

# Stochastische Modelle in der Biologie

VON PETER PFAFFELHUBER

Version: 21. Januar 2018

## Inhaltsverzeichnis

<b>0</b>	<b>Vorbemerkung</b>	<b>2</b>
<b>1</b>	<b>Chemische Reaktionsnetzwerke</b>	<b>2</b>
1.1	Stochastische Modellierung . . . . .	3
	<a href="#">Wiederholung zu Markov-Prozessen</a> . . . . .	3
1.2	Deterministische Modellierung . . . . .	9
1.3	Stationäre Verteilungen . . . . .	9
	<a href="#">Stationäre Verteilungen von Markov-Prozessen</a> . . . . .	10
1.4	Konvergenz von stochastischen zu deterministischen Dynamiken . . . . .	14
	<a href="#">Konvergenz des reskalierten Poisson-Prozesses</a> . . . . .	14
	<a href="#">Einiges über Martingal-Probleme</a> . . . . .	16
<b>2</b>	<b>Populationsgenetik</b>	<b>23</b>
2.1	Das neutrale Moran-Modell mit zwei Typen . . . . .	24
	<a href="#">Geburts-Todes-Prozesse</a> . . . . .	26
2.2	Der Coaleszent . . . . .	29
2.3	Mutationsmodelle . . . . .	32
2.4	Selektion . . . . .	36
	<a href="#">Ein-dimensionale Diffusionen</a> . . . . .	37
<b>3</b>	<b>Neurowissenschaft</b>	<b>46</b>
3.1	Modelle eines einzelnen Neurons . . . . .	47
	<a href="#">Der Ornstein-Uhlenbeck-Prozess</a> . . . . .	48
	<a href="#">Treffzeiten der Brown'schen Bewegung</a> . . . . .	50
3.2	Modelle mehrerer gekoppler Neuronen . . . . .	51

## 0 Vorbemerkung

**Notation 0.1.** Sei  $(E, r)$  ein vollständiger und separabler metrischer Raum. Wir bezeichnen mit  $\mathcal{M}(E)$  die Menge der messbaren, reellwertigen Funktionen auf  $E$ , mit  $\mathcal{B}(E)$  die Teilmenge der beschränkten, messbaren Funktionen, mit  $\mathcal{C}_b(E)$  die Teilmenge der stetigen, beschränkten Funktionen und mit  $\mathcal{C}_B^k(E)$  die Teilmenge der  $k$ -mal stetig differenzierbaren Funktionen mit beschränkten Ableitungen.

## 1 Chemische Reaktionsnetzwerke

Chemische Gleichungen schreibt man oft in der Form



was soviel bedeutet wie die Reaktion eines Moleküls der Sorte  $A$  und eines der Sorte  $B$  zu einem Molekül der Sorte  $C$ . Kombiniert man mehrere solcher Gleichungen, spricht man von einem chemischen Reaktionsnetzwerk.

**Definition 1.1** (Chemisches Reaktionsnetzwerk). *Ein chemisches Reaktionsnetzwerk ist ein Tripel  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$ , wobei*

- $\mathcal{S} = \{S_1, \dots, S_n\}$  die Menge der Molekülsorten,
- $\mathcal{C}$  die Menge der im Reaktionsnetzwerk vorkommenden Komplexe (d.h. Linearkombination mit positiven, ganzzahligen Koeffizienten aus Molekülsorten) und
- $\mathcal{R}$  die Menge der im Reaktionsnetzwerk vorkommenden Reaktionen  $\nu_k \rightarrow \nu'_k$  zwischen den Komplexen

ist. Für eine Reaktion  $\nu_k \rightarrow \nu'_k$  ist  $\nu_k$  der Edukt-Komplex und  $\nu'_k$  der Produkt-Komplex. Wir bezeichnen mit

$$\zeta_k = \nu'_k - \nu_k$$

den  $k$ -ten Reaktionsvektor des Reaktionsnetzes,  $k \in \mathcal{R}$ .

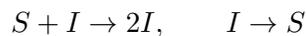
Die Ordnung von  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ist  $\max\{|\nu_k| : k = 1, \dots, |\mathcal{R}|\}$ , also die maximale Anzahl an Molekülen, die miteinander reagieren.

**Beispiel 1.2.** 1. In obigem Beispiel ist

$$\mathcal{S} = \{A, B, C\}, \quad \mathcal{C} = \{A + B, C\}, \quad \mathcal{R} = \{A + B \rightarrow C\}.$$

Dies ist ein Netzwerk zweiter Ordnung.

2. In der Tat lassen sich viele stochastische Modelle in obiger Formulierung beschreiben, auch wenn sie gar nichts mit Chemie zu tun haben: Das Reaktionsnetzwerk



modelliert infizierte ( $I$ ) und suszeptible ( $S$ ) Personen einer Population konstanter Größe. Im Abschnitt 2 werden wir das Netzwerk



kennenlernen, das die Ausbreitung zweier Allele  $A$  und  $B$  in einer Population modelliert.

## 1.1 Stochastische Modellierung

Unser Ziel ist es nun, die Dynamik der Molekülzahl der einzelnen Substanzen zu studieren. Hier ist es wichtig, dass die Dynamik (der Anzahl an Molekülen) eines chemisches Reaktionsnetzwerk noch nicht durch das Tripel  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  festgelegt ist. Vielmehr muss man – im Fall einer stochastischen, Markov’schen Modellierung – für jeden Zustand  $x \in \mathbb{Z}_+^{\mathcal{S}}$  Raten angeben, mit denen die Reaktionen ausgeführt werden.

**Definition 1.3** (Stochastische Dynamik des chemischen Reaktionsnetzwerks). *Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und<sup>1</sup>  $\lambda_k : \mathbb{Z}_+^{\mathcal{S}} \rightarrow \mathbb{R}_+$ ,  $k \in \mathcal{R}$ . Ein Markov-Sprung-Prozess  $X = (X_t)_{t \geq 0}$  mit Zustandsraum  $\mathbb{Z}_+^{\mathcal{S}}$ , der gegeben  $X_t = x$  mit Rate  $\lambda_k(x)$  nach  $x + \zeta_k$  springt, heißt Dynamik des stochastischen Reaktionsnetzwerks mit Raten  $\lambda_k, k \in \mathcal{R}$ .*

### Wiederholung zu Markov-Prozessen

Zunächst jedoch wiederholen wir ein paar Grundlagen zu stochastischen Prozessen, insbesondere zu Markov-Sprung-Prozessen. Wir beginnen mit dem einfachsten solchen Prozess.

**Definition 1.4** (Poisson-Prozess). *Ein Poisson-Prozess mit Rate  $\lambda$  ist der einzige stochastische Prozess  $Y = (Y_t)_{t \geq 0}$  mit  $Y_0 = 0$ , unabhängigen Inkrementen (d.h.  $(Y_{t_1} - Y_{t_0}, \dots, Y_{t_n} - Y_{t_{n-1}})$  ist für jede Wahl von  $t_0 \leq \dots \leq t_n$  eine unabhängige Familie) und  $Y_t - Y_s \sim \text{Poi}(\lambda(t - s))$ .*

Aus den stochastischen Prozessen ist bekannt, dass die Wartezeit zwischen zwei Sprüngen im Poisson-Prozess exponentialverteilt ist. Genauer gilt folgendes.

**Proposition 1.5** (Konstruktion mittels Exponentialverteilungen). *Seien  $S_1, S_2, \dots$  unabhängig, exponentialverteilt mit Parameter  $\lambda$ . Weiter sei  $X = (X_t)_{t \geq 0}$  gegeben durch*

$$X_t := \max\{i : S_1 + \dots + S_i \leq t\}$$

mit  $\max \emptyset = 0$ . Dann ist  $X$  ein Poisson-Prozess mit Intensität  $\lambda$ .

Der Poisson-Prozess ist der einfachste Markov’sche Sprungprozess.

**Definition 1.6** (Markov’scher Sprungprozess). *Sei  $\lambda : \mathbb{Z}^d \rightarrow \mathbb{R}_+$  und sei  $\mu$  ein Markov’scher Übergangskern auf  $\mathbb{Z}^d$ . Sei  $X = (X_t)_{t \geq 0}$  ein  $\mathbb{Z}^d$ -wertiger Prozess, der wie folgt definiert ist. Gegeben  $X_t = x$ , warte eine  $\exp(\lambda(x))$ -verteilte Zeit und springe dann gemäß  $\mu(x, \cdot)$  zu einem neuen Ort. Dann heißt  $X$  Markov-Sprung-Prozess mit Sprungraten  $\lambda$  und Übergangskern  $\mu$ .*

Folgendes ist aus den stochastischen Prozessen bekannt.

**Proposition 1.7** (Äquivalente Formulierung für Markov-Sprung-Prozesse). *Seien  $\lambda_1, \dots, \lambda_n : \mathbb{Z}^d \rightarrow \mathbb{R}_+$  und  $\zeta_1, \dots, \zeta_n \in \mathbb{Z}^d$ . Sei  $X = (X_t)_{t \geq 0}$  ein  $\mathbb{Z}^d$ -wertiger Prozess, der wie folgt definiert ist. Gegeben  $X_t = x$ , seien  $S_k \sim \exp(\lambda_k(x))$ ,  $k = 1, \dots, n$  und  $S = \min S_k$  und  $K = \operatorname{argmin} S_k$ . Dann springt der Prozess zur Zeit  $S$  nach  $x + \zeta_K$ .*

Sei weiter

$$\mu(x, y) = \begin{cases} \frac{\lambda_k(x)}{\sum_j \lambda_j(x)}, & y = x + \zeta_k, \\ 0, & \text{sonst} \end{cases}$$

<sup>1</sup>Im Folgenden identifizieren wir die Reaktion  $\nu_k \rightarrow \nu'_k$  mit  $k$  und schreiben dann etwa  $k \in \mathcal{R}$ .

ein Markov-Übergangskern. Dann ist  $X$  ein Markov-Sprungprozess mit Sprungraten  $\lambda_1 + \dots + \lambda_n$  und Übergangskern  $\mu$ .

Oftmals ist es für Markov-Prozesse am einfachsten, die infinitesimalen Charakteristiken des Prozesses anzugeben, um ihn zu definieren. Dies geschieht immer durch Angabe des Generators.

**Definition 1.8** (Generator). Sei  $X = (X_t)_{t \geq 0}$  ein  $\mathbb{R}^d$ -wertiger Markov-Prozess und  $f \in \mathcal{M}(\mathbb{R}^d)$ . Wir definieren, falls existent,

$$Gf(x) := \frac{1}{h}(\mathbf{E}_x[f(X_h)] - f(x)).$$

Dann heißt  $G$  Generator von  $X$ . Die Menge aller  $f$ , für die der Grenzwert überall existiert, heißt Domain von  $G$  oder  $\mathcal{D}(G)$ . Manchmal definieren wir  $G$  auch nur auf einer Teilmenge  $\mathcal{D} \subseteq \mathcal{D}(G)$ .

Der Generator gibt die Dynamik von  $f(X)$  im Mittel an. Zieht man dieses mittlere Verhalten von der Dynamik von  $f(X)$  ab, so erhält man typischerweise ein Martingal.

**Theorem 1.9** (Martingal-Problem). Sei  $X$  ein Markov-Prozess mit Zustandsraum  $\mathbb{Z}^d$  und Generator  $G$ . Für jedes  $f \in \mathcal{D}(G) \cap \mathcal{B}(\mathbb{R}^d)$  mit  $Gf \in \mathcal{B}(\mathbb{R}^d)$  ist

$$\left( f(X_t) - f(X_0) - \int_0^t Gf(X_s) ds \right)_{t \geq 0}$$

ein Martingal.

**Proposition 1.10** (Generator für Markov-Sprung-Prozesse). 1. Sei  $X$  ein Markov-Sprungprozess mit Sprungrate  $\lambda : \mathbb{Z}^d \rightarrow \mathbb{R}_+$  und Übergangskern  $\mu$  wie in Definition 1.6. Dann ist für  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$

$$Gf(x) = \lambda(x) \int (f(y) - f(x)) \mu(x, dy)$$

der Generator von  $X$ .

2. Seien  $\lambda_1, \dots, \lambda_n, \zeta_1, \dots, \zeta_n, X$  wie in Proposition 1.7 und  $X$  der dort konstruierte Markov-Sprung-Prozess. Dann ist für  $f : \mathbb{Z}^d \rightarrow \mathbb{R}$

$$Gf(x) = \sum_{k=1}^n \lambda_k(x) (f(x + \zeta_k) - f(x))$$

der Generator von  $X$ .

Für die stochastische Dynamik  $X$  eines chemischen Reaktionsnetzwerkes finden wir nun äquivalente Beschreibungen. Die Gleichung in 1. wird oft auch Master-Gleichung genannt.

**Theorem 1.11** (Zeit-Transformation, Master-Gleichung, Martingal-Problem). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $(\lambda_k)_{k \in \mathcal{R}}$  mit  $\lambda_k : \mathbb{Z}_+^n \rightarrow \mathbb{R}_+$  beschränkt,  $k \in \mathcal{K}$ .

Weiter sei  $X = (X_t)_{t \geq 0}$  die stochastische Dynamik von  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  mit Raten  $\{\lambda_k : k \in \mathcal{R}\}$ .

1. Es gilt

$$\frac{d}{dt}\mathbf{P}(X_t = x) = \sum_{k \in \mathcal{K}} \lambda_k(x - \zeta_k)\mathbf{P}(X_t = x - \zeta_k) - \lambda_k(x)\mathbf{P}(X_t = x). \quad (1.1)$$

2. Für alle  $f \in \mathcal{B}(\mathbb{Z}_+^n)$  ist

$$f(X_t) - \int_0^t Gf(X_s)ds$$

ein Martingal, wobei

$$Gf(x) = \sum_{k \in \mathcal{K}} \lambda_k(x)(f(x + \zeta_k) - f(x)).$$

3.  $X$  erfüllt die Gleichung

$$X_t = X_0 + \sum_{k \in \mathcal{K}} \zeta_k Y_k \left( \int_0^t \lambda_k(X_s)ds \right)$$

für unabhängige Poisson-Prozesse  $(Y_k)_{k \in \mathcal{K}}$  mit Rate 1.

Ist andersherum  $X = (X_t)_{t \geq 0}$  ein stochastischer Prozess, der 1., 2. oder 3. erfüllt, dann ist  $X$  ein Markov-Sprungprozess mit Zustandsraum  $\mathbb{Z}_+^S$ , der gegeben  $X = x$  mit Rate  $\lambda_k(x)$  zu  $x + \zeta_k$  springt,  $k \in \mathcal{K}$ .

*Beweis.* 1. Wir argumentieren etwas heuristisch. Sei hierfür  $h$  eine kleine Zahl. Dann ist

$$\mathbf{P}(X_{t+h} = x) = \mathbf{P}(X_t = h) \left( 1 - \sum_{k \in \mathcal{K}} \lambda_k(x)h \right) + \sum_{k \in \mathcal{K}} \mathbf{P}(X_t = x - \zeta_k) \lambda_k(x - \zeta_k) + O(h^2),$$

wobei  $O(h^2)$  alle Terme beinhaltet, bei denen mindestens zwei Sprünge zwischen  $t$  und  $t+h$  passiert sein müssen. Die Wahrscheinlichkeit hierfür ist bekanntlich  $O(h^2)$ . Stellt man diese Gleichung nun um und betrachtet  $h \rightarrow 0$ , so folgt direkt 1.

2. Wir schreiben zunächst

$$\begin{aligned} \mathbf{E} \left[ f(X_t) - \int_0^t Gf(X_r)dr \middle| \mathcal{F}_s \right] &= f(X_s) - \int_0^s Gf(X_r)dr \\ &\quad + \mathbf{E} \left[ f(X_t) - f(X_s) - \int_s^t Gf(X_r)dr \middle| X_s \right]. \end{aligned}$$

Da  $X$  homogen ist, müssen wir also zeigen, dass

$$\mathbf{E} \left[ f(X_t) - f(X_0) - \int_0^t Gf(X_r)dr \middle| X_0 = x \right] = 0$$

für alle  $t$ . Für  $t = 0$  ist diese Behauptung klar. Betrachten wir zunächst eine Indikatorfunktion  $f(x) = 1_y(x)$ . Dann ist

$$G1_y(x) = \sum_{k \in \mathcal{K}} \lambda_k(x)(1_y(x + \zeta_k) - 1_y(x)) = \sum_{k \in \mathcal{K}} \lambda_k(x)(1_{y-\zeta_k}(x) - 1_y(x)).$$

Damit ist

$$\begin{aligned}
& \frac{d}{dt} \mathbf{E} \left[ 1_y(X_t) - 1_y(X_0) - \int_0^t G 1_y(X_r) dr \mid X_0 = x \right] \\
&= \frac{d}{dt} \mathbf{P}(X_t = y) - \mathbf{E}[G 1_y(X_t)] \\
&= \frac{d}{dt} \mathbf{P}(X_t = y) - \mathbf{E} \left[ \sum_{k \in \mathcal{K}} \lambda_k(X_t) (1_{y-\zeta_k}(X_t) - 1_y(X_t)) \right] \\
&= \frac{d}{dt} \mathbf{P}(X_t = y) - \mathbf{E} \left[ \sum_{k \in \mathcal{K}} \lambda_k(y - \zeta_k) (1_{y-\zeta_k}(X_t) - 1_y(X_t)) \right] \\
&= \frac{d}{dt} \mathbf{P}(X_t = y) - \sum_{k \in \mathcal{K}} \lambda_k(y - \zeta_k) (\mathbf{P}(X_t = y - \zeta_k) - \mathbf{P}(X_t = y)) = 0
\end{aligned}$$

nach 1. Allgemeine, beschränkte Funktionen  $f : \mathbb{Z}_+^n \rightarrow \mathbb{R}$  stellt man als Summe von Indikatorfunktionen dar. Nun ist  $G$  linear, und  $f(X_t) - \int_0^t Gf(X_s) ds$  ist als beschränkte Summe von Martingalen ein Martingal.

3. Wir konstruieren den behaupteten Zusammenhang von Sprung zu Sprung. Ist  $X_t = x$ , so ist die Wartezeit bis zum nächsten Sprung von  $Y_k \left( \int_0^\cdot \lambda_k(X_s) ds \right)$  exponentialverteilt mit Rate  $\lambda_k(x)$ ,  $k \in \mathcal{K}$ . Daraus folgt bereits die Behauptung.

