rho'^{n_0}}{\|X_{n_0}\|} \right) \middle| X_0 = x \right] < 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \log \left( \frac{\|X_0\|}{\|X_{n_0}\| \cdot q^{n_0}} \right) \middle| X_0 = x \right] < 0 \\ \Leftrightarrow & \limsup_{\|x\| \rightarrow \infty} \frac{1}{n_0} E \left[ \sum_{k=1}^{n_0} \log \left( \frac{\|X_{j-1}\|}{\|X_j\| \cdot q} \right) \middle| X_0 = x \right] < 0. \end{aligned}$$

### 3 Die Piggyback-Methode

Demnach ist die zweite Bedingung für  $D = \mathbb{R}^p$  und  $q = \frac{1}{p'} < 1$  erfüllt. Die erste Bedingung folgt aber direkt mit Lemma 3.10. Damit ist die Behauptung gezeigt.  $\square$

### 3.7 Ein einfaches Beispiel

Als Beispiel betrachten wir einen eindimensionalen Prozess gegeben durch:

$$\begin{aligned} X_t &= a^*(X_{t-1}) + b^*(X_{t-1})e_t + a_0(X_{t-1}) + b_0(X_{t-1})e_t \\ \text{mit } a^*(x) &= (a_1 \mathbf{1}_{\{x < 0\}}(x) + a_2 \mathbf{1}_{\{x > 0\}}(x)) \cdot x \\ \text{und } b^*(x) &= (b_1 \mathbf{1}_{\{x < 0\}}(x) + b_2 \mathbf{1}_{\{x > 0\}}(x)) \cdot |x|, \end{aligned}$$

$b_1, b_2 > 0$  sowie  $a_0(x) = \sqrt{|x|} + 1$  und  $b_0(x) = \sqrt{|x|} + 1$ . Dann sind  $a^*$  und  $b^*$  homogen,  $a_0 = o(\|x\|)$  und  $b_0 = (\|x\|)$  und es gilt  $b(x) = b^*(x) + b_0(x) > 0, \forall x \in \mathbb{R}$ . Des Weiteren sind die Annahmen 1-6 erfüllt. Wir wollen nun das Stabilitätskriterium aus Satz 3.11 überprüfen. Da der Prozess  $(X_t)_{t \geq 0}$  auf  $\mathbb{R}$  definiert ist, gilt  $\Theta = \{-1, +1\}$ . Der normierte Prozess  $(\theta_t)_{t \geq 0}$  ist also nur auf den beiden Zuständen  $-1$  und  $+1$  definiert. In diesem Fall lässt sich die invariante Verteilung  $\pi$  leicht berechnen. Hierzu definieren wir

$$p_1 := P(\theta_1^* = +1 | \theta_0^* = -1) \quad \text{und} \quad p_2 := P(\theta_1^* = -1 | \theta_0^* = +1).$$

Für die invariante Verteilung muss dann

$$\begin{aligned} \pi(-1) &= \pi(-1)(1 - p_1) + \pi(+1)p_2, \\ \pi(+1) &= \pi(-1)p_1 + \pi(+1)(1 - p_2) \end{aligned}$$

gelten. Stellt man eine der beiden Gleichungen um, erhält man

$$\pi(-1)p_1 = \pi(+1)p_2,$$

woraus direkt  $\pi(-1) = k \cdot p_2$  und  $\pi(+1) = k \cdot p_1$  für eine Konstante  $k \neq 0$  folgt. Normierung liefert das eindeutige invariante Wahrscheinlichkeitsmaß

$$\pi(-1) = \frac{p_2}{p_1 + p_2} \quad \text{und} \quad \pi(+1) = \frac{p_1}{p_1 + p_2}.$$

Um Stabilität nachzuweisen, muss  $\log \rho$  berechnet werden. In diesem Fall gilt

$$\begin{aligned} \log \rho &= \int_{\Theta} \int_{\mathbb{R}} \log(|a^*(\theta) + b^*(\theta)e_1|) f(u) du \pi(d\theta) \\ &= \int_{\mathbb{R}} \log(|a^*(-1) + b^*(-1)e_1|) f(u) du \cdot \pi(-1) \\ &\quad + \int_{\mathbb{R}} \log(|a^*(+1) + b^*(+1)e_1|) f(u) du \cdot \pi(+1) \\ &= \frac{p_2}{p_1 + p_2} E \left[ \log |a_1 - b_1 e_1| \right] + \frac{p_1}{p_1 + p_2} E \left[ \log |a_2 + b_2 e_1| \right]. \end{aligned}$$