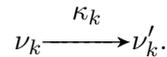
Um die letzte Behauptung zu zeigen, bemerken wir zunächst, dass 3. genau den Markov-Sprungprozess beschreibt, der mit Rate  $\lambda_k(x)$  von  $x$  nach  $x + \zeta_k$  springt. Weiter ist 1. ein lineares System gewöhnlicher Differentialgleichungen, das also eine eindeutige Lösung hat. Da der Markov-Sprungprozess dieses System löst, muss  $X$  dieselbe Verteilung wie dieser Sprungprozess haben. Schließlich haben wir gesehen, dass 2. im Spezialfall von Indikatorfunktionen äquivalent ist zu 1., so dass aus 2. ebenfalls folgt, dass  $X$  der Markov-Sprungprozess ist.  $\square$

Die Funktionen  $\lambda_k$  werden oftmals durch das Massenwirkungsgesetz gegeben.

**Definition 1.12** (Stochastisches Massenwirkungsgesetz). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk. Ist die Dynamik durch Funktionen

$$\lambda_k(x) = \kappa_k \prod_{s \in \mathcal{S}} \nu_{ks}! \binom{x_s}{\nu_{ks}}$$

mit  $\kappa_k > 0$ ,  $k \in \mathcal{R}$  gegeben, so spricht man davon, dass das chemische Reaktionsnetzwerk dem Massenwirkungsgesetz<sup>2</sup> mit Reaktionsraten  $\kappa_k$ ,  $k \in \mathcal{R}$  folgt. In diesem Fall schreiben wir die  $k$ -te Reaktion als



**Beispiel 1.13** ( $M/M/\infty$ -Warteschlange). Wir betrachten ein sehr einfaches Beispiel



<sup>2</sup>Auf englisch: *law of mass action*

Hier gibt es also nur  $\mathcal{S} = \{A\}, \mathcal{C} = \{\emptyset, A\}$  und zwei Reaktionen, wobei  $\lambda_1(x) = \lambda$  und  $\lambda_2(x) = \rho x$ . Nach Theorem 1.11 kann man also

$$X_t = x_0 + Y_1(\lambda t) - Y_2\left(\int_0^t \rho X_s ds\right)$$

schreiben. In diesem Beispiel kann man die Verteilung  $X_t$ , der Anzahl an Molekülen der Sorte  $A$  zur Zeit  $t$ , explizit berechnen. Stellen wir hierzu zunächst eine heuristische Überlegung an. Ist anfangs eine Anzahl  $x_0$  Moleküle vorhanden, so ist jedes dieser Moleküle mit Wahrscheinlichkeit  $e^{-\rho t}$  zur Zeit  $t$  noch nicht zu  $\emptyset$  reagiert. Andererseits ist die Wahrscheinlichkeit, dass zwischen  $s$  und  $s + ds$  ein neues Molekül entstanden ist, gerade  $\lambda ds$ . Dieses Molekül ist dann mit Wahrscheinlichkeit  $e^{-\rho(t-s)}$  zur Zeit  $t$  noch nicht wieder zu  $\emptyset$  reagiert. Da die Wahrscheinlichkeit, dass zwischen  $s$  und  $s + ds$  ein Molekül entsteht, sehr klein ist, es jedoch viele solcher kleiner Intervalle gibt, ist die Gesamtzahl an entstehenden Molekülen Poisson-verteilt mit Parameter  $\lambda t$ . Betrachtet man nur die Moleküle, die zur Zeit  $t$  noch da sind, muss man diese Poisson-Verteilung ausdünnen, was bekanntlich wieder zu einer Poisson-Verteilung führt. Der Parameter ist dann

$$\int_0^t e^{-\rho(t-s)} \lambda ds = \frac{\lambda}{\rho} (1 - e^{-\rho t}).$$

Deshalb sollte  $X_t \sim Y + Z$  gelten, wobei  $Y \sim B(x, e^{-\rho t})$  und  $Z \sim \text{Poi}\left(\frac{\lambda}{\rho}(1 - e^{-\rho t})\right)$  unabhängig sind, also für  $\theta_t := \frac{\lambda}{\rho}(1 - e^{-\rho t})$

$$\mathbf{P}(X_t = k) = \frac{\lambda^k}{\rho^k} \sum_{i=0}^k \binom{x}{i} \left(\frac{\rho}{\lambda} e^{-\rho t}\right)^i (1 - e^{-\rho t})^{x+k-2i} e^{-\theta_t} \frac{1}{(k-i)!}. \quad (1.2)$$

Alternativ berechnen wir die charakteristische Funktion von  $X_t$  durch

$$\mathbf{E}[s^Y] = \sum_{i=0}^x s^i \binom{x}{i} e^{-\rho t i} (1 - e^{-\rho t})^{x-i} = \sum_{i=0}^x \binom{x}{i} (s e^{-\rho t})^i (1 - e^{-\rho t})^{x-i} = (1 - e^{-\rho t} + s e^{-\rho t})^x,$$

$$\mathbf{E}[s^Z] = \sum_{i=0}^{\infty} s^i e^{-\theta_t} \frac{\theta_t^i}{i!} = e^{-\theta_t} e^{s \theta_t} = e^{-\theta_t(1-s)},$$

$$\mathbf{E}[s^{X_t}] = \mathbf{E}[s^Y] \mathbf{E}[s^Z] = (1 - (1-s)e^{-\rho t})^x e^{-\theta_t(1-s)}.$$

Multipliziert man andererseits (1.4) mit  $s^k$  und summiert über  $k$ , entsteht, mit  $\zeta_1 = 1, \zeta_2 = -1$ ,

$$\begin{aligned} \frac{d}{dt} \mathbf{E}[s^{X_t}] &= \lambda \sum_{k=1}^{\infty} s^k \mathbf{P}(X_t = k-1) - \lambda \sum_{k=0}^{\infty} s^k \mathbf{P}(X_t = k) + \rho \sum_{k=0}^{\infty} (k+1) s^k \mathbf{P}(X_t = k+1) - k s^k \mathbf{P}(X_t = k) \\ &= -\lambda(1-s) \mathbf{E}[s^{X_t}] + \rho(1-s) \frac{d}{ds} \mathbf{E}[s^{X_t}]. \end{aligned}$$

Nun ist

$$\begin{aligned} \frac{d}{dt} \mathbf{E}[s^{X_t}] &= \mathbf{E}[s^{X_t}] \left( \frac{x(1-s)\rho e^{-\rho t}}{1 - (1-s)e^{-\rho t}} - \lambda(1-s)e^{-\rho t} \right) \\ \frac{d}{ds} \mathbf{E}[s^{X_t}] &= \mathbf{E}[s^{X_t}] \left( \frac{x e^{-\rho t}}{1 - (1-s)e^{-\rho t}} + \frac{\lambda}{\rho} (1 - e^{-\rho t}) \right), \end{aligned}$$

also

$$\frac{d}{dt} \mathbf{E}[s^{X_t}] - \rho(1-s) \frac{d}{ds} \mathbf{E}[s^{X_t}] = -\lambda(1-s) \mathbf{E}[s^{X_t}]$$

Damit erfüllt die Verteilung (1.2) die Master-Gleichung und wir haben nochmals die Verteilung von  $X_t$  gezeigt.

Bei dem letzten Beispiel handelt es sich um ein chemisches Reaktionsnetzwerk erster Ordnung. Bei chemischen Reaktionsnetzwerken nullter und erster Ordnung ist es möglich, alle Momente der Verteilung zu berechnen. Der Grund hierfür ist das folgende Resultat.

**Proposition 1.14.** *Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk. Dann gilt für  $f : \mathbb{R}_+^{\mathcal{S}} \rightarrow \mathbb{R}$*

$$\frac{d}{dt} \mathbf{E}[f(X_t)] = \sum_{k \in \mathcal{K}} \mathbf{E}[\lambda_k(X_t)(f(X_t + \zeta_k) - f(X_t))].$$

*Beweis.* Die Aussage folgt direkt aus Theorem 1.11.2.  $\square$

**Beispiel 1.15** (*M/M/ $\infty$ -Warteschlange II*). Betrachten wir nochmals Beispiel 1.13. Es handelt sich hier um eine Netzwerk erster Ordnung. Hier ist nach der letzten Proposition

$$\frac{d}{dt} \mathbf{E}[X_t] = \lambda - \rho \mathbf{E}[X_t].$$

Diese Differentialgleichung ist schnell gelöst. Setzt man  $g(t) = \mathbf{E}[X_t] - \frac{\lambda}{\rho}$ , so ist  $g'(t) = -\rho g(t)$ , also

$$\mathbf{E}[X_t] = \frac{\lambda}{\rho} + \left(x - \frac{\lambda}{\rho}\right) e^{-\rho t} = x e^{-\rho t} + \frac{\lambda}{\rho} (1 - e^{-\rho t}).$$

In der Tat ist dies auch der Erwartungswert der in (1.2) angegebenen Verteilung.

Um das zweite Moment zu berechnen, schreiben wir

$$\begin{aligned} \frac{d}{dt} \mathbf{E}[X_t^2] &= \lambda \mathbf{E}[(X_t + 1)^2 - X_t^2] + \rho \mathbf{E}[X_t((X_t - 1)^2 - X_t^2)] \\ &= (2\lambda + \rho) \mathbf{E}[X_t] + \lambda - 2\rho \mathbf{E}[X_t^2] \end{aligned}$$

oder

$$\frac{d}{dt} \mathbf{Var}[X_t] = \lambda + \rho \mathbf{E}[X_t] - 2\rho \mathbf{Var}[X_t].$$

Nach der allgemeinen Lösungsformel für lineare Differentialgleichungen<sup>3</sup> ist nun

$$\begin{aligned} \mathbf{Var}[X_t] &= e^{-2\rho t} \left( \int_0^t (\lambda + \rho \mathbf{E}[X_s]) e^{2\rho s} ds \right) \\ &= e^{-2\rho t} \int_0^t (2\lambda + (\rho x - \lambda) e^{-\rho s}) e^{2\rho s} ds \\ &= \frac{\lambda}{\rho} (1 - e^{-2\rho t}) + (\rho x - \lambda) \frac{1}{\rho} e^{-\rho t} (1 - e^{-\rho t}) \\ &= x e^{-\rho t} (1 - e^{-\rho t}) + \frac{\lambda}{\rho} (1 - e^{-\rho t}). \end{aligned}$$

<sup>3</sup>Die lineare Differentialgleichung  $\dot{x} = a + bx$  mit glatten Funktionen  $a$  und  $b$  mit  $x_0 = x$  wird eindeutig gelöst von

$$x(t) = e^{\int_0^t b(s) ds} \left( \int_0^t a(s) e^{-\int_0^s b(r) dr} ds + x \right).$$

Dies ist auch die Varianz, die sich für die Verteilung in (1.2) ergibt.

## 1.2 Deterministische Modellierung

Neben der stochastischen Dynamik von chemischen Reaktionsnetzwerken gibt es noch die deterministische Modellierung. Diese kommt häufig dann zum Einsatz, wenn durch eine hohe Teilchenzahl die einzelnen Substanzen kaum schwanken.

**Definition 1.16** (Deterministisches chemisches Reaktionsnetzwerk). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $\lambda_k : \mathbb{R}_+^{\mathcal{S}} \rightarrow \mathbb{R}_+$ ,  $k \in \mathcal{R}$ . Eine Lösung des Systems von Differentialgleichungen

$$\dot{X}_s = \sum_k \zeta_k \lambda_k(X)$$

heißt deterministische Dynamik des Reaktionsnetzwerks.

Ein Spezialfall von stochastischen Dynamiken haben wir in Definition 1.12 als das Massenwirkungsgesetz kennengelernt. Dieses gibt es auch für deterministische Dynamiken.

**Definition 1.17** (Deterministisches Massenwirkungsgesetz). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk. Ist die Dynamik durch Funktionen

$$\lambda_k(x) = \kappa_k \prod_{s \in \mathcal{S}} x_s^{y_{ks}} \quad (1.3)$$

mit  $\kappa_k > 0$ ,  $k \in \mathcal{R}$  gegeben, so spricht man davon, dass das chemische Reaktionsnetzwerk dem Massenwirkungsgesetz mit Reaktionsraten  $\kappa_k$ ,  $k \in \mathcal{R}$  folgt.

**Beispiel 1.18.** Für Beispiel 1.13, also das Reaktionsnetzwerk



sei  $X_A$  die Menge des Stoffes  $A$ . Gilt dann

$$\dot{X}_A = \lambda - \rho X_A,$$

so ist dies die deterministische Dynamik des Reaktionsnetzwerkes, die dem Massenwirkungsgesetz folgt. Diese Differentialgleichung ist übrigens schnell gelöst: Sei  $Y_A = \frac{\lambda}{\rho} - X_A$ . Dann muss  $\dot{Y}_A = -\rho Y_A$ , also  $Y_A(t) = Y_A(0)e^{-\rho t}$  oder

$$X_A(t) = \frac{\lambda}{\rho} - \left( \frac{\lambda}{\rho} - X_A(0) \right) e^{-\rho t} = X_A(0)e^{-\rho t} + \frac{\lambda}{\rho} (1 - e^{-\rho t}).$$

gelten.

## 1.3 Stationäre Verteilungen

Sowohl stochastische als auch deterministische Dynamiken können Gleichgewichte haben. Im stochastischen Fall sind dies Verteilungen, im deterministischen Fall sind dies feste Zustände. In beiden Fällen ändern sich diese durch die Dynamik nicht mehr.

**Definition 1.19** (Gleichgewicht des stochastischen und des deterministischen Systems). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $\lambda_k, k \in \mathcal{R}$  die Reaktionsraten.

1. Jede Verteilung  $\pi$  auf  $\mathbb{Z}_+^{\mathcal{S}}$  mit

$$\sum_{k \in \mathcal{K}} \lambda_k(x - \zeta_k) \pi(x - \zeta_k) - \lambda_k(x) \pi(x) = 0. \quad (1.4)$$

heißt Gleichgewicht der stochastischen Dynamik.

2. Jedes  $c \in \mathcal{R}_+^{\mathcal{S}}$  mit

$$\sum_{k \in \mathcal{R}} \zeta_k \lambda_k(c) = 0.$$

heißt Gleichgewicht der deterministischen Dynamik des chemischen Reaktionsnetzwerkes.

**Beispiel 1.20** ( $M/M/\infty$ -Warteschlange III). Wir haben bereits die Verteilung von  $X_t$  in Beispiel 1.13 berechnet. Für große Zeiten  $t$  wird sich diese Verteilung nicht mehr ändern, also vermuten wir, dass

$$\pi = \text{Poi}\left(\frac{\lambda}{\rho}\right), \text{ also } \pi(k) = e^{-\lambda/\rho} \frac{(\lambda/\rho)^k}{k!}$$

eine stationäre Verteilung für die stochastische Dynamik ist. In der Tat berechnen wir

$$\begin{aligned} & \lambda\pi(k-1) + \rho(k+1)\pi(k+1) - (\lambda + \rho k)\pi(k) \\ &= e^{-\lambda/\rho} \left( \lambda \frac{(\lambda/\rho)^{k-1}}{(k-1)!} + \rho(k+1) \frac{(\lambda/\rho)^{k+1}}{(k+1)!} - (\lambda + \rho k) \frac{(\lambda/\rho)^k}{k!} \right) \\ &= e^{-\lambda/\rho} \left( \frac{\lambda^k}{\rho^{k-1}} \frac{1}{(k-1)!} + \frac{\lambda^{k+1}}{\rho^k} \frac{1}{k!} - \frac{\lambda^{k+1}}{\rho^k} \frac{1}{k!} - \frac{\lambda^k}{\rho^{k-1}} \frac{1}{(k-1)!} \right) = 0. \end{aligned}$$

## Stationäre Verteilungen von Markov-Prozessen

**Proposition 1.21.** Sei  $X$  ein Markov-Prozess mit Zustandsraum  $\mathbb{R}^d$  und Generator  $G$ . Eine Verteilung  $\pi$  auf  $\mathbb{R}^d$  ist genau dann stationär, wenn

$$\mathbf{E}_\pi[Gf(X)] = 0$$

für alle  $f \in \mathcal{D}(G)$  mit  $Gf \in \mathcal{C}_b(\mathbb{R}^d)$ .

**Definition 1.22** (Irreduzibler Markov-Sprung-Prozess). Sei  $X$  ein Markov-Sprungprozess mit Zustandsraum  $E \subseteq \mathbb{Z}^d$ . Der Prozess heißt irreduzibel, wenn für alle  $A \subseteq E$  gilt, dass  $\mathbf{P}(X_t \in A | X_0 = x) < 1$ .

**Bemerkung 1.23** (Zerlegung eines Markov-Sprung-Prozesses). Ist ein Markov-Sprung-Prozesses  $X$  mit Zustandsraum  $E \subseteq \mathbb{Z}^d$  nicht irreduzibel, so gibt es  $E_1, E_2, \dots$ , so dass der in  $E_k$  gestartete Prozess  $X$  (mit Zustandsraum  $E_k$ ) irreduzibel ist. Hier besteht  $E_k$  aus den Zuständen, die von jedem  $x_k \in E_k$  aus erreichbar sind.

**Theorem 1.24** (Markov-Ketten-Konvergenzsatz in stetiger Zeit). Sei  $X = (X_t)_{t \geq 0}$  ein irreduzibler Markov-Sprung-Prozess mit Zustandsraum  $\mathbb{Z}^d$ . Dann gilt genau eine der beiden Aussagen

1. Es gibt genau eine invariante Verteilung  $\pi$  von  $X$ . Ist  $X_0$  beliebig verteilt, so ist

$$\|\mathcal{L}(X_t) - \pi\| \xrightarrow{t \rightarrow \infty} 0. \quad (1.5)$$

In diesem Fall ist

$$\pi(x) = \frac{1}{\lambda(x) \cdot \mathbf{E}[T_x | X_0 = x]},$$

wobei  $T_x$  die erste Sprungzeit von  $x$  ist.

2. Es gibt keine invariante Verteilung von  $X$  und es gilt

$$\mathbf{P}(X_t = y | X_0 = x) \xrightarrow{t \rightarrow \infty} 0 \text{ für alle } x, y \in E. \quad (1.6)$$

**Beispiel 1.25** (Ein chemisches Reaktionsnetzwerk mit einer Erhaltungsgröße). Für chemische Reaktionsnetzwerke treten reduzible (d.h. nicht irreduzible) Markov-Sprung-Prozesse immer dann auf, wenn das System Erhaltungsgrößen hat. Sei hierzu



und  $X_A(t), X_B(t)$  die Anzahlen der Moleküle der Sorten  $A$  und  $B$ . Hier ist  $t \mapsto X_A(t) + X_B(t)$  konstant und  $X$  (mit Zustandsraum  $\mathbb{Z}_+^2$ ) ist reduzibel. Setzen wir  $E_N := \{(x_A, x_B) \in \mathbb{Z}_+^2 : x_A + x_B = N\}$ , so ist der Prozess  $X$  mit Zustandsraum  $E_N$  irreduzibel.

**Definition 1.26** (Stoichiometrische Kompatibilitätsklasse). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk. Für  $x \in \mathbb{Z}_+^{\mathcal{S}}$  heißt

$$E_x := x + \text{span}\{\nu'_k - \nu_k : k \in \mathcal{R}\}$$

stoichiometrische Kompatibilitätsklasse von  $x$ . Mit anderen Worten ist  $E_x$  die Menge der Zustände in  $\mathbb{Z}_+^{\mathcal{S}}$ , die der die stochastische Dynamik des chemischen Reaktionsnetzwerkes erreichen kann.

**Bemerkung 1.27** (Dynamik auf einer stoichiometrischen Kompatibilitätsklasse). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $X$  die stochastische Dynamik mit Raten  $\lambda_k, k \in \mathcal{R}$ . Weiter seien  $E_1, E_2, \dots$  die stoichiometrischen Kompatibilitätsklassen. Dann ist der Markov-Sprung-Prozess auf jedem  $E_N$  irreduzibel. Insbesondere gibt es dann – laut Theorem 1.24 entweder eine (eindeutige) stationäre Verteilung  $\pi$  auf  $E_N$ , oder der Prozess divergiert.

Nun beschäftigen wir uns ausschließlich mit Dynamiken, die dem Massenwirkungsgesetz folgen. Wir bemerken kurz, dass in diesem Fall die Bedingung für eine stationäre Verteilung im stochastischen Fall durch

$$\sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{X}} \nu_{ks}! \binom{x_s - \zeta_{ks}}{\nu_{ks}} \pi(x - \zeta_k) = \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{X}} \nu_{ks}! \binom{x_s}{\nu_{ks}} \pi(x) \quad (1.7)$$

gegeben ist. Im deterministischen Fall ist die Bedingung

$$\sum_{k \in \mathcal{R}} \zeta_k \kappa_k \prod_{s \in \mathcal{S}} x_s^{\nu_{ks}} = 0,$$

die man auch als

$$\sum_{\eta \in \mathcal{C}} \eta \sum_{k \in \mathcal{R}: \nu_k = \eta} \kappa_k \prod_{s \in \mathcal{S}} x_s^{\nu_{ks}} = \sum_{\eta \in \mathcal{C}} \eta \sum_{k \in \mathcal{R}: \nu'_k = \eta} \kappa_k \prod_{s \in \mathcal{S}} x_s^{\nu'_{ks}}$$

schreiben kann. In der Tat gibt es in diesem Fall eine wichtige Klasse von stationären Punkten für deterministisch Dynamiken von chemischen Reaktionsnetzwerken.

**Definition 1.28** (Komplex-balanzieretes Gleichgewicht). *Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $X$  dessen Dynamik, die sich deterministisch mit  $\lambda_k, k \in \mathcal{R}$  wie in (1.3) entwickelt. Gilt für alle  $\eta \in \mathcal{C}$*

$$\sum_{\{k \in \mathcal{R}: \nu_k = \eta\}} \kappa_k \prod_{s \in \mathcal{S}} x_s^{\nu_{ks}} = \sum_{\{k \in \mathcal{R}: \nu'_k = \eta\}} \kappa_k \prod_{s \in \mathcal{S}} x_s^{\nu'_{ks}},$$

so heißt  $x$  komplex-balanziert.

**Theorem 1.29** (Produkt-Form stationäre Verteilung). *Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk, das sich nach dem Massenwirkungsgesetz mit Raten  $\kappa_k, k \in \mathcal{R}$ , entwickelt. Angenommen,  $c \in \mathbb{R}_{>0}^{\mathcal{S}}$  ist ein komplex-balanzieretes Gleichgewicht für die deterministische Dynamik, dann hat die stochastische Dynamik eine stationäre Verteilung der Form*

$$\pi(x) = \prod_{s \in \mathcal{S}} \frac{c_s^{x_s}}{x_s!} e^{-c_s}, \quad x \in \mathbb{Z}_{\geq 0}^{\mathcal{S}}. \quad (1.8)$$

*Beweis.* Wir müssen zeigen, dass die angegebene Verteilung die Gleichung (1.7) löst, also müssen wir

$$\sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \nu_{ks}! \binom{x_s - \zeta_{ks}}{\nu_{ks}} \frac{c_s^{x_s - \zeta_{ks}}}{(x_s - \zeta_{ks})!} e^{-c_s} = \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \nu'_{ks}! \binom{x_s}{\nu'_{ks}} \frac{c_s^{x_s}}{x_s!} e^{-c_s},$$

d.h.

$$\begin{aligned} \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \nu_{ks}! \binom{x_s - \zeta_{ks}}{\nu_{ks}} \frac{c_s^{-\zeta_{ks}}}{(x_s - \zeta_{ks})!} &= \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \nu_{ks}! \binom{x_s}{\nu_{ks}} \frac{1}{x_s!}, \\ \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \frac{1}{(x_s - \nu'_{ks})!} c_s^{-\nu'_{ks} + \nu_{ks}} &= \sum_{k \in \mathcal{R}} \kappa_k \prod_{s \in \mathcal{S}} \frac{1}{(x_s - \nu_{ks})!} \end{aligned}$$

zeigen. Die letzte Gleichung legt schon eine Symmetrie zwischen Edukt- und Produkt-Komplexen. Um diese Gleichheit zu sehen, schreiben wir statt  $\sum_{k \in \mathcal{R}}$  auf der linken Seite die Summe  $\sum_{\eta \in \mathcal{C}} \sum_{k \in \mathcal{R}: \nu'_k = \eta}$  beziehungsweise  $\sum_{\eta \in \mathcal{C}} \sum_{k \in \mathcal{R}: \nu_k = \eta}$  auf der rechten Seite. Da zwei Summen  $\sum_{\eta \in \mathcal{C}} \dots$  insbesondere dann gleich sind, wenn alle Summanden gleich sind, fixieren wir (neben  $x \in \mathbb{Z}_+^{\mathcal{S}}$ ) den Komplex  $\eta \in \mathcal{C}$  und behaupten

$$\sum_{k \in \mathcal{R}: \nu'_k = \eta} \kappa_k \prod_{s \in \mathcal{S}} \frac{1}{(x_s - \eta_s)!} c_s^{-\eta_s + \nu_{ks}} = \sum_{k \in \mathcal{R}: \nu_k = \eta} \kappa_k \prod_{s \in \mathcal{S}} \frac{1}{(x_s - \eta_s)!}.$$

Da hier  $\eta$  und  $x$  fixiert sind, können wir beide Seiten mit  $\prod c_s^{\eta_s}$  multiplizieren und erhalten äquivalent

$$\sum_{k \in \mathcal{R}: \nu'_k = \eta} \kappa_k c_s^{\nu_{ks}} = \sum_{k \in \mathcal{R}: \nu_k = \eta} \kappa_k c_s^{\eta_s},$$

also der komplexen Balanziertheit des Gleichgewichts  $c$ , die ja vorausgesetzt war.  $\square$

**Beispiel 1.30** ( $M/M/\infty$ -Warteschlange III). Wieder betrachten wir Beispiel 1.13. Wir wissen bereits aus Beispiel 1.20, dass die Poisson-Verteilung mit Parameter  $\lambda/\rho$  die stationäre Verteilung ist. (Es kann keine weitere geben, da es nur eine einzige stochiometrische Kompatibilitätsklasse gibt; siehe Theorem 1.24.) Dies können wir auch folgendermaßen berechnen: Da die deterministische Dynamik des chemischen Reaktionsnetzwerkes durch

$$\dot{x} = \lambda - \rho x$$

gegeben ist, ist das einzige Gleichgewicht  $c = \lambda/\rho$ . Dieses ist komplex balanziert, da  $A$  der einzige Komplex ist. Nach Theorem 1.29 ist demnach  $\text{Poi}(\lambda/\rho)$  eine stationäre Verteilung.

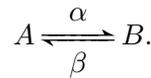
**Korollar 1.31.** Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk, das sich nach dem Massenwirkungsgesetz mit Raten  $\kappa_k, k \in \mathcal{R}$ , entwickelt. Weiter sei  $c \in \mathbb{R}_{>0}^{\mathcal{S}}$  ein komplex-balanzierteres Gleichgewicht für die deterministische Dynamik und  $E$  eine stochiometrische Kompatibilitätsklasse. Dann hat die stochastische Dynamik eine stationäre Verteilung auf  $E$  der Form

$$\pi(x) = \begin{cases} M_E \prod_{s \in \mathcal{S}} \frac{c_s^{x_s}}{x_s!}, & x \in E, \\ 0, & \text{sonst,} \end{cases}$$

wobei  $M_E$  eine Normierungskonstante ist (d.h.  $M_E$  ist so gewählt, dass  $\sum_x \pi(x) = 1$ ).

*Beweis.* Aus Theorem 1.29 wissen wir bereits, dass (1.8) eine stationäre Verteilung ist. Ist jedoch  $X_0 \sim \pi$  mit  $X_0 \in E$ , so muss  $X_t \in E$  für alle  $t$ . Da  $X_0 \sim X_t$  wegen der Stationarität von  $\pi$ , muss also  $\pi$  von der Form  $\pi = \sum_E \pi(E) \pi_E$  sein, wobei die Summe über alle Kompatibilitätsklassen geht,  $\pi(E)$  die Wahrscheinlichkeit ist, dass  $X_0 \in E$  ist, und  $\pi_E$  eine stationäre Verteilung auf  $E$ . Daraus folgt bereits die behauptete Form der stationären Verteilung auf  $E$ .  $\square$

**Beispiel 1.32** (Switching). Wir betrachten das chemische Reaktionsnetzwerk



Für die deterministische Dynamik ist

$$\dot{x}_A = -\dot{x}_B = \beta x_B - \alpha x_A.$$

Daraus lesen wir ab, dass etwa  $c_A = 1 - c_B = \frac{\beta}{\alpha + \beta}$  ein Gleichgewicht ist, das auch komplex balanziert ist. Weiter gibt es – für die stochastische Dynamik – die stochiometrischen Kompatibilitätsklassen  $E_1, E_2, \dots$  mit  $E_N = \{(x_A, x_B) : x_A + x_B = N\}$ . Deshalb ist die Verteilung  $\pi$  auf  $E_N$ , gegeben durch

$$\pi(x_A, x_B) \sim \left(\frac{\beta}{\alpha + \beta}\right)^{x_A} \frac{1}{x_A!} \left(\frac{\alpha}{\alpha + \beta}\right)^{x_B} \frac{1}{x_B!} \sim \binom{N}{x_A} c_A^{x_A} (1 - c_A)^{N - c_A}$$

für  $x \in E_N$  und  $\pi(x_A, x_B) = 0$  sonst. Mit anderen Worten ist  $X_A$  im Gleichgewicht  $B(N, c_A)$ -verteilt.

Wir wollen nun noch – ohne Beweis – ein klassisches Resultat angeben, das für die deterministische Dynamik die Existenz und Eindeutigkeit eines komplex balanzierten Gleichgewichts sichert. Hierfür benötigen wir noch ein paar Definitionen.

**Definition 1.33** ((Schwach) reversibles Netzwerk, Defizient). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk.

1. Ist für jedes  $(\nu \rightarrow \nu') \in \mathcal{R}$  auch  $(\nu' \rightarrow \nu) \in \mathcal{R}$ , so heißt das Netzwerk reversibel.
2. Gibt es für jedes  $(\nu \rightarrow \nu') \in \mathcal{R}$  eine Folge  $\nu = \nu_1, \nu_2, \dots, \nu_n = \nu'$ , so dass  $\nu_k \rightarrow \nu_{k+1} \in \mathcal{R}, k = 1, \dots, n - 1$ , so heißt das Netzwerk schwach reversibel.
3. Erstellt man eine Graphen mit den Komplexen als Knoten und den Reaktionen als Kanten, so heißen die Zusammenhangskomponenten des Graphen auch Linkage-Klassen.
4. Sei  $|\mathcal{C}|$  die Anzahl an Komplexen,  $\ell$  die Anzahl der Linkage-Klassen und  $s$  die Dimension des stoichiometrischen Kompatibilitätsklasse (also die Dimension von  $E := \text{span}\{\nu'_k - \nu_k : k \in \mathcal{R}\} \subseteq \mathbb{R}^{\mathcal{S}}$ ). Dann ist  $\delta := |\mathcal{C}| - \ell - s$  die Defizienz des Netzwerks.

**Proposition 1.34.** Die Defizienz eines chemischen Reaktionsnetzwerks ist nicht-negativ.

*Beweis.* Sei  $\mathcal{C}_i$  die Komplexe der  $i$ -ten Linkage-Klasse und  $E_i$  der stoichiometrische Teilraum, der von der  $i$ -ten Linkage-Klasse erzeugt wird. Der Raum  $E_i$  hat Dimension höchstens  $|\mathcal{C}_i| - 1$ . (Um die Dimension herauszufinden, entfernt man zunächst die Reaktionen, die durch Hintereinanderausführung von anderen Reaktionen erreicht werden können. Anschließend hat der Graph der  $i$ -ten Linkage-Klasse genau  $|\mathcal{C}_i|$  Knoten und  $|\mathcal{C}_i| - 1$  Kanten. Nur wenn die Reaktionsvektoren dieser Kanten linear unabhängig sind, ist die Dimension  $|\mathcal{C}_i| - 1$ , andernfalls kleiner.) Nun schreiben wir

$$s = \dim E \leq \sum_{i=1}^{\ell} \dim E_i \leq \sum_{i=1}^{\ell} (|\mathcal{C}_i| - 1) = |\mathcal{C}| - \ell.$$

□

Ohne Beweis folgt nun ein Hauptresultat der deterministischen Theorie chemischer Reaktionsnetzwerke von Martin Feinberg aus den 1970er Jahren.

**Theorem 1.35** (Defizienz-0-Theorem). Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $X$  die deterministische Dynamik, die dem Massenwirkungsgesetz mit Reaktionsraten  $\kappa_k, k \in \mathcal{R}$  folgt. Ist die Defizienz des Netzwerks 0, dann sind folgende Aussagen äquivalent:

1. Das Netzwerk ist schwach reversibel.
2. Es gibt ein komplex-balanziertes Gleichgewicht  $c \in \mathbb{R}_+^{\mathcal{S}}$  mit  $c_s > 0, x \in \mathcal{S}$ .

Weiter gibt es dann in jeder nicht-negativen stoichiometrischen Kompatibilitätsklasse genau ein Gleichgewicht, und dieses ist komplex-balanziert.

## 1.4 Konvergenz von stochastischen zu deterministischen Dynamiken

Wir wollen nun einen Zusammenhang zwischen stochastischen und deterministischen Dynamiken von chemischen Reaktionsnetzwerken herstellen. Intuitiv ist klar, dass es bei einer großen Molekülnzahl zu weniger relativen Schwankungen in der Molekülnzahl gibt. Formal müssen wir hierfür etwas ausholen und stellen zunächst wieder ein paar allgemeine Resultate zusammen.

## Konvergenz des reskalierten Poisson-Prozesses

Aus den stochastischen Prozessen ist der Satz von Donsker über die Konvergenz einer Irrfahrt gegen die Brown'sche Bewegung bekannt. Wir wollen nun sogar etwas einfacher die Konvergenz von Markov-Sprung-Prozessen gegen die Lösung einer Differentialgleichung zeigen. Zunächst ein Resultat aus den stochastischen Prozessen.

**Proposition 1.36** (Schwache und fdd-Konvergenz). *Seien  $X, X^1, X^2, \dots$  Zufallsvariable mit Werten in  $\mathcal{C}_{\mathbb{R}^d}([0, \infty))$ . Dann sind äquivalent*

1.  $X^n \xrightarrow{n \rightarrow \infty} X$
2.  $X^n \xrightarrow{n \rightarrow \infty} fdd X$  und  $\{X^n : n = 1, 2, \dots\}$  ist straff in  $\mathcal{C}_{\mathbb{R}^d}([0, \infty))$ .

Um diesem Resultat Sinn zu verleihen, benötigt man die Topologie der gleichmäßigen Konvergenz auf Kompakta auf dem Raum  $\mathcal{C}_{\mathbb{R}^d}([0, \infty))$ . Wir erwähnen hier, dass im Fall, dass  $X^1, X^2, \dots$  Markov-Sprung-Prozesse sind, die Voraussetzungen des Satzes nicht erfüllt sind, da diese keine stetigen Pfade besitzen. In diesem Fall muss man zum Raum  $\mathcal{D}_{\mathbb{R}^d}([0, \infty))$  der rechtsstetigen Pfade mit linksseitigen Grenzwerten übergehen. Für diesen Raum gilt Proposition 1.36 völlig analog, mit der Ausnahme, dass die Topologie auf  $\mathcal{D}_{\mathbb{R}^d}([0, \infty))$  eine andere ist. Da wir hier aber nicht in diese Feinheiten einsteigen wollen, begnügen wir uns im Folgenden mit der fdd-Konvergenz. Wir formulieren zunächst ein einfaches Resultat zu Poisson-Prozessen.