Gilt  $\log \rho < 1$ , so ist der Prozess ergodisch (vgl. Satz 3.11). Um  $\log \rho$  wenigstens numerisch berechnen zu können, müssen wir uns überlegen, wie man die Übergangswahrscheinlichkeiten  $p_1$  und  $p_2$  bestimmen kann. Betrachten wir  $p_1 = P(\theta_1^* = +1 | \theta_0^* = -1)$ .

### 3 Die Piggyback-Methode

Da  $\theta_1^* = \frac{a^*(-1)+b^*(-1)e_1}{|a^*(-1)+b^*(-1)e_1|}$  bleibt zu überlegen, wann  $a^*(-1) + b^*(-1)e_1 > 0$  gegeben ist. Es gilt:

$$\begin{aligned} & a^*(-1) + b^*(-1)e_1 > 0 \\ \Leftrightarrow & -a_1 + b_1e_1 > 0 \\ \Leftrightarrow & a_1 - b_1e_1 < 0. \end{aligned}$$

Also erhält man

$$p_1 = P(a_1 - b_1e_1 < 0).$$

Analog ergibt sich

$$p_2 = P(a_2 + b_2e_1 < 0).$$

Eine kleine Simulation des Prozesses  $(X_t)_{t \geq 0}$  wurde für unabhängige, identisch normalverteilte  $e_t$  durchgeführt. Abgebildet sind zwei Parametersätze, wobei einmal das Stabilitätskriterium erfüllt ist und einmal nicht. Bei dem Parametersatz

$$a_1 = 0,9; a_2 = 0,8; b_1 = 0,3; b_2 = 0,5$$

ist das Stabilitätskriterium mit  $\log \rho = -0,2915$  erfüllt. Abbildung 3.1 stellt eine Rea-

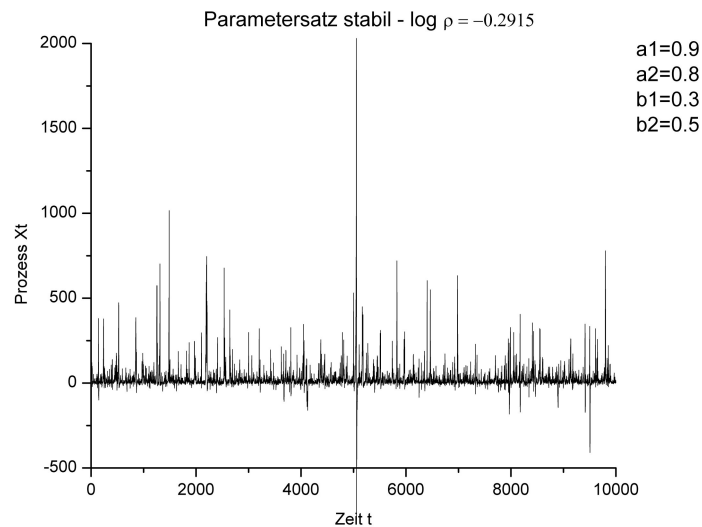


Abbildung 3.1: Parametersatz "stabil"

lisierung des eben spezifizierten Modells für diesen Parametersatz dar. Wie zu erwarten hält sich der Prozess überwiegend in einem begrenzten Intervall auf und kehrt immer wieder in die Nähe des Startpunktes zurück. Bei dem Parametersatz

$$a_1 = 1,2; a_2 = 0,4; b_1 = 0,2; b_2 = 0,85$$

ist das Stabilitätskriterium mit  $\log \rho = 0,0250$  nicht erfüllt.