**Proposition 1.37** (Konvergenz von reskalierten Poisson-Prozessen). *Sei  $Y$  ein Poisson-Prozess mit Rate 1 und  $N > 0$ . Dann ist  $(Y_{Nt})_{t \geq 0}$  ein Poisson-Prozess mit Rate  $N$ . Weiter gilt*

$$\left(\frac{1}{N}Y_{Nt}\right)_{t \geq 0} \xrightarrow{N \rightarrow \infty} fdd id$$

und

$$\sqrt{N} \left(\frac{1}{N}Y_{Nt} - t\right)_{t \geq 0} \xrightarrow{N \rightarrow \infty} fdd W,$$

wobei  $W = (W_t)_{t \geq 0}$  eine Brown'sche Bewegung (mit  $W_0 = 0$ ) ist.

*Beweis.* Es gilt

$$\begin{aligned} \mathbf{E} \left[ \frac{1}{N} Y_{Nt} \right] &= t, \\ \mathbf{Var} \left[ \frac{1}{N} Y_{Nt} \right] &= \frac{Nt}{N^2} \xrightarrow{N \rightarrow \infty} 0. \end{aligned}$$

Deshalb folgt die ( $L^2$ -Konvergenz, also auch die) stochastische Konvergenz  $\frac{1}{N}Y_{Nt} \xrightarrow{t \rightarrow \infty}_p t$ . Die Konvergenz der endlich-dimensionalen Randverteilungen ist dann eine Übungsaufgabe. Für die Konvergenz gegen die Brown'sche Bewegung stellen wir zunächst fest, dass  $Y_t^N \sim P_1 + \dots + P_N$ , wobei  $P_1, \dots, P_N \sim \text{Poi}(t)$ . Nach dem zentralen Grenzwertsatz gilt demnach

$$\sqrt{\frac{N}{t}} \left( \frac{1}{N} Y_{Nt} - t \right) = \frac{P_1 + \dots + P_N - N\mathbf{E}[P_1]}{\sqrt{Nt}} \xrightarrow{N \rightarrow \infty} W_1,$$

also  $\sqrt{N} \left( \frac{1}{N} Y_{Nt} - t \right) \xrightarrow{N \rightarrow \infty} W_t$ . Für  $t_1 < \dots < t_n$  ist mir derselben Argumentation

$$\begin{aligned} & \left( \sqrt{N} \left( \left( \frac{1}{N} Y_{Nt_2} - \frac{1}{N} Y_{Nt_1} - (t_2 - t_1) \right), \dots, \left( \frac{1}{N} Y_{Nt_n} - \frac{1}{N} Y_{Nt_{n-1}} - (t_n - t_{n-1}) \right) \right) \right. \\ & \quad \left. \xrightarrow{N \rightarrow \infty} (W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}) \right). \end{aligned}$$

Daraus folgt die Behauptung.  $\square$

**Bemerkung 1.38** (Konvergenz der Generatoren). 1. Der Prozess  $X^N = (\frac{1}{N} Y_{Nt})_{t \geq 0}$  ist Markov mit Generator

$$G^N f(x) = N \left( f \left( x + \frac{1}{N} \right) - f(x) \right) = \frac{f \left( x + \frac{1}{N} \right) - f(x)}{1/N} \xrightarrow{N \rightarrow \infty} f'(x)$$

für  $f \in \mathcal{C}^1(\mathbb{R})$ . Weiter stellen wir fest, dass  $Gf = f'$  der Generator eines deterministischen Prozesses  $X$  ist, der  $\dot{X} = 1$  löst, also  $X(t) = t$ . Mit anderen Worten: Hier konvergiert nach Proposition 1.37  $X^N \xrightarrow{N \rightarrow \infty}_{fdd} X$  und gleichzeitig ist  $G^N f(x) \xrightarrow{N \rightarrow \infty} Gf(x)$  für alle  $f \in \mathcal{D}(G)$ .

2. Der Prozess  $X^N = (\sqrt{N} (\frac{1}{N} Y_{Nt} - t))_{t \geq 0}$  ist Markov mit Generator

$$\begin{aligned} G^N f(x) &= N \left( f \left( x + \frac{1}{\sqrt{N}} \right) - f(x) \right) - \sqrt{N} f'(x) \\ &= N \left( \frac{1}{\sqrt{N}} f'(x) + \frac{1}{N} \frac{1}{2} f''(x) + o(1/N) \right) - \sqrt{N} f'(x) \xrightarrow{N \rightarrow \infty} \frac{1}{2} f''(x) \end{aligned}$$

für  $f \in \mathcal{C}^2(\mathbb{R})$ . Bekanntermaßen ist  $f \mapsto Gf = \frac{1}{2} f''$  der Generator einer Brown'schen Bewegung  $X$ . Wieder ist es also so, dass  $X^N \xrightarrow{N \rightarrow \infty}_{fdd} X$  und gleichzeitig ist  $G^N f(x) \xrightarrow{N \rightarrow \infty} Gf(x)$  für alle  $f \in \mathcal{D}(G)$ .

## Einiges über Martingal-Probleme

Im Folgenden sei  $X$  immer ein  $\mathbb{R}^d$ -wertiger, stochastischer Prozess und  $\mathcal{F}^X = (\mathcal{F}_t^X)_{t \geq 0}$  die von  $X$  erzeugte Filtration. Für  $f, g \in \mathcal{C}_b(\mathbb{R}^d)$  ist  $M^{f,g} = (M_t^{f,g})_{t \geq 0}$  mit

$$M_t^{f,g} = f(X_t) - f(X_0) - \int_0^t g(X_s) ds. \quad (1.9)$$

(Dies wird weiter unten oft ein Martingal sein.) Wir schreiben auch  $Gf$  anstatt  $g$ , womit wir andeuten wollen, dass es eine Abbildung  $G : f \mapsto g$  gibt, so dass  $M^{f,Gf}$  ein Martingal ist. In Theorem 1.9 haben wir wiederholt, dass ein Markov-Prozess sein eigenes Martingal-Problem löst, in folgendem Sinne.

**Definition 1.39** (Martingal-Problem). Sei  $\mathbf{P}_0$  ein Wahrscheinlichkeitsmaß auf  $\mathbb{R}^d$  und<sup>4</sup>  $G : \mathcal{D} \subseteq \mathcal{B}(\mathbb{R}^d) \rightarrow \mathbb{M}(\mathbb{R}^d)$  ein Operator. Sind für einen  $\mathbb{R}^d$ -wertigen stochastischen Prozess  $X = (X_t)_{t \geq 0}$  mit Pfaden in  $D_{\mathbb{R}^d}([0, \infty))$  alle abgeleiteten Prozesse  $M^{f,Gf}$  Martingale, und ist  $X_0 \stackrel{d}{=} \mathbf{P}_0$ , so heißt  $X$  Lösung des lokalen  $(\mathbf{P}_0, G, \mathcal{D})$  Martingalproblems (auf  $\mathcal{D}_{\mathbb{R}^d}([0, \infty))$ ). Ist keine Verteilung für  $X_0$  spezifiziert, spricht man auch vom  $(G, \mathcal{D})$ -Martingalproblem.

Gibt es genau eine Lösung  $X$  für das  $(\mathbf{P}_0, G, \mathcal{D})$  Martingalproblem, so sagt man, dass es gut gestellt sei.

<sup>4</sup>Der Raum  $\mathcal{M}(E)$  ist beinhaltet genau die messbaren Funktionen  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , der Raum  $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{M}(\mathbb{R}^d)$  ist der Raum der messbaren, beschränkten Funktionen.

Das folgende Resultat ist eigentlich schon mit Definition 1.8 klar.

**Lemma 1.40** (Martingaleigenschaft und zeitliche Ableitung). *Sei  $X = (X_t)_{t \geq 0}$  ein  $\mathbb{R}^d$ -wertiger stochastischer Prozess und  $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$  so, dass  $M^{f,g}$  wie in (1.9) ein Martingal ist. Dann ist für  $t \geq 0$*

$$\lim_{h \rightarrow 0} \frac{\mathbf{E}[f(X_{t+h}) - f(X_t) | \mathcal{F}_t]}{h} = g(X_t).$$

Inbesondere ist

$$\frac{d}{dt} \mathbf{E}[f(X_t)] = \mathbf{E}[g(X_t)].$$

Ist  $X$  ein Markov-Prozess mit Generator  $G$ , dann ist also, für  $f \in \mathcal{D}(G)$

$$\lim_{h \rightarrow 0} \frac{\mathbf{E}[f(X_{t+h}) - f(X_t) | \mathcal{F}_t]}{h} = Gf(X_t).$$

*Beweis.* Wir verwenden direkt die Martingaleigenschaft und schreiben direkt

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\mathbf{E}[f(X_{t+h}) - f(X_t) | \mathcal{F}_t]}{h} &= \lim_{h \rightarrow 0} \frac{\mathbf{E}[M_{t+h}^f - M_t^f + \int_t^{t+h} g(X_s) ds | \mathcal{F}_t]}{h} \\ &= \lim_{h \rightarrow 0} \mathbf{E} \left[ \frac{1}{h} \int_t^{t+h} g(X_s) ds \middle| \mathcal{F}_t \right] = g(X_t). \end{aligned}$$

□

Der Generator eines Markov-Prozesses beschreibt seine infinitesimalen Eigenschaften. Das folgende Resultat zeigt, wie man sich dies zunutze machen kann, um etwa Momente des Prozesses zu berechnen.

**Proposition 1.41** (Quadratische Variation). *Sei  $X = (X_t)_{t \geq 0}$  ein  $\mathbb{R}^d$ -wertiger stochastischer Prozess,  $f, f^2, Gf, Gf^2 : \mathbb{R}^d \rightarrow \mathbb{R}$  so, dass  $M_t^{f, Gf}$  und  $M_t^{f^2, Gf^2}$  Martingale sind. Dann ist*

$$\mathbf{E}[(M_t^f)^2] = \int_0^t \mathbf{E}[Gf^2(X_s) - 2f(X_s)Gf(X_s)] ds.$$

und

$$\mathbf{Var}[f(X_t)] = \mathbf{E}[(M_t^f)^2] + \int_0^t 2\mathbf{Cov}[f(X_s), Gf(X_s)] ds$$

*Beweis.* OBdA sei  $f(X_0) = 0$ . Wir schreiben

$$\begin{aligned} \mathbf{E}[(M_t^f)^2] &= \mathbf{E} \left[ f(X_t)^2 - 2 \int_0^t \mathbf{E}[f(X_t) | \mathcal{F}_s] Gf(X_s) ds + \left( \int_0^t Gf(X_s) ds \right)^2 \right] \\ &= \mathbf{E} \left[ f(X_t)^2 - 2 \int_0^t \left( f(X_s) + \int_s^t Gf(X_r) dr \right) Gf(X_s) ds + \left( \int_0^t Gf(X_s) ds \right)^2 \right] \\ &= \mathbf{E} \left[ f(X_t)^2 - 2 \int_0^t f(X_s) Gf(X_s) ds \right. \\ &\quad \left. - \int_0^t \int_0^t Gf(X_r) Gf(X_s) dr ds + \left( \int_0^t Gf(X_s) ds \right)^2 \right] \\ &= \mathbf{E} \left[ f(X_t)^2 - 2 \int_0^t f(X_s) Gf(X_s) ds \right] \end{aligned}$$

und

$$\begin{aligned}
\mathbf{Var}[f(X_t)] &= \mathbf{E}[f^2(X_t)] - \mathbf{E}[f(X_t)]^2 = \int_0^t \frac{d}{ds} \mathbf{E}[f(X_s)^2] - \frac{d}{ds} \mathbf{E}[f(X_s)]^2 ds \\
&= \int_0^t \mathbf{E}[Gf^2(X_s)] - 2\mathbf{E}[f(X_s)]\mathbf{E}[Gf(X_t)] ds \\
&= \int_0^t \mathbf{E}[Gf^2(X_s)] - 2\mathbf{E}[f(X_s)Gf(X_t)] + 2\mathbf{Cov}[f(X_s), Gf(X_s)] ds \\
&= \mathbf{E}[(M_t^f)^2] + \int_0^t 2\mathbf{Cov}[f(X_s), Gf(X_s)] ds
\end{aligned}$$

□

Da wir mit Hilfe eines Generators nun Momente berechnen können, können wir auch den Spezialfall eines Martingalproblems eines deterministischen Prozesses behandeln.

**Proposition 1.42** (Martingal-Problem für Lösungen von DGLs). *Sei  $X = (X_t)_{t \geq 0}$  ein Markov-Prozess mit Zustandsraum  $\mathbb{R}^d$ , so dass für alle  $f \in \mathcal{C}_b^1(\mathbb{R}^d)$  der Prozess  $M^{f, Gf}$  mit*

$$Gf(x) = \sum_{i=1}^d \mu_i(x) \frac{\partial f(x)}{\partial x_i}$$

ein Martingal ist. Dann ist  $X$  deterministisch und löst

$$\frac{d}{dt} X = \mu(X).$$

*Beweis.* Wir zeigen die Aussage nur für  $d = 1$  und  $X_0 = 0$ . Der allgemeine Fall folgt dann mit Hilfe von Verschiebungen und Projektionen. Nach Lemma 1.41 und

$$Gf^2(x) - 2f(x)Gf(x) = \mu(x) \left( \frac{df^2(x)}{dx} - 2f(x) \frac{df(x)}{dx} \right) = 0$$

ist

$$f(X_t) = f(X_0) + \int_0^t \mu(X_s) f'(X_s) ds$$

fast sicher oder

$$\frac{d}{dt} f(X_t) = \mu(X_t) f'(X_t).$$

Damit ist  $(f(X_t))_{t \geq 0}$  als Lösung einer Differentialgleichung deterministisch. Da dies für alle  $f \in \mathcal{C}_b(\mathbb{R}^d)$  gilt, ist auch  $X$  deterministisch. Setzt man  $f(x) = x$  folgt die zweite Behauptung. □

Zuletzt benötigen wir noch eine Charakterisierung des Martingals  $M^{f, g}$ , die ohne bedingte Erwartungen auskommt.

**Proposition 1.43** (Charakterisierung der Martingaleigenschaft). *Der Prozess  $M^{f, g}$  ist genau dann ein  $\mathcal{F}^X$ -Martingal, wenn für  $h_1, \dots, h_n \in \mathcal{C}_b(\mathbb{R}^d)$  und  $t_1 \leq \dots \leq t_n \leq t$  und  $h > 0$*

$$\mathbf{E} \left[ \left( f(X_{t+h}) - f(X_t) - \int_t^{t+h} Gf(X_s) ds \right) h_1(X_{t_1}) \cdots h_n(X_{t_n}) \right] = 0. \quad (1.10)$$

*Beweis.* Ist zunächst  $M^{f,g}$  ein Martingal, so ist

$$\begin{aligned} & \mathbf{E} \left[ \left( f(X_{t+h}) - f(X_t) - \int_t^{t+h} g(X_s) ds \right) h_1(X_{t_1}) \cdots h_n(X_{t_n}) \right] \\ &= \mathbf{E} \left[ \mathbf{E} \left[ f(X_{t+h}) - f(X_t) - \int_t^{t+h} g(X_s) ds \middle| \mathcal{F}_t \right] h_1(X_{t_1}) \cdots h_n(X_{t_n}) \right] = \\ &= \mathbf{E} \left[ \mathbf{E} [M_{t+h}^{f,g} - M_t^{f,g} | \mathcal{F}_t] h_1(X_{t_1}) \cdots h_n(X_{t_n}) \right] = 0. \end{aligned}$$

Gilt andersherum (1.11) für alle  $t_1 \leq \cdots \leq t_n \leq t$  und  $h > 0$ , so folgt mit einem Approximationsargument (Indikatorfunktionen von Quadern lassen sich durch stetige Funktionen approximieren), dass

$$\mathbf{E} \left[ \left( f(X_{t+h}) - f(X_t) - \int_t^{t+h} Gf(X_s) ds \right) 1_{A_1}(X_{t_1}) \cdots 1_{A_n}(X_{t_n}) \right] = 0. \quad (1.11)$$

für beliebige Quader  $A_1, \dots, A_n$ . Nun ist aber

$$\mathcal{E} \{ \{X_{t_1} \in A_1\} \cap \cdots \cap \{X_{t_n} \in A_n\} : t_1 \leq \cdots \leq t_n \leq t, A_1, \dots, A_n \subseteq \mathbb{R}^d \text{ offener Quader} \}$$

ein schnittstabiler Erzeuger von  $\sigma((X_s)_{0 \leq s \leq t})$ . Deshalb gilt für alle  $A \in \mathcal{F}_t^X$

$$\mathbf{E}[(M_{t+h}^{f,g} - M_t^{f,g})^+, A] = \mathbf{E}[(M_{t+h}^{f,g} - M_t^{f,g})^-, A], \text{ d.h. } \mathbf{E}[M_{t+h}^{f,g} - M_t^{f,g}, A] = 0,$$

d.h.  $M^{f,g}$  ist ein Martingal. □

**Theorem 1.44** (Konvergenz von Lösungen von Martingalproblemen). *Seien  $X^1, X^2, \dots$  Markov-Prozesse mit Generatoren  $G^1, G^2, \dots$ . Gilt für ein  $G : \mathcal{D} \subseteq \mathcal{D}(G^1) \cap \mathcal{D}(G^2) \cap \cdots \rightarrow \mathcal{C}_b(\mathbb{R}^d)$ , dass für  $f \in \mathcal{D}(G)$*

$$\lim_{N \rightarrow \infty} \|G^N f - Gf\| \xrightarrow{N \rightarrow \infty} 0 \quad (1.12)$$

uniform auf Kompakta und gilt  $X^N \xrightarrow{N \rightarrow \infty}_{fdd} X$  für einen stochastischen Prozess  $X$ , dann löst  $X$  das  $(G, \mathcal{D})$ -Martingal-Problem.

*Beweis.* Nach Proposition 1.43 müssen wir zeigen, dass (1.11) gilt. Sei also  $f \in \mathcal{D}$ ,  $t_1 \leq \cdots \leq t_n \leq t, h > 0$  und  $h_1, \dots, h_n \in \mathcal{C}_b(\mathbb{R}^d)$ . Dann gilt

$$\begin{aligned} & \mathbf{E} \left[ \left( f(X_{t+h}) - f(X_t) - \int_t^{t+h} Gf(X_s) ds \right) h_1(X_{t_1}) \cdots h_n(X_{t_n}) \right] \\ &= \lim_{N \rightarrow \infty} \mathbf{E} \left[ \left( f(X_{t+h}^N) - f(X_t^N) - \int_t^{t+h} Gf(X_s^N) ds \right) h_1(X_{t_1}^N) \cdots h_n(X_{t_n}^N) \right] \\ &= \lim_{N \rightarrow \infty} \mathbf{E} \left[ \left( f(X_{t+h}^N) - f(X_t^N) - \int_t^{t+h} G^N f(X_s^N) ds \right) h_1(X_{t_1}^N) \cdots h_n(X_{t_n}^N) \right] = 0. \end{aligned}$$

Dabei folgt das erste Gleichheitszeichen mit majorisierter Konvergenz und das zweite wegen (1.12). Damit ist also  $M^{f,Gf}$  für jedes  $f \in \mathcal{D}(G)$  ein Martingal und die Behauptung ist gezeigt. □

Es folgt nun der gewünschte Zusammenhang zwischen stochastischer und deterministischer Modellierung.

**Theorem 1.45** (Gesetz der Großen Zahlen für stochastische Dynamiken). *Sei  $(\mathcal{S}, \mathcal{C}, \mathcal{R})$  ein chemisches Reaktionsnetzwerk, sowie  $\lambda_k^N : \mathcal{Z}_+^{\mathcal{S}} \rightarrow \mathbb{R}_+$ ,  $k \in \mathcal{R}$ ,  $N = 1, 2, \dots$ . Sei  $X^N$  die Dynamik des stochastischen Reaktionsnetzwerkes mit Raten  $\lambda_k^N$ ,  $k \in \mathcal{R}$  und  $V^N := \frac{1}{N}X^N$ . Gelte weiterhin für  $v \in \frac{1}{N}\mathcal{Z}_+^{\mathcal{S}}$*

$$\frac{1}{N}\lambda_k^N(Nv) \xrightarrow{N \rightarrow \infty} \lambda_k(v)$$

*uniform auf Kompakta und sei  $V$  die deterministische Dynamik des chemischen Reaktionsnetzwerkes mit Raten  $\lambda_k$ ,  $k \in \mathcal{R}$ . Falls  $V^N$  konvergiert und falls  $V_0^N \xrightarrow{N \rightarrow \infty} V_0$ , dann gilt*

$$V^N \xrightarrow{N \rightarrow \infty} fdd V.$$

**Bemerkung 1.46.** Als Vorüberlegung gehen wir mal von einer Konvergenz von  $V^N$  gegen einen Grenzprozess  $V$  aus. Es gilt

$$X_t^N = X_0^N + \sum_{k \in \mathcal{K}} \zeta_k Y_k \left( \int_0^t \lambda_k^N(X_s^N) ds \right)$$

für unabhängig Poisson-Prozesse  $Y_k$ ,  $k \in \mathcal{R}$  mit Rate 1, und damit

$$\begin{aligned} V_t^N &= V_0^N + \sum_{k \in \mathcal{K}} \zeta_k \frac{1}{N} Y_k \left( \int_0^t \lambda_k^N(NV_s^N) ds \right) \\ &\approx V_0^N + \sum_{k \in \mathcal{K}} \zeta_k \frac{1}{N} Y_k \left( N \int_0^t \lambda_k(V_s) ds \right) \\ &\approx V_0 + \sum_{k \in \mathcal{K}} \zeta_k \int_0^t \lambda_k(V_s) ds, \end{aligned}$$

wegen Proposition 1.37. Aus der letzten Zeile lesen wir ab, dass

$$\dot{V} = \sum_{k \in \mathcal{K}} \zeta_k \lambda_k(V).$$

*Beweis.* Wir verwenden Theorem 1.44. Der Generator von  $V^N$  ist für  $f \in \mathcal{C}^1(\mathbb{R}^d)$

$$\begin{aligned} G^{V^N} f(v) &= \sum_{k \in \mathcal{R}} \lambda_k^N(Nv) \left( f\left(v + \frac{1}{N}\zeta_k\right) - f(v) \right) \\ &= \sum_{k \in \mathcal{R}} \lambda_k^N(Nv) \left( \frac{1}{N}\zeta_k f'(v) + o(1/N) \right) \\ &= \sum_{k \in \mathcal{R}} (\lambda_k(v) + o(1)) (\zeta_k f'(v) + o(1)) \\ &= G^V f(v) + o(1). \end{aligned}$$

□

**Korollar 1.47** (Stochastisches und deterministisches Massenwirkungsgesetz). Sei  $(S, C, \mathcal{R})$  ein chemisches Reaktionsnetzwerk und  $\kappa_k, k \in \mathcal{R}$  Reaktionsraten. Sei weiter  $X^N$  das stochastische System, das sich nach dem Massenwirkungsgesetz mit Raten  $\kappa_k N^{1-|\nu_k|}, k \in \mathcal{R}$  entwickelt. Weiter sei  $X$  das deterministische System, das sich nach dem Massenwirkungsgesetz mit Raten  $\kappa_k, k \in \mathcal{R}$  entwickelt. Konvergiert  $X^N/N$ , dann gilt

$$\frac{1}{N} X^N \xrightarrow{N \rightarrow \infty} fdd X.$$

*Beweis.* Es gilt

$$\lambda_k^N(x) = \kappa_k N^{1-|\nu_k|} \prod_{s \in S} x_s \cdots (x_s - \nu_{ks} + 1) = N \kappa_k \prod_{s \in S} \frac{x_s}{N} \cdots \frac{x_s - \nu_{ks} + 1}{N}.$$

Damit gilt

$$\frac{1}{N} \lambda_k^N(Nv) = \kappa_k \prod_{s \in S} v_s \cdots \left( v_s - \frac{\nu_{ks} + 1}{N} \right) \xrightarrow{N \rightarrow \infty} \kappa_k \prod_{s \in S} v_s^{\nu_{ks}} = \lambda_k(v).$$

Damit folgt aus Theorem 1.45 die Behauptung.  $\square$

**Beispiel 1.48.** In Beispiel 1.13 sei  $\lambda = N\kappa$ . Dann sind die Voraussetzungen des Korollars gegeben und mit  $V_t^N := X_t^N/N$  ist

$$\begin{aligned} V_t^N &= V_0^N + \frac{1}{N} Y_1(N\kappa t) - \frac{1}{N} Y_2 \left( N \int_0^t \rho V_s^N ds \right) \\ &\approx V_0 + \kappa t - \int_0^t \rho V_s ds, \end{aligned}$$

d.h. in der Tat erfüllt der Grenzwert  $V$  die Differentialgleichung

$$\dot{V} = \kappa - \rho V.$$

**Theorem 1.49** (Zentraler Grenzwertsatz für stochastische Dynamiken). Sei  $(S, C, \mathcal{R})$  ein chemisches Reaktionsnetzwerk, sowie  $\lambda_k^N, k \in \mathcal{R}, X^N, V^N, N \in \mathbb{N}$  und  $\lambda_k, k \in \mathcal{R}$  so, dass die Voraussetzungen aus Theorem 1.45 erfüllt sind, d.h.

$$\lambda_k^N(Nv) = N\lambda_k(v) + O(1) \text{ und glatt}$$

$$V^N \xrightarrow{N \rightarrow \infty} V. \text{ Dann gilt}$$

$$\sqrt{N}(V^N - V) \xrightarrow{N \rightarrow \infty} U,$$

wobei  $U$  das System stochastischer Differentialgleichungen

$$dU = \sum_{k \in \mathcal{K}} \zeta_k (\nabla \lambda_k(V_s) \cdot U ds + \sqrt{\lambda_k(V)} dW_k)$$

löst.

**Bemerkung 1.50.** Als Vorüberlegung bemerken wir folgendes. Nach Proposition 1.37 lässt sich in gewisser Hinsicht

$$\frac{1}{N} Y_{Nt} \approx t + \frac{1}{\sqrt{N}} W_t$$

für eine Brown'sche Bewegung  $W$  schreiben. Weiter ist  $(W_{ct})_{t \geq 0} \stackrel{d}{=} (\sqrt{c}W_t)_{t \geq 0}$  für  $c > 0$ . Deshalb ist – wie in Bemerkung 1.46 –

$$\begin{aligned} V_t^N &= V_0^N + \sum_{k \in \mathcal{K}} \zeta_k \frac{1}{N} Y_k \left( \int_0^t \lambda_k^N(NV_s^N) ds \right) \\ &\approx V_0^N + \sum_{k \in \mathcal{K}} \zeta_k \frac{1}{N} Y_k \left( N \int_0^t \lambda_k(V_s^N) ds \right) \\ &\approx V_0 + \sum_{k \in \mathcal{K}} \zeta_k \int_0^t \lambda_k(V_s^N) ds + \frac{1}{\sqrt{N}} W_k \left( \int_0^t \lambda_k(V_s) ds \right) \\ &\approx V_0 + \sum_{k \in \mathcal{K}} \zeta_k \int_0^t \lambda_k(V_s) + \nabla \lambda_k(V_s) \cdot (V_s^N - V_s) ds + \frac{1}{\sqrt{N}} \int_0^t \sqrt{\lambda_k(V_s)} dW_k(s) \end{aligned}$$

wegen Proposition 1.37, wobei  $W_k, k \in \mathcal{R}$  unabhängige Brown'sche Bewegungen sind. Aus der letzten Zeile lesen wir ab, dass

$$\sqrt{N}(V^N - V) = \left( \sum_{k \in \mathcal{K}} \zeta_k \int_0^t \nabla \lambda_k(V_s) \sqrt{N}(V_s^N - V_s) ds + \int_0^t \lambda_k(V) dW_k \right)_{t \geq 0}.$$

**Bemerkung 1.51.** Stochastische Differentialgleichungen lesen sich etwas leichter als Martingalprobleme, da man bei der Schreibweise

$$dX = \mu(X)dt + \sigma(X) \cdot dW \quad (1.13)$$

genau ablesen kann, was die infinitesimale Drift und infinitesimale Varianz ist. Dabei verhält es sich so (siehe die Vorlesung Stochastische Integration), dass  $X = (X_s)_{s \in \mathcal{S}}$  genau dann die Gleichung (1.13) mit  $\mu : \mathbb{R}^{\mathcal{S}} \rightarrow \mathbb{R}^{\mathcal{S}}$  und  $\sigma : \mathbb{R}^{\mathcal{S}} \rightarrow \mathbb{R}^{\mathcal{S}} \times \mathbb{R}^{\mathcal{K}}$  und unabhängigen Brown'schen Bewegungen  $(W_k)_{k \in \mathcal{R}}$  löst, wenn  $X$  das Martingalproblem für

$$(G^X f)(x) = \sum_{s \in \mathcal{S}} \mu^s(x) \frac{\partial f}{\partial x_s}(x) + \frac{1}{2} \sum_{s, t \in \mathcal{S}} \sum_{k \in \mathcal{R}} \sigma^{sk}(x) \sigma^{tk}(x) \frac{\partial^2 f}{\partial x_s \partial x_t}(x) \quad (1.14)$$

löst.

*Beweis von Theorem 1.49.* Sei  $G^N$  der Generator von  $U^N := \sqrt{N}(V^N - V)$ . Dann gilt mit  $v^N = v + u/\sqrt{N}$  und der Hessematrix  $H_f$  der glatten Funktion  $f$

$$\begin{aligned} \lim_{N \rightarrow \infty} G^N f(u) &= \lim_{N \rightarrow \infty} \sum_{k \in \mathcal{R}} \lambda_k^N(Nv^N) (f(u + \zeta_k/\sqrt{N}) - f(u)) - \sqrt{N} \lambda_k(v) \zeta_k \cdot \nabla f(u) \\ &= \lim_{N \rightarrow \infty} \sum_{k \in \mathcal{R}} N \lambda_k(v^N) \cdot (\zeta_k/\sqrt{N} \cdot \nabla f(u) + \frac{1}{2N} \zeta_k H_f(u) \zeta_k^\top) - \sqrt{N} \lambda_k(v) \zeta_k \cdot \nabla f(u) \\ &= \lim_{N \rightarrow \infty} \sum_{k \in \mathcal{R}} \sqrt{N} (\lambda_k(v) + \nabla \lambda_k(v) \cdot u) (\zeta_k \cdot \nabla f(u) + \frac{1}{2\sqrt{N}} \zeta_k H_f(u) \zeta_k^\top) - \sqrt{N} \lambda_k(v) \zeta_k \cdot \nabla f(u) \\ &= \sum_{k \in \mathcal{R}} \zeta_k (\nabla \lambda_k(v) \cdot u) \cdot \nabla f(u) + \frac{1}{2} \lambda_k(v) (\zeta_k H_f(u) \zeta_k^\top), \end{aligned}$$

woraus die geforderte Konvergenz folgt.  $\square$

**Beispiel 1.52.** In Beispiel 1.13 mit  $\lambda = N\kappa$  und  $V_t^N := X_t^N/N$  ist  $V^N \xrightarrow{V} \dot{V}$  mit  $\dot{V} = \kappa - \rho V$ . Weiter ist  $\sqrt{N}(V^N - V) \xrightarrow{N \rightarrow \infty} U$  mit

$$dU = -\rho U ds + \sqrt{\kappa} dW_1 + \sqrt{\rho V} dW_2 = -\rho U ds + \sqrt{\kappa + \rho V} dW.$$

für  $W_1, W_2$  unabhängig und  $W$  geeignet. Ist weiter  $V_0 = \kappa/\rho$ , so ist  $V = \kappa/\rho$  und

$$dU = -\rho U + \sqrt{2\kappa} dW,$$

also ist  $U$  ein Ornstein-Uhlenbeck-Prozess.

## 2 Populationsgenetik

Als Teil der Evolutionstheorie beschäftigt sich die Populationsgenetik mit der Ausbreitung von (genetischen) Typen in Populationen. Wir führen zunächst das Moran-Modell als Standard-Modell der Populationsgenetik ein.

**Definition 2.1** (Moran-Modell).

*Eine Population bestehe aus  $N$  Individuen (Teilchen). Jeder dieser Individuen habe einen (Geno-)Typ  $u \in I = [0, 1]$ . Der Typ von Individuum  $i$  zur Zeit  $t$  bezeichnen wir mit  $X_i(t)$ . Zwischen je zwei Individuen findet mit Rate 1 ein Resampling-Ereignis statt. Bei diesem wird einer der beiden Individuen ausgesucht, produziert einen Nachkommen, der das andere Individuum (und insbesondere seinen Typen) ersetzt.*

*Mit anderen Worten ist  $X = (X_i)_{i=1, \dots, N}$  ein Markov-Sprungprozess und  $(x_1, \dots, x_N)$  spring nach  $(x_i, \dots, x_{j-1}, x_i, x_{j+1}, \dots, x_N)$  für jedes  $i, j$  mit Rate  $\frac{1}{2}$ .*

**Bemerkung 2.2** (Interpretationen). 1. Unter einer Generation versteht man typischerweise die Zeitspanne, die vergeht, bis etwa  $N$  neue Individuen geboren wurden. Da im Moran-Modell mit  $N$  Individuen die Gesamt-Geburtenrate  $\binom{N}{2}$  ist, haben nach etwa einer Zeit  $2/N$  insgesamt  $N$  Geburtseignisse stattgefunden. Wir sagen deshalb, dass im Moran-Modell eine Generation etwa  $2/N$  lange dauert.

2. Entscheidend bei allen Populationsmodellen ist es, dass Individuen Vorfahren haben. Das Moran-Modell ist besonders einfach, weil jedes Individuum zur Zeit  $t$  genau einen Vorfahr zur Zeit  $s \leq t$  besitzt.
3. Man bezeichnet das Moran-Modell auch als Modell mit überlappenden Generationen. Ein Modell mit nicht-überlappenden Generationen ist das Wright-Fisher-Modell, das in den Übungen besprochen wird.
4. Die grafische Repräsentation des Moran-Modells ist etwa in Figur 2.1 dargestellt.
5. Es ist möglich, weitere biologische Mechanismen in das Moran-Modell zu integrieren, insbesondere Mutation, Selektion und Rekombination. Dazu später mehr.
6. Das Moran-Modell gehört zur Klasse der interagierenden Teilchensysteme. Diese sind typischerweise durch einen Typenraum  $I^G$  gekennzeichnet, wobei  $G$  ein Graph ist und  $I$  eine endliche Menge. Im Moran-Modell ist  $G$  der vollständige Graph zwischen  $N$  Knoten und  $I$  der Typenraum.

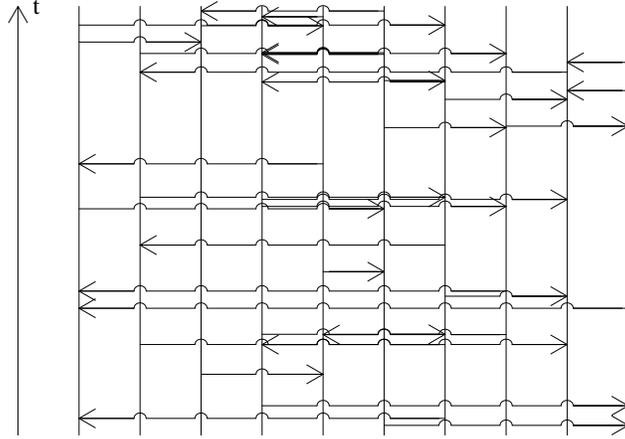


Abbildung 2.1: Grafische Repräsentation eines Moran-Modells. Pfeile bedeuten Resampling-Ereignisse zwischen Individuen. Ein Nachkomme des Individuums am Schaft des Pfeiles ersetzt das Individuum an der Spitze des Pfeiles.

## 2.1 Das neutrale Moran-Modell mit zwei Typen

In diesem Abschnitt betrachten wir ein Moran-Modell der Größe  $N$  mit zwei Allelen  $A$  und  $a$ . Jedes Individuum zur Zeit 0 hat also entweder Typ 0 oder Typ 1. Jeder Nachkomme hat außerdem denselben Typ wie der Vorfahr, d.h. Typen werden vererbt. Wir bezeichnen mit  $X^N$  den Prozess (mit Zustandsraum  $\{0, \dots, N\}/N$ ) der Häufigkeit von Allel  $A$ .