### 3 Die Piggyback-Methode

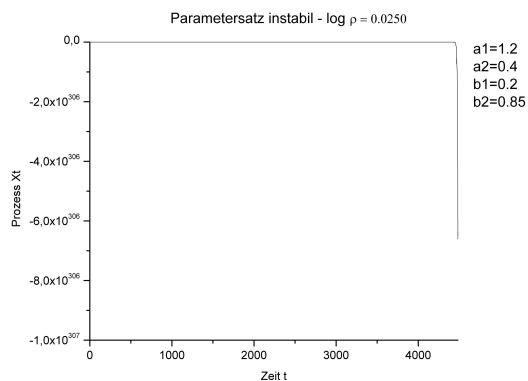
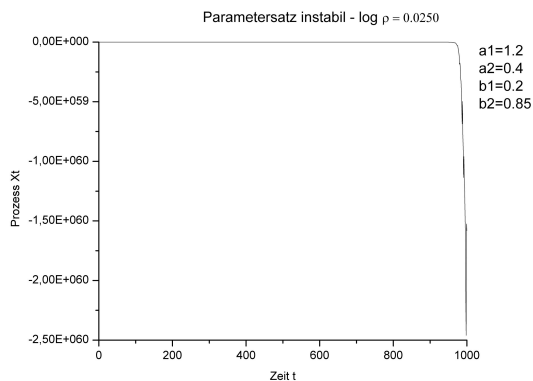


Abbildung 3.2: Parametersatz "instabil" a)    Abbildung 3.3: Parametersatz "instabil" b)

Die Abbildungen 3.2 und 3.3 bilden eine Realisierung für diesen Parametersatz mit zwei verschiedenen Laufzeiten dar. Auch hier stimmt die Beobachtung mit dem erwarteten Ergebnis überein.

Bei beiden Parametersätzen, lag der entscheidende Term  $\log \rho$  nahe bei 1. Die Wahl der Realisierungen diente dazu die Aufmerksamkeit darauf zu lenken, dass die Stabilitätsbedingung  $\log \rho < 1$  scharf ist.

Die Grafiken wurden mit dem Intel Fortran Compiler 10.0, der IMSL Library 6.0 und OriginPro 7G erzeugt.



# Symbolverzeichnis

$\mathbb{N}$	Menge der positiven natürlichen Zahlen 1,2,...
$\mathbb{R}$	Menge der reellen Zahlen
$\mathbb{R}_0$	Menge der reellen Zahlen ohne 0
$\mathbb{Z}$	Menge der ganzen Zahlen
$\mathcal{B}(S)$	$\sigma$ -Algebra über S
$\mathcal{B}(\mathbb{R}^p)$	Borel- $\sigma$ -Algebra über $\mathbb{R}^p$
$\mathbf{1}_A$	Indikatorfunktion der Menge A
$\delta_x$	Dirac-Maß im Punkt x
$\lambda$	Lebesgue-Maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
$\lambda_p$	Lebesgue-Maß auf $(\mathbb{R}^p, \mathcal{B}(\mathbb{R}^p))$
$\lambda_\Theta$	Oberflächenmaß auf $\Theta$
$\eta_A$	Anzahl der Besuche der Menge A, $\eta_A = \sum_{n=1}^{\infty} \mathbf{1}_{\{\Phi_n \in A\}}$
$\tau_A$	Rückkehrzeit zur Menge A, $\tau_A = \min\{n \geq 1, \Phi_n \in A\}$
$\sigma_A$	erste Eintrittszeit in die Menge A, $\sigma_A = \min\{n \geq 0, \Phi_n \in A\}$
$U(x, A)$	erwartete Anzahl der Besuche von A bei Start in x, $U(x, A) = E_x[\eta_A]$
$L(x, A)$	Rückkehrwahrscheinlichkeit zur Menge A, $L(x, A) = P_x(\tau_A < \infty)$
$Q(x, A)$	Wahrscheinlichkeit, dass $\Phi$ die Menge A unendlich oft besucht, $Q(x, A) = P_x(\Phi \in A \text{ unendlich oft})$
$x \rightarrow y$	y ist von x aus erreichbar ( $\Leftrightarrow L(x, y) > 0$ )
$x \leftrightarrow y$	x kommuniziert mit y ( $\Leftrightarrow x \rightarrow y$ und $y \rightarrow x$ )
$C(x)$	Äquivalenzklasse der mit x kommunizierenden Zustände
$\varphi$	Irreduzibilitätsmaß
$\psi$	Irreduzibilitätsmaß mit Maximalitätseigenschaft
$d$	Periode eines Zustands bzw. einer Äquivalenzklasse C
$\pi$	stationäres Maß
$\Delta V(x)$	Drift von $\Phi$ , $\Delta V(x) = E_x[V(\Phi_1) - V(\Phi_0)]$
$\ \mu\ $	totale Variation Norm eines Maßes $\mu$ , $\ \mu\  = \sup_{A \in \mathcal{B}(S)}  \mu(A)  - \inf_{A \in \mathcal{B}(S)}  \mu(A) $