**Proposition 2.3** (Allelfrequenzen im Moran-Modell). *Der Prozess  $X^N$  ist ein Geburts-Todes-Prozess, der von  $x = k/N$  nach*

$$\begin{aligned} x + \frac{1}{N} & \text{ mit Rate } \frac{1}{2}x(1-x)N^2, \\ x - \frac{1}{N} & \text{ mit Rate } \frac{1}{2}x(1-x)N^2. \end{aligned}$$

springt. Weiter konvergiert  $X^N \xrightarrow{N \rightarrow \infty} fdd X$ , wobei  $X$  die stochastische Differentialgleichung

$$dX = \sqrt{X(1-X)}dW \quad (2.1)$$

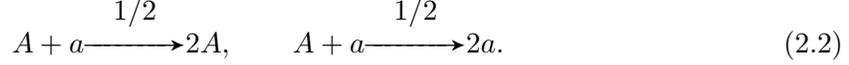
löst.

*Beweis.* Der Prozess  $X^N$  ist ein Markov-Prozess mit Generator (für  $f \in \mathcal{C}_b^2([0, 1])$ )

$$\begin{aligned} G^N f(x) &= \frac{1}{2}x(1-x)N^2(f(x + \frac{1}{N}) + f(x - \frac{1}{N}) - 2f(x)) \\ &= \frac{1}{2}x(1-x)N^2 f''(x) + o(1). \end{aligned}$$

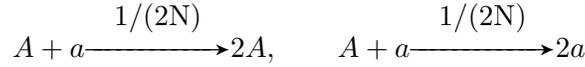
Die Aussage folgt deshalb aus Theorem 1.44. □

**Bemerkung 2.4** (Zeitskalen). Ein Moran-Modell lässt sich auch als chemisches Reaktionsnetzwerk schreiben, nämlich



Zu beachten ist hier die Reaktionsrate  $1/2$ , die dazu führt, dass – falls es  $xN$  Individuen  $A$  und  $(1-x)N$  Individuen  $a$  gibt – die Gesamt-Sprungrate in  $x$  als  $x(1-x)N^2$  berechnet wird. Zu beachten ist dies deshalb, da diese Skalierung anders ist als in Korollar 1.47.

Um dieses Korollar anzuwenden, wäre



gefordert gewesen, so dass sich die Allelfrequenz nur sehr langsam ändert. Interessanterweise führt dies ebenfalls zu einer Grenzodynamik. In diesem Fall ist nämlich als einfache Anwendung des Korollars (wobei wir  $\tilde{X}_t = X_{t/N}$  schreiben, so dass die Raten um einen Faktor  $N$  kleiner werden)  $\tilde{X}^N \xrightarrow{N \rightarrow \infty} \tilde{X}$ , wobei  $\tilde{X}$  die Differentialgleichung des deterministischen Reaktionsnetzwerkes, also

$$\dot{\tilde{X}} = \frac{1}{2}\tilde{X}(1 - \tilde{X}) - \frac{1}{2}\tilde{X}(1 - \tilde{X}) = 0$$

löst. Weiter ist es in diesem Fall möglich, für  $\tilde{X}_0 = x$  einen Grenzwert von  $\sqrt{N}(\tilde{X}^N - x)$  zu berechnen. Theorem 1.49 ergibt in diesem Fall einen Grenzwert  $U$ , der

$$dU = (1 - 2x)Udt + \sqrt{\frac{1}{2}x(1-x)}dW_1 - (1 - 2x)Udt + \sqrt{\frac{1}{2}x(1-x)}dW_2 = \sqrt{x(1-x)}dW$$

für geeignete Brown'sche Bewegungen  $W_1, W_2, W$  folgt. Entscheidend ist in beiden Grenzwerten, dass die Zeit deutlich langsamer läuft als im Grenzwert der Proposition 2.3. Der letzte Grenzwert passt andersherum auch zur Lösung von  $dX = \sqrt{X(1-X)}dW$ . Betrachtet man diesen Prozess und berechnet den Generator  $G_U$  von  $U = \sqrt{N}(X_{t/N} - x)$ , so erhält man

$$\begin{aligned} G_U f(u) &= \lim_{h \rightarrow 0} \frac{\mathbf{E}_x[f(\sqrt{N}X_{h/N})] - f(\sqrt{N}x)}{h} = \frac{1}{N}G_X f(\sqrt{N}x) = \frac{1}{2N}Nx(1-x)f''(u) \\ &= \frac{1}{2}x(1-x)f''(u). \end{aligned}$$

**Proposition 2.5.** Die Größe  $2X(1-X)$  heißt Heterozygotie, da sie die Wahrscheinlichkeit beschreibt, beim zweimaligen Ziehen (mit Zurücklegen) aus der Population zwei verschiedene Typen zu ziehen. Ist  $X(0) = x$ , so ist

$$\mathbf{E}_x[X_t(1-X_t)] = x(1-x)e^{-t}.$$

**Bemerkung 2.6** (Abnahme der Heterozygotie in der Wright-Fisher-Diffusion). Dieselbe Formel gilt auch, wenn  $X$  der Wright-Fisher-Diffusion aus (2.1) folgt.

Denn:  $X$  hat den Generator  $Gf(x) = \frac{1}{2}x(1-x)f''(x)$ . Aus der Martingaleigenschaft aus Theorem 1.9 folgt

$$\mathbf{E}_x[X_t(1-X_t)] - x(1-x) = \int_0^t \mathbf{E}_x[GX_s(1-X_s)]ds = - \int_0^t \mathbf{E}_x[X_s(1-X_s)]ds. \quad (2.3)$$

Daraus folgt die exponentielle Abnahme der Heterozygotie.

*Beweis.* Wir wenden den Generator auf  $f(x) = x(1-x)$  an und erhalten

$$\begin{aligned} Gf(x) &= \frac{1}{2}x(1-x)N^2\left(f\left(x + \frac{1}{N}\right) + f\left(x - \frac{1}{N}\right) - 2f(x)\right) \\ &= \frac{1}{2}x(1-x)N^2\left(\left(x + \frac{1}{N}\right)\left(1 - x - \frac{1}{N}\right) + \left(x - \frac{1}{N}\right)\left(1 - x + \frac{1}{N}\right) - 2x(1-x)\right) \\ &= -x(1-x), \end{aligned}$$

also folgt die Aussage genau wie in (2.3).  $\square$

## Geburts-Todes-Prozesse

Das zwei-Allel-Moran-Modell ist ein Geburts-Todesprozess. Solche Prozesse kommen in Anwendungen häufig vor, deshalb werden wir sie nun ein wenig beleuchten.

**Definition 2.7** (Geburts-Todes-Prozess). *Ein Geburts-Todes-Prozess  $X$  ist ein Markov-Sprungprozess mit Zustandsraum  $E = \mathbb{Z}_+$ , wobei der Prozess aus  $X_t = n$  nach*

$$\begin{aligned} n+1 &\text{ mit Rate } \lambda_n, \\ n-1 &\text{ mit Rate } \mu_n \end{aligned}$$

(mit  $\mu_0 = 0$ ) springt.

Im Folgenden treten immer wieder Produkte  $a_n \cdots a_i$  auf. Dies werden wir für  $i > n$  als leeres Produkt, also als 1 interpretieren. Genauso definieren wir leere Summen als 0.

**Proposition 2.8** (Fixationswahrscheinlichkeit eines Geburts-Todes-Prozesses). *Sei  $X$  ein Geburts-Todes-Prozess mit Geburtsraten  $(\lambda_n)_{n=0,1,2,\dots}$  und Todesraten  $(\mu_n)_{n=1,2,\dots}$ . Sei weiter  $T_N := \inf\{t : X_t = N\}$  (mit  $\inf \emptyset = \infty$ ). Dann gilt*

$$\mathbf{P}_n(T_N < T_0) = \frac{\sum_{i=0}^{n-1} \frac{\mu_i \cdots \mu_1}{\lambda_i \cdots \lambda_1}}{\sum_{i=0}^{N-1} \frac{\mu_i \cdots \mu_1}{\lambda_i \cdots \lambda_1}}.$$

*Beweis.* Wir suchen zunächst ein  $S$ , so dass  $S(0) = 0$ ,  $S(1) = 1$  und  $S(X)$  ein Martingal ist. Hierzu muss  $GS = 0$  sein, wie aus der Martingaleigenschaft aus Theorem 1.9 folgt. Dies ergibt für alle  $n$

$$\lambda_n(S(n+1) - S(n)) + \mu_n(S(n-1) - S(n)) = 0,$$

also

$$S(n+1) - S(n) = \frac{\mu_n}{\lambda_n}(S(n) - S(n-1)).$$

Durch die Vorgabe von  $S(1) - S(0) = 1$  entsteht dadurch

$$S(n+1) - S(n) = \frac{\mu_n \cdots \mu_1}{\lambda_n \cdots \lambda_1},$$

also

$$S(n) = \sum_{i=0}^{n-1} \frac{\mu_i \cdots \mu_1}{\lambda_i \cdots \lambda_1}.$$

Für  $T = T_0 \wedge T_N$  ist  $S_{t \wedge T}$  ein nicht-negatives Martingal, das also fast sicher und in  $L^1$  (da es beschränkt ist) konvergiert. Es muss also sicher gegen 0 oder  $N$  konvergieren, es gilt also, falls

$$S(k) = \mathbf{E}_k[S_T] = 0 \cdot \mathbf{P}_k[T_0 < T_N] + S(N) \cdot \mathbf{P}_k[T_N < T_0],$$

woraus die Behauptung folgt.  $\square$

**Proposition 2.9** (Stationäre Verteilung eines Geburts-Todes-Prozesses). *Sei  $X$  ein Geburts-Todes-Prozess mit Geburtsraten  $(\lambda_n)_{n=0,1,2,\dots}$  und Todesraten  $(\mu_n)_{n=1,2,\dots}$ . Sei weiter  $T_N := \inf\{t : X_t = N\}$  (mit  $\inf \emptyset = \infty$ ). Dann ist  $X$  genau dann positiv rekurrent, wenn*

$$\sum_{n=1}^{\infty} \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} < \infty.$$

In diesem Fall ist die Verteilung

$$\pi(n) = \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1} \pi(0), \quad n = 1, 2, \dots$$

mit

$$\pi(0) = \frac{1}{\sum_{n=1}^{\infty} \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1}}$$

stationär. Weiterhin ist diese Verteilung reversibel.

*Beweis.* Wir wissen bereits, dass stationäre Verteilungen  $\pi$  die Bedingung – genau wie in (1.4)

$$(\lambda_k + \mu_k \pi(k) = \lambda_{k-1} \pi(k-1) + \mu_{k+1} \pi(k+1)).$$

erfüllen müssen. Diese Rekursion lösen wir ähnlich wie im letzten Beweis. Wir schreiben

$$\mu_{k+1} \pi(k+1) - \mu_k \pi(k) = \lambda_k \pi(k) - \lambda_{k-1} \pi(k-1)$$

Summiert man beide Seiten über  $k$  bis  $n$  auf, so erhält man

$$\begin{aligned} \mu_{n+1} \pi(n+1) &= \lambda_n \pi(n), \\ \frac{\pi(n)}{\pi(0)} &= \frac{\lambda_{n-1} \cdots \lambda_0}{\mu_n \cdots \mu_1}. \end{aligned}$$

Summiert man über  $n$ , ergibt sich  $\pi(0)$  und damit die stationäre Verteilung (falls  $\pi(0) > 0$ ). Für die Reversibilität bemerken wir, dass

$$\lambda_n \frac{\pi(n)}{\pi(0)} = \frac{\lambda_n \cdots \lambda_0}{\mu_n \cdots \mu_0} = \mu_{n+1} \frac{\pi(n+1)}{\pi(0)}.$$

$\square$

**Proposition 2.10** (Fixationswahrzeit eines Geburts-Todes-Prozesses). *Sei  $X$  ein Geburts-Todes-Prozess mit Geburtsraten  $(\lambda_n)_{n=0,1,2,\dots}$  und Todesraten  $(\mu_n)_{n=1,2,\dots}$ . Sei weiter  $T_n := \inf\{t : X_t = n\}$  (mit  $\inf \emptyset = \infty$ ). Dann ist für  $0 \leq k \leq N$  und  $T := T_0 \wedge T_N$*

$$\mathbf{E}_n[T] = \sum_{k=1}^{N-1} \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i} \frac{\sum_{k=0}^{n-1} \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1}}{\sum_{k=0}^{N-1} \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1}} - \sum_{k=1}^{n-1} \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i}.$$

*Beweis.* Zunächst ist klar, dass  $\mathbf{E}_N[T] = \mathbf{E}_0[T] = 0$ . Wieder setzen wir eine Rekursion an, nämlich für  $1 \leq n \leq N-1$

$$\begin{aligned}\mathbf{E}_k[T] &= \frac{1}{\lambda_k + \mu_k} + \frac{\lambda_k}{\lambda_k + \mu_k} \mathbf{E}_{k+1}[T] + \frac{\mu_k}{\lambda_k + \mu_k} \mathbf{E}_{k-1}[T], \\ \mathbf{E}_{k+1}[T] - \mathbf{E}_k[T] &= \frac{\mu_k}{\lambda_k} (\mathbf{E}_k[T] - \mathbf{E}_{k-1}[T]) - \frac{1}{\lambda_k}.\end{aligned}$$

Dies ergibt

$$\begin{aligned}\mathbf{E}_{k+1}[T] - \mathbf{E}_k[T] &= \frac{\mu_k}{\lambda_k} (\mathbf{E}_k[T] - \mathbf{E}_{k-1}[T]) - \frac{1}{\lambda_k} \\ &= \frac{\mu_k}{\lambda_k} \frac{\mu_{k-1}}{\lambda_{k-1}} (\mathbf{E}_{k-1}[T] - \mathbf{E}_{k-2}[T]) - \frac{\mu_k}{\lambda_k \lambda_{k-1}} - \frac{1}{\lambda_k} \\ &= \dots = \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1} \mathbf{E}_1[T] - \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i}.\end{aligned}$$

Summation von  $k=0$  bis  $k=N-1$  ergibt

$$\sum_{k=0}^{N-1} \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1} \mathbf{E}_1[T] = \sum_{k=0}^{N-1} \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i},$$

also

$$\mathbf{E}_n[T] = \sum_{k=0}^{N-1} \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i} \frac{\sum_{k=0}^{n-1} \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1}}{\sum_{k=0}^{N-1} \frac{\mu_k \cdots \mu_1}{\lambda_k \cdots \lambda_1}} - \sum_{k=0}^{n-1} \sum_{i=1}^k \frac{\mu_k \cdots \mu_{i+1}}{\lambda_k \cdots \lambda_i}.$$

□

Für den Häufigkeitsprozess  $X$  im 2-Allel-Moran-Modell ist  $NX$  ein Geburts-Todesprozess mit  $\lambda_n = \mu_n = \frac{1}{2}n(N-n)$  ist. Deshalb können wir nun Fixationswahrscheinlichkeiten und -zeiten berechnen.

**Proposition 2.11** (Fixationswahrscheinlichkeit und -zeit im Moran-Modell). *Die Fixationswahrscheinlichkeit im 2-Allel-Moran-Modell ist*

$$\mathbf{P}_n[T_N < T_0] = \frac{n}{N}.$$

*Die Fixationszeit ist*

$$\mathbf{E}_n[T] = 2 \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i} + \frac{N-n}{N} \sum_{i=N-n}^{N-1} \frac{1}{i}.$$

**Bemerkung 2.12.** Ist  $x = n/N$  und  $N$  groß, so ist

$$\begin{aligned}\mathbf{E}_n[T] &= 2x(\log N - \log n) + 2(1-x)(\log N - \log(N-n)) + o(1) \\ &= -2x \log x - 2(1-x) \log(1-x) + o(1).\end{aligned}$$

*Beweis.* Für die Fixationswahrscheinlichkeit ist wegen  $\lambda_n = \mu_n$  aus Proposition 2.8

$$\mathbf{P}_n[T_N < T_0] = \frac{\sum_{i=0}^{n-1} 1}{\sum_{i=0}^{N-1} 1} = \frac{n}{N}.$$

Die Fixationszeit berechnen wir aus Proposition 2.10 mittels

$$\begin{aligned} \mathbf{E}_n[T] &= \sum_{k=1}^{N-1} \sum_{i=1}^k \frac{2}{i(N-i)} \frac{n}{N} - \sum_{k=1}^{n-1} \sum_{i=1}^k \frac{2}{i(N-i)} \\ &= \frac{n}{N} \sum_{i=1}^{N-1} \sum_{k=i}^{N-1} \frac{2}{i(N-i)} - \sum_{i=1}^{n-1} \sum_{k=i}^{n-1} \frac{2}{i(N-i)} \\ &= \frac{n}{N} \sum_{i=1}^{N-1} \frac{2}{i} - \sum_{i=1}^{n-1} \frac{2(n-i)}{i(N-i)} \\ &= \frac{n}{N} \sum_{i=1}^{N-1} \frac{2}{i} + \frac{N-n}{N} \sum_{i=1}^{n-1} \left( \frac{2}{i} + \frac{2}{N-i} \right) - \sum_{i=1}^{n-1} \frac{2}{i} \\ &= \frac{n}{N} \sum_{i=n}^{N-1} \frac{2}{i} + \frac{N-n}{N} \sum_{i=N-n}^{N-1} \frac{2}{i}. \end{aligned}$$

□

## 2.2 Der Coaleszent

Aus der grafischen Repräsentation des Moran-Modells kann man Vorfahrens-Beziehungen ablesen. Diese erweisen sich bei der Analyse des Modells (oder auch bei Datenauswertungen) als äußerst hilfreich. Entscheidend ist folgende Beobachtung:

Verfolgt man die Vorfahren von  $n$  Individuen gleichzeitig im Coaleszenten zurück, so findet zwischen jedem der  $\binom{n}{2}$  Paare mit Rate 1 ein Resampling-Ereignis statt. Die Individuen, die in diesem Ereignis involviert sind, finden dadurch einen gemeinsamen Vorfahren.

**Definition 2.13** ( $n$ -Coaleszent). *Der  $n$ -Coaleszent  $\mathcal{T}_n$  ist ein zufälliger (ultrametrischer<sup>5</sup>) Baum, der folgendermaßen zustande kommt: Ausgehend von  $n$  Linien verschmilzt jedes Paar von Linien mit Rate 1. Sind noch  $2 \leq k \leq n$  Linien übrig, ist also die Gesamtrate, mit der eine Verschmelzung stattfindet,  $\binom{k}{2}$ . Jedes der  $\binom{k}{2}$  Paare hat dabei die gleiche Wahrscheinlichkeit, zu verschmelzen. Der Prozess endet, wenn nur noch eine Linie übrig ist.*

Die Zeit, bei der der Coaleszent zum ersten mal  $k$  Linien hat, bezeichnen wir mit  $S_k$  und  $T_k = S_k - S_{k+1} \sim \exp\left(\binom{k}{2}\right)$  ist die Zeitdauer, die der Coaleszent mit genau  $k$  Linien verbringt.

Den Prozess, der die Linien im Coaleszenten zählt, nennen wir  $(K_t)_{t \geq 0}$ . Er ist ein reiner Todesprozess mit  $\mu_k = \binom{k}{2}$ .

**Bemerkung 2.14.** 1. Ein paar Realisierungen des Coaleszenten für  $n = 4$  und  $n = 20$  sind in Abbildung 2.2 abgedruckt.

<sup>5</sup>Bekanntlich ist eine Ultrametrik eine Metrik  $r$ , bei der die Dreiecksungleichung zu  $r(x, z) \leq r(x, y) \vee r(y, z)$  verschärft werden kann. Ein ultrametrischer Baum ist durch die Metrik auf den Blättern gegeben, d.h. der Abstand zwischen zwei Blättern im Baum ist die Weglänge im Baum von einem zum anderen Blatt.

2. Unsere Konstruktion des Coaleszenten geht von den Blättern des Baumes aus und endet bei der Wurzel. Stochastisch äquivalent ist die Beschreibung von der Wurzel zu den Blättern: startend mit der Wurzel und zwei daranhängenden Linien vergeht im Zustand von  $k$  Linien eine  $\text{exp} \binom{k}{2}$ -verteilte Zeit. Ist diese zu Ende, wird eine zufällig ausgewählte Linie verzweigt. Dies geschieht bis zu der Zeit, bis kurz bevor die  $n + 1$ -te Linie erzeugt wird.
3. An dieser Vorwärts-Konstruktion sieht man, dass es auch den Coaleszent mit unendlich vielen Blättern gibt. Es ist nämlich

$$\mathbf{E} \left[ \sum_{k=2}^{\infty} T_k \right] = \sum_{k=2}^{\infty} \frac{1}{\binom{k}{2}} < \infty,$$

also insbesondere  $\mathbf{P} \left( \sum_{k=2}^{\infty} T_k < \infty \right) = 1$ .

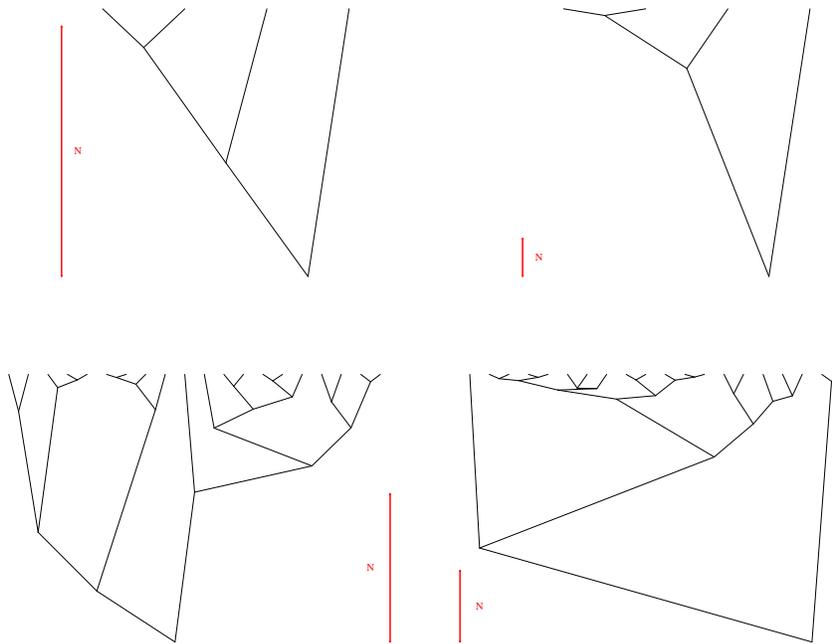


Abbildung 2.2: Coaleszenten für  $n = 4$  und  $n = 20$  Blätter.

**Proposition 2.15** (Eigenschaften des  $n$ -Coaleszenten). *Für einen  $n$ -Coaleszenten sei  $S_1$  die Zeit des jüngsten gemeinsamen Stichproben-Vorfahrens, und*

$$L_n := \sum_{k=2}^n kT_k$$

*die gesamte Baumlänge.*

1. Es gilt

$$\begin{aligned}\mathbf{E}[S_1] &= 2\left(1 - \frac{1}{n}\right), \\ \mathbf{V}[S_1] &= 8\left(\sum_{k=2}^n \frac{1}{k^2}\right) - 4\left(1 - \frac{1}{n}\right)^2.\end{aligned}$$

2. Seien  $X_1, \dots, X_{n-1} \sim \exp(1)$  unabhängig. Dann ist  $L_n \sim 2 \max_{i=1, \dots, n-1} X_i$ . Weiter gilt

$$\begin{aligned}\mathbf{E}[L_n] &= 2 \sum_{k=1}^{n-1} \frac{1}{k}, & \mathbf{V}[L_n] &= 4 \sum_{k=1}^{n-1} \frac{1}{k^2}, \\ \frac{1}{2}L_n - \log n &\xrightarrow{N \rightarrow \infty} G,\end{aligned}$$

wobei  $G$  Gumbel-verteilt ist, d.h. die Verteilungsfunktion von  $G$  ist  $F(x) = e^{-e^{-x}}$ .

*Beweis.* Wir erinnern daran, dass  $S_1 = T_n + \dots + T_2$  und  $T_k \sim \exp\left(\frac{1}{k}\right)$ . 1. Daraus folgt

$$\begin{aligned}\mathbf{E}[S_1] &= \sum_{k=2}^n \frac{2}{k(k-1)} = 2 \sum_{k=2}^n \frac{1}{k-1} - \frac{1}{k} = 2\left(1 - \frac{1}{n}\right), \\ \mathbf{V}[S_1] &= \sum_{k=2}^n \frac{4}{k^2(k-1)^2} = 4 \sum_{k=2}^n \left(\frac{1}{(k-1)} - \frac{1}{k}\right)^2 \\ &= 4 \left[ 2 \left(\sum_{k=2}^n \frac{1}{k^2}\right) + 1 - \frac{1}{n^2} - \sum_{k=2}^{n-1} \frac{2}{k(k-1)} \right] \\ &= 8 \left(\sum_{k=2}^n \frac{1}{k^2}\right) - 4\left(1 - \frac{1}{n}\right)^2.\end{aligned}$$

Es gilt  $kT_k \sim \exp((k-1)/2)$ . Daraus folgt bereits

$$\begin{aligned}\mathbf{E}[L_n] &= \sum_{k=2}^n \frac{2}{k-1} = 2 \sum_{k=1}^{n-1} \frac{1}{k}, \\ \mathbf{V}[L_n] &= 4 \sum_{k=1}^{n-1} \frac{1}{k^2}.\end{aligned}$$

Für die Darstellung  $L_n = 2 \max_{i=1, \dots, n-1} X_i$  bemerken wir folgendes. Die Zufallsgrößen  $X_1, \dots, X_{n-1}$  sind die Zeiten, zu denen  $n-1$  unabhängige, exponentialverteilte Wecker klingeln. Die Gesamtrate, mit der ein Wecker klingelt, ist zunächst  $n-1$ . Nach Ablauf des ersten Weckers sind noch  $n-2$  übrig, so dass die Gesamtrate  $n-2$  ist. Der letzte Wecker klingelt also zu einer Zeit, die nach  $Y_{n-1} + \dots + Y_1$  verteilt ist, wobei  $Y_k \sim \exp(k)$ . Also ist

$$L_n = \sum_{k=2}^n kT_k \sim 2 \sum_{k=1}^{n-1} Y_k = 2 \max_{i=1, \dots, n-1} X_i.$$

Nun ist

$$\begin{aligned} \mathbf{P}\left(\frac{1}{2}L_n - \log n \leq x\right) &= \mathbf{P}\left(\max_{i=1, \dots, n-1} X_i - \log n \leq x\right) = (1 - e^{-(x+\log n)})^{n-1} \\ &= \left(1 - \frac{e^{-x}}{n}\right)^{n-1} \xrightarrow{n \rightarrow \infty} e^{-e^{-x}}. \end{aligned}$$

Daraus folgt die Konvergenz gegen die Gumbel-Verteilung.  $\square$

**Bemerkung 2.16** (Fixationszeit). In Bemerkung 2.16 haben wir bereits die Fixationszeit im Moran-Modell für große  $N$  berechnet. Dies können wir nochmals mittels des Coaleszenten tun. Grundlegend ist es hier, dass bei einer Startfrequenz von  $x$  die Fixationszeit von  $A$  ( $a$ ) vor  $t$  liegt genau dann, wenn alle Individuen zur Zeit  $t$  einen Vorfahren zur Zeit 0 haben, der  $A$  ( $a$ ) trägt. Deshalb gilt

$$\mathbf{E}_x[T] = \int_0^\infty \mathbf{P}_x[T > t] dt = \int_0^\infty \sum_{n=2}^\infty \mathbf{P}[K_t = n] (1 - x^n - (1 - x)^n) dt$$

und

$$\int_0^\infty \mathbf{P}_x[K_t = n] dt = \mathbf{E}_x \left[ \int_0^\infty 1_{K_t=n} dt \right] = \mathbf{E}[T_n] = \frac{2}{n(n-1)}$$

Nun ist also

$$\begin{aligned} \sum_{n=2}^\infty \frac{1}{n(n-1)} x^n &= \int_0^x \int_0^y \sum_{n=2}^\infty z^{n-2} dz = - \int_0^x \log(1-y) dy \\ &= (1-y) \log(1-y) + y \Big|_0^x = (1-x) \log(1-x) + x \end{aligned}$$

und damit

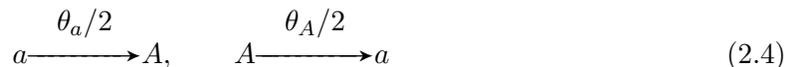
$$\begin{aligned} \mathbf{E}_x[T] &= \sum_{n=2}^\infty \frac{2}{n(n-1)} (1 - x^n - (1-x)^n) \\ &= -2(-1 + (1-x) \log(1-x) + x + x \log(x) + 1 - x) \\ &= -2(x \log(x) + (1-x) \log(1-x)) \end{aligned}$$

### 2.3 Mutationsmodelle

Aus der Biologie ist bekannt, dass sich bei der Vererbung von Genmaterial von einer auf die nächste Generation Fehler einschleichen, wobei man auch von Mutationen spricht.

**Definition 2.17** (Moran-Modell mit bi-allelischer Mutation). *Im Moran-Modell mit bi-allelischer Mutation gibt es neben Resampling-Ereignissen wie in Definition 2.1 auch Mutationsereignisse. Es gibt Raten  $\theta_A$  und  $\theta_a$ , so dass der Zustand  $X = (x_1, \dots, x_N)$  für jedes  $i$  mit Rate  $\theta_A/2$  nach  $(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_N)$  und mit Rate  $\theta_a/2$  nach  $(x_1, \dots, x_{i-1}, A, x_{i+1}, \dots, x_N)$  springt.*

**Bemerkung 2.18.** Das Moran-Modell mit bi-allelischer Mutation ist mit anderen Worten ein chemisches Reaktionsnetzwerk, bei dem neben den Reaktionen 2.2 noch die Reaktionen



stattfinden.

**Proposition 2.19** (Gleichgewichtsverteilung). *Die Allelfrequenz von A in einem Moran-Modell mit bi-allelischer Mutation ist ein Geburts-Todes-Prozess, der von  $NX_t = n$  nach*

$$\begin{aligned} n+1 & \text{ mit Rate } \frac{1}{2}n(N-n) + \frac{1}{2}\theta_a(N-n), \\ n-1 & \text{ mit Rate } \frac{1}{2}n(N-n) + \frac{1}{2}\theta_A n \end{aligned}$$

*springt. Dieser besitzt eine Gleichgewichtsverteilung, nämlich*

$$\frac{\pi(n)}{\pi(0)} = \prod_{i=1}^n \left(1 + \frac{\theta_a - 1}{i}\right) \left(1 - \frac{\theta_A - 1}{\theta_A + N - i}\right).$$

**Bemerkung 2.20.** Betrachten wir die stationäre Verteilung für große  $N$ . Für  $n = xN$  und große  $N$  ist

$$\begin{aligned} \frac{\pi(xN)}{\pi(0)} & \approx \exp\left(\sum_{i=1}^{xN} \log\left(1 + \frac{\theta_a - 1}{i}\right) + \log\left(1 - \frac{\theta_A - 1}{\theta_A + N - i}\right)\right) \\ & \approx \exp\left(\sum_{i=1}^{xN} \frac{\theta_a - 1}{i} - \frac{\theta_A - 1}{\theta_A + N - i}\right) \\ & \approx \exp\left((\theta_a - 1)\log(xN) - (\theta_A - 1)(\log N - \log((1-x)N))\right) \\ & \sim x^{\theta_a - 1}(1-x)^{\theta_A - 1}. \end{aligned}$$

Dies legt nahe, dass für große  $N$  die Gleichgewichtsverteilung die Dichte

$$f(x) = \frac{1}{\beta(\theta_a, \theta_A)} x^{\theta_a - 1} (1-x)^{\theta_A - 1}$$

besitzt. Mit anderen Worten ist sie eine Beta-Verteilung.

*Beweis.* Wir verwenden Proposition 2.9 mit

$$\begin{aligned} \lambda_n & = \frac{1}{2}n(N-n) + \frac{1}{2}\theta_a(N-n) = \frac{1}{2}(n + \theta_a)(N-n), \\ \mu_n & = \frac{1}{2}n(N-n) + \frac{1}{2}\theta_A n = \frac{1}{2}(N-n + \theta_A)n. \end{aligned}$$

Wir schreiben direkt

$$\begin{aligned} \frac{\pi(n)}{\pi(0)} & = \frac{(\theta_a + n - 1) \cdots \theta_a}{n \cdots 1} \frac{N \cdots (N - n + 1)}{(\theta_A + N - 1) \cdots (\theta_A + N - n)} \\ & = \prod_{i=1}^n \left(1 + \frac{\theta_a - 1}{i}\right) \left(1 - \frac{\theta_A - 1}{\theta_A + N - i}\right). \end{aligned}$$

□

**Bemerkung 2.21.** Während das bi-allelische Moran-Modell mathematisch gut analysiert werden kann, ist es jedoch biologisch eher unrealistisch. Ein Beispiel genetischer Daten sieht man etwa in Abbildung 2.3. Anzumerken ist jedoch, dass diese Daten aus den Zeiten stammt, bei denen die DNA-Sequenzierung recht teuer war. Heutzutage werden genetische Daten nicht mehr abgedruckt, sondern in Datenbanken zur Analyse geladen.

Reference sequence	5' Flanking sequence	Adult leader (exon 1)	Intron 1 (Adult intron, larval non-coding)	Larval leader	Translated region of exon 2	Intron 2	Exon 3	Intron 3	Translated region of exon 4	3'-Untranslated region	3' Flanking sequence
	C C G		C A A T A T G G G T C V 2 G	C	T A C	C C C C	G G A A T	C T C C A C T A G	A V 3 C	A G C V 4 C V 5 T A 6	
Strain											
Wa-S	. . .	. . . . . A T . . . . .	. . . . .	. . . . .	. . . . .	. . . . .	T T . A	C A . T A	A C . . . . .	. . . . .	. . . . . 3
Fl-S	. . . C	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	T T . A	C A . T A	A C . . . . .	. . . . .	. . . . . 3
At-S	. . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . . A	. . . . .	. . . T V . 1 A .
Fr-S	. . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	G T . . . . .	. . . . .	. . . . .	. . . . . A	. . . . . T A . . . . .
Fl-Ss	. . .	A G . . . A . T C . . . . .	. . . . .	A	G	G T . . . . .	. . . . .	. . . . .	. . . . .	. . . . .	C 3 . . . . .
Ja-S	. . . C	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . T . T . C A	. . . . .	C 4 . . . . . T . . .
Fl-P	. . . C	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . G T C T C C	. . . . .	C 4 . . . . .
Fr-P	T G C	A G . . . A . T C V G V .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . G T C T C C	. . . . .	C 4 G . . . . .
Wa-P	T G C	A G . . . A . T C V G V .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . G T C T C C	. . . . .	C 4 G . . . . .
At-P	T G C	A G . . . A . T C V G V .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . G T C T C C	. . . . .	C 5 G . . . . .
Ja-P	T G C	A G G G G A . . . V . . T	. . . . .	. . . . .	. . . . .	. . . . .	. . . A .	. . . G . . .	. . . G T C T C C	. . . . .	C 4 . . . . . -1 . . .
No. of polymorphic sites	3	0	11	1	1	2	4	5	9	2	5
Average no. of nucleotides compared	63	87	620	70	99	65	405	70	204	178	767
% Sites polymorphic	4.7	0	1.8	1.4	1.0	3.1	1.0	7.1	3.5	1.1	0.6

Abbildung 2.3: Daten aus dem Paper von M. Kreitman. Nucleotide polymorphism at the alcohol dehydrogenase locus of *Drosophila melanogaster*. Nature, Vol. 304, 412-417, 1983. Damals war das Sequenzieren von DNA sehr teuer, so dass man nur wenige Basen in kleinen Stichproben sequenziert hat. Jede Zeile steht hier für den adh-locus einer Fliege. Die Referenzabfolge der Basen steht in der ersten Zeile, darunter sind jeweils nur Unterschiede zur Referenz aufgelistet. Wie man sieht, gibt es an jedem Site höchstens zwei verschiedene Allele, was eine Grundvoraussetzung für die Gültigkeit des unendlich-viele-Sites-Modells ist.

Um ein realistischeres Modell für genetische Daten zu bekommen, definieren wir das Modell für unendlich viele Sites.