## Symbolverzeichnis

$\mu(f)$	$\mu(f) = \int_S f(s)\mu(ds)$
$\ \mu\ _f$	$f$ -Norm eines Maßes $\mu$ , $\ \mu\ _f = \sup_{g,  g  \leq f}  \mu(g) $
$\ P_1(x, \cdot) - P_2(x, \cdot)\ $	$V$ -Norm Abstand zweier Übergangsdichten, $\ P_1(x, \cdot) - P_2(x, \cdot)\ _V = \sup_{x \in S} \frac{\ P_1(x, \cdot) - P_2(x, \cdot)\ _V}{V(x)}$ .
$o(\ x\ ) = f(x)$	$\Leftrightarrow \lim_{\ x\  \rightarrow \infty} \frac{f(x)}{\ x\ } = 0$
$H_0$	Achsenflächen des $\mathbb{R}^p$ , $H_0 = \{x \in \mathbb{R}^p \mid \min_i x_i = 0\}$
$H_j$	homogene Hyperflächen in $\mathbb{R}^p$
$\mathbb{R}_\times^p$	$\mathbb{R}_\times^p = \mathbb{R}^p \setminus H_0$
$\Theta$	$\Theta = \{\theta \in \mathbb{R}^p, \ \theta\  = 1\}$
$\Theta_\times$	$\Theta_\times = \Theta \setminus H_0$
$\Theta_{\times \times}$	$\Theta_{\times \times} = \{\theta \in \Theta_\times, F^k(\theta) \notin H_j, \forall j, k\}$
$w(x, u)$	$w(x, u) = (a(x) + b(x)u, x_1, \dots, x_{p-1})$
$w^*(x, u)$	$w^*(x, u) = (a^*(x) + b^*(x)u, x_1, \dots, x_{p-1})$

# Literaturverzeichnis

- [1] Alsmeyer, G. (2005): Stochastische Prozesse, Teil 1, Skripten zur Mathematischen Statistik Nr. 33, Münster
- [2] Arnold, L. (1998): Random Dynamical Systems, Springer Monographs in Mathematics
- [3] Borkovec, M. und Klüppelberg, M. (2001): The tail of the stationary distribution of an autoregressive process with ARCH(1) errors, The Annals of Applied Probability, Vol.11, No. 4, 1220-1241
- [4] Cline, D.B.H. und Pu, H.H. (1998): Verifying irreducibility and continuity of a nonlinear time series, Statistics & Probability Letters 40, 139-148
- [5] Cline, D.B.H. und Pu, H.H. (2001): Geometric Transience of nonlinear time series, Statistica Sinica 11, 237-287
- [6] Cline, D.B.H. und Pu, H.H. (2004): Stability and the Lyapounov exponent of threshold ARCH-models, The Annals of Applied Probability, Vol. 14, No. 4, 1920-1949
- [7] Embrechts, P., Klüppelberg, C. und Mikosch, T. (2003): Modelling Extremal Events for Insurance and Finance, Applications of Mathematics, Springer-Verlag
- [8] Engle, R. F. (1995): ARCH Selected Readings, Advanced Texts in Econometrics
- [9] Gantert, N. (2007): Skript zu Markovketten und Potentialtheorie, Münster
- [10] Gantert, N. (2004): Skript zur Wahrscheinlichkeitstheorie I, Münster
- [11] Gantert, N. (2005): Skript zur Wahrscheinlichkeitstheorie II, Münster
- [12] Meyn, S.P. und Tweedie, R.L. (1993): Markov Chains and Stochastic Stability, Springer-Verlag
- [13] Tong, H. (1990): Non-linear Time series, Oxford Statistical Science Series



Münster, den 30. September 2008

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.