**Definition 2.22** (Der Coaleszent im unendlich-viele-Sites-Modell). Sei  $\mathcal{T}_n$  ein Coaleszent wie in Definition 2.13 und  $\theta > 0$ . Beginnend an der Wurzel des Baumes definieren wir eine  $\mathcal{T}_n$ -indizierte Markov-Kette  $\Lambda$  mit Zustandsraum  $\mathcal{N}_f(I)^6$  für  $I = [0, 1]$ . Die Übergänge sind dabei von  $\lambda$  nach  $\lambda + \delta_U$ , wobei  $U \sim U(I)$ , mit Rate  $\theta/2$ . Solche  $x \in I$ , die eine  $\delta$ -Masse erhalten, heißen auch SNPs (Single Nucleotide Polymorphisms).

**Bemerkung 2.23.** Eine Markov-Kette ist baum-indiziert, wenn die Zeit nicht linear, sondern eben entlang eines Baumes abläuft. An den Knoten sind dabei linker und rechter Teilast unabhängig gegeben den Zustand am Knoten.

Wir haben bereits berechnet, dass  $\mathcal{T}_n$  eine erwartete Länge von  $\mathbf{E}[L(\mathcal{T}_n)] = 2 \sum_{k=1}^{n-1} \frac{1}{k}$  hat. Insbesondere ist er fast sicher nur endlich lang. Das bedeutet für das unendlich-viele-Sites-Modell, dass fast sicher keine zwei SNPs dieselbe Stelle  $x \in I$  treffen.

Weiter sei  $S$  die Gesamtzahl der SNPs auf  $\mathcal{T}_n$ . Da dies ein Zählprozess ist, der nach exponentialverteilten Raten springt, ist  $S$  gegeben  $L(\mathcal{T}_n)$  Poisson-verteilt mit Parameter  $\theta L(\mathcal{T}_n)/2$ , also

$$\mathbf{E}[S] = \mathbf{E}[\mathbf{E}[S|L(\mathcal{T}_n)]] = \mathbf{E}[\theta L(\mathcal{T}_n)/2] = \theta \sum_{k=1}^{n-1} \frac{1}{k}.$$

Wir berechnen nun im unendlich-viele-Sites-Modell weitere Statistiken, die bei der Datenanalyse behilflich sein können.

<sup>6</sup>Wir bezeichnen die Menge der endlichen Zählmaße auf einem metrischen Raum  $(E, r)$  mit  $\mathcal{N}_f(E)$ .

**Definition 2.24** (Das Frequenzspektrum). Sei  $\Lambda = (\Lambda_t)_{t \in \mathcal{T}_n}$  die baum-indizierte Markov-Kette aus Definition 2.22 und  $\mathcal{L}_n \subseteq \mathcal{T}_n$  die Menge der Blätter in  $\mathcal{T}_n$ . Wir definieren

$$\Lambda_{\mathcal{L}_n} := \sum_{t \in \mathcal{L}_n} \Lambda_t$$

und damit

$$S_i := \frac{1}{i} \int 1_{\Lambda_{\mathcal{L}_n}(dx)=i} \Lambda_{\mathcal{L}_n}(dx),$$

also die Anzahl der SNPs, die in genau  $k$  der  $n$  Genome  $\Lambda_t, t \in \mathcal{L}_n$  enthalten sind.

**Bemerkung 2.25** (Polya-Urne). Aus der Stochastik ist die Polya-Urne bekannt, die uns im Beweis hilfreich sein wird. Wir stellen uns eine Urne mit  $k$  Kugeln vor, die alle eine unterschiedliche Farbe besitzen. Wir ziehen eine Kugel, notieren die Farbe, und legen die Kugel zusammen mit einer Kugel derselben Farbe zurück. Dies wiederholen wir, bis  $n$  Kugeln in der Urne sind. Wenn wir uns nun fragen, wie groß die Wahrscheinlichkeit ist, dass  $i$  blaue Kugeln in der Urne am Ende (wenn  $n$  Kugeln in der Urne sind) sind, wenn am Anfang (wenn  $k$  Kugeln in der Urne sind) eine blaue Kugel enthalten ist, so lässt sich dies folgendermaßen berechnen: Insgesamt müssen in den  $n-k$  Ziehungen von Kugeln aus der Urne genau  $i-1$ -mal eine blaue gezogen worden sein, damit am Ende  $i$  blaue Kugeln in der Urne sind. Betrachten wir etwa die Möglichkeit, dass alle  $i-1$  blauen Kugeln in die Urne gelegt werden, bevor irgendeine andersfarbige Kugel gezogen wird, so erhalten wir

$$\frac{1 \cdots (i-1) \cdot (k-1) \cdots (n-i-1)}{k \cdots (n-1)}.$$

Insgesamt gibt es aber  $\binom{n-k}{i-1}$  Möglichkeiten, in welchen Schritten die blauen Kugeln in die Urne gelegt werden können. Deshalb ist die oben gesuchte Wahrscheinlichkeit gleich

$$\binom{n-k}{i-1} \frac{(i-1)!(k-1) \cdots (n-i-1)}{k \cdots (n-1)} = \frac{k-1}{i} \binom{n-k}{i-1} \frac{i!}{(n-i) \cdots (n-1)} = \frac{\binom{n-k}{i-1}}{\binom{n-1}{i}} \frac{k-1}{i}.$$

**Proposition 2.26.** Im Coaleszent im unendlich-viele-Sites-Modell ist das erwartete Frequenzspektrum gegeben durch

$$\mathbf{E}[S_i] = \frac{\theta}{i}.$$

*Beweis.* Im unendlich-viele-Sites-Modell sagen wir, ein SNP  $x \in I$  habe Größe  $i$ , wenn er in genau  $i$  der  $n$  Blätter von  $\mathcal{T}_n$  vorkommt. Außerdem sagen wir, der Coaleszent ist in Zustand  $k$ , wenn er gerade  $k$  Linien hat. Dabei muss man bedenken, dass jeder SNP in einer Zeit im Coaleszenten vorkommt, in der dieser einen Zustand  $k$  hat und der SNP auf einer Linie sitzen muss, die insgesamt zu  $i$  Blättern führt. Nun ist es so, dass man den Coaleszenten auch von  $k \leq n$  nach  $n$  Linien generieren kann, indem man bei  $k$  Linien eine  $\exp\left(\frac{k}{2}\right)$ -verteilte Zeit wartet und dann eine zufällig ausgewählte Linie nimmt, und diese verzweigt. Das Verfahren endet dann kurz bevor die Verzweigung in  $n+1$  Linien ientreten würde. Mit anderen Worten ist die Wahrscheinlichkeit, dass eine von  $k$  Linien genau zu  $i$  Blättern führt, dieselbe wie die oben berechnete Wahrscheinlichkeit, dass eine Farbe in der Polya-Urne zu der Zeit, wenn  $n$  Kugeln in der Urne sind, genau  $i$ -mal vorkommt, wenn zu der Zeit, als in der Urne  $k$  Linien waren, genau einmal vorkam, also

$$\mathbf{P}[l\text{-te Linie im Zustand } k \text{ hat Größe } i] = \frac{\binom{n-k}{i-1}}{\binom{n-1}{i}} \frac{k-1}{i}.$$

Insgesamt folgt

$$\begin{aligned}
\mathbb{E}[S_i] &= \sum_{k=2}^n \sum_{l=1}^k \mathbb{P}[l\text{-te Linie im Zustand } k \text{ hat Größe } i] \cdot \\
&\quad \mathbb{E}[\text{Anzahl der Mutation auf der } i\text{-ten Linie im Zustand } k] \\
&= \sum_{k=2}^n \sum_{l=1}^k \frac{\binom{n-k}{i-1}}{\binom{n-1}{i}} \frac{k-1}{i} \frac{\theta}{k(k-1)} \\
&= \frac{\theta}{i} \frac{1}{\binom{n-1}{i}} \sum_{k=2}^n \binom{n-k}{i-1} \\
&= \frac{\theta}{i} \frac{1}{\binom{n-1}{i}} \sum_{k=2}^n \left( \binom{n-(k-1)}{i} - \binom{n-k}{i} \right) \\
&= \frac{\theta}{i} \frac{1}{\binom{n-1}{i}} \left( \sum_{k=1}^{n-1} \binom{n-k}{i} - \sum_{k=2}^n \binom{n-k}{i} \right) = \frac{\theta}{i}.
\end{aligned}$$

□

## 2.4 Selektion

Bisher haben alle Typen die gleiche Chance gehabt, Nachkommen zu erzeugen. Biologisch besonders relevant sind jedoch Situationen, in denen dies nicht mehr so ist.

**Definition 2.27** (Moran-Modell mit Mutation und Selektion). *Im zwei-Typ-Moran-Modell mit Mutation und Selektion gibt es neben Resampling-Ereignissen wie in Definition 2.1 und den Mutationsereignissen wie in Definition 2.17 auch Selektionsereignisse. Hier wird im Zustand  $X = (x_1, \dots, x_N)$  für alle  $i, j$  mit Rate  $\alpha 1_{x_i=1}/N$  der Zustand nach  $(x_1, \dots, x_{j-1}, x_i = 1, x_{j+1}, \dots, x_N)$  geändert. Hier heißt  $\alpha$  auch Selektionskoeffizient.*

**Proposition 2.28.** *Sei  $X^N$  der Prozess der Allelfrequenz von Typ 1 im zwei-Typ-Moran-Modell mit Mutation und Selektion (mit Mutationsraten  $\theta_A, \theta_a$  und Selektionskoeffizient  $\alpha$ ). Dann gilt  $X^N \xrightarrow{N \rightarrow \infty} f d d X$ , wobei  $X$  die SDE*

$$dX = \frac{\theta_a}{2}(1-X)dt - \frac{\theta_A}{2}Xdt + \alpha X(1-X)dt + \sqrt{X(1-X)}dW \quad (2.5)$$

löst.

*Beweis.* Wieder rechnen wir nur die Konvergenz der Generatoren nach. Sei also  $f \in \mathcal{C}^2([0, 1])$ . Der Prozess  $X^N$  ist ein Geburts-Todes-Prozess mit Übergängen von  $x$  nach

$$\begin{aligned}
x + \frac{1}{N} &\text{ mit Rate } \frac{1}{2}N^2x(1-x) + \frac{1}{2}\theta_a N(1-x) + \frac{\alpha}{N}N^2x(1-x), \\
x + \frac{1}{N} &\text{ mit Rate } \frac{1}{2}N^2x(1-x) + \frac{1}{2}\theta_A N x.
\end{aligned}$$

Daraus folgt für den Generator  $G^N$  von  $X^N$

$$\begin{aligned}
G^N f(x) &= \frac{1}{2}N^2x(1-x)\left(f\left(x + \frac{1}{N}\right) + f\left(x - \frac{1}{N}\right) - 2f(x)\right) \\
&\quad + \left(\frac{1}{2}\theta_a N(1-x) + \alpha N x(1-x)\right)\left(f\left(x + \frac{1}{N}\right) - f(x)\right) + \frac{1}{2}\theta_A N x\left(f\left(x - \frac{1}{N}\right) - f(x)\right) \\
&\xrightarrow{N \rightarrow \infty} \frac{1}{2}x(1-x)f''(x) + \left(\frac{1}{2}\theta_a(1-x) - \frac{1}{2}\theta_A x + \alpha N x(1-x)\right)f'(x).
\end{aligned}$$

Daraus folgt die Behauptung. □

Um Typ-Frequenz  $X$  aus (2.5) analysieren zu können, brauchen wir die Theorie der ein-dimensionalen Diffusionen. Genau wie bei Geburts-Todes-Prozessen führt die ein-dimensionale Struktur dazu, dass man einiges explizit berechnen kann.

### Ein-dimensionale Diffusionen

Wir nennen einen zeithomogenen Markov-Prozess  $X$  mit Zustandsraum  $\mathbb{R}$  auch ein-dimensionale Diffusion, falls  $X$  stetige Pfade besitzt und

$$\begin{aligned}\mu(x) &:= \lim_{t \rightarrow 0} \frac{1}{t} (\mathbf{E}_x[X_t] - x), \\ \sigma^2(x) &:= \lim_{t \rightarrow 0} \frac{1}{t} \mathbf{V}_x[X_t]\end{aligned}$$

existieren und stetig sind, sowie

$$\lim_{t \rightarrow 0} \frac{1}{t} \mathbf{E}_x[(X_t - x)^k] = 0, \quad k = 3, 4, \dots$$

Hierbei heißt  $\mu$  die infinitesimale Drift und  $\sigma^2$  die infinitesimale Varianz von  $X$ . Man spricht auch von den (infinitesimalen) Charakteristiken von  $X$ . Ohne Beweis bemerken wir:

**Proposition 2.29.** *Sei  $X$  stochastischer Prozess mit mit Zustandsraum  $\mathbb{R}$ , sowie  $\mu, \sigma^2$  stetig. Dann sind äquivalent:*

1.  $X$  ist eine Diffusion mit infinitesimalen Charakteristiken  $\mu$  und  $\sigma^2$ .
2.  $X$  ist ein Markov-Prozess mit Generator

$$Gf(x) = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)$$

für  $f \in C_b^\infty(\mathbb{R})$ .

3.  $X$  löst die stochastische Differentialgleichung

$$dX = \mu(X)dt + \sqrt{\sigma^2(X)}dW.$$

**Proposition 2.30** (Rückwärts-Gleichung). *Sei  $X$  eine Diffusion mit Charakteristiken  $\mu, \sigma^2$  und  $g$  eine glatte Funktion. Dann erfüllt  $(t, x) \mapsto \mathbf{E}_x[g(X_t)]$  die partielle Differentialgleichung*

$$\frac{\partial u}{\partial t} = \mu(x) \frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(x) \frac{\partial^2 u}{\partial x^2}.$$

*Beweis.* Wir schreiben

$$\begin{aligned}\frac{\partial u(t, x)}{\partial t} &= \frac{1}{h} \mathbf{E}_x[g(X(t+h)) - g(X(t))] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{E}_x[\mathbf{E}_{X(t+h)}[g(X(t))] - g(X(t))] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \mathbf{E}_x[u(t, X(t+h)) - u(t, x)] \\ &= Gu(t, \cdot)(x) = \mu(x) \frac{\partial u(t, x)}{\partial x} + \frac{1}{2}\sigma^2(x) \frac{\partial^2 u(t, x)}{\partial x^2}.\end{aligned}$$

□

**Proposition 2.31.** Sei  $X$  die Lösung von (2.5) und  $g(x) = x(1-x)$ . Dann ist  $u(t, x) := \mathbf{E}_x[g(X_t)]$  die erwartete Heterozygotität zur Zeit  $t$  und es gilt

$$\frac{\partial}{\partial t} u(t, x) = \left(\frac{1}{2}\theta_a(1-x) - \frac{1}{2}\theta_A x + \alpha x(1-x)\right) \frac{\partial}{\partial x} u(t, x) + \frac{1}{2}x(1-x) \frac{\partial^2}{\partial x^2} u(t, x).$$

*Beweis.* Das ist eine direkte Anwendung der obigen Proposition.  $\square$

**Proposition 2.32** (Vorwärts-Gleichung). Sei  $X$  eine Diffusion mit Zustandsraum  $[l, r]$ , Charakteristiken  $\mu, \sigma^2$ , so dass für jedes  $t$  die Verteilung von  $X_t$  unter  $\mathbf{P}_x$  eine Dichte  $p(t, x, dy)$  besitzt mit  $p(t, x, l) = p(t, x, r) = 0$ . Dann löst  $p$  die partielle Differentialgleichung

$$\frac{\partial p(t, x, y)}{\partial t} = -\frac{\partial}{\partial y} (\mu(y)p(t, x, y)) + \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma^2(y)p(t, x, y)).$$

*Beweis.* Mit Hilfe der Chapman-Kolmogoroff-Gleichungen schreiben wir

$$\begin{aligned} \frac{\partial p(t, x, z)}{\partial t} &= \lim_{s \rightarrow 0} \frac{\partial}{\partial s} \int p(t, x, y) p(s, y, z) dy \\ &= \lim_{s \rightarrow 0} \int p(t, x, y) \frac{\partial p(s, y, z)}{\partial s} dy \\ &= \lim_{s \rightarrow 0} \int p(t, x, y) \left( \mu(y) \frac{\partial p(s, y, z)}{\partial y} + \frac{1}{2} \sigma^2(y) \frac{\partial^2 p(s, y, z)}{\partial y^2} \right) dy \\ &= \lim_{s \rightarrow 0} \int -p(s, y, z) \left( \frac{\partial}{\partial y} (p(t, x, y) \mu(y)) \right. \\ &\quad \left. - \frac{1}{2} \frac{\partial^2}{\partial y^2} (p(t, x, y) \sigma^2(y)) \right) \\ &= -\frac{\partial}{\partial z} (\mu(z)p(t, x, z)) + \frac{1}{2} \frac{\partial^2}{\partial z^2} (\sigma^2(z)p(t, x, z)). \end{aligned}$$

$\square$

**Proposition 2.33.** Sei  $X$  eine Diffusion mit Zustandsraum  $[l, r]$ , Charakteristiken  $\mu, \sigma^2$ , so dass für jedes  $t$  die Verteilung von  $X_t$  unter  $\mathbf{P}_x$  eine Dichte  $p(t, x, dy)$  besitzt mit  $p(t, x, l) = p(t, x, r) = 0$ . Falls es eine eindeutige Gleichgewichtsverteilung von  $X$  gibt und diese eine Dichte  $\psi$  besitzt, so gilt

$$\psi(y) = \frac{C}{\sigma^2(y)} \exp \left( 2 \int_{\eta}^y \frac{\mu(z)}{\sigma^2(z)} dz \right)$$

für ein beliebiges  $\eta \in (l, r)$  und geeignete Normierungskonstante  $C$ .

*Beweis.* Zunächst integrieren wir die Vorwärts-Gleichung und erhalten

$$\begin{aligned} \frac{\partial}{\partial t} P_x[X_t \leq y] &= \int_{-\infty}^y \left( -\frac{\partial}{\partial z} (\mu(z)p(t, x, z)) + \frac{1}{2} \frac{\partial^2}{\partial z^2} (\sigma^2(z)p(t, x, z)) \right) dz \\ &= -\mu(y)p(t, x, y) + \frac{1}{2} \frac{\partial}{\partial y} \sigma^2(y)p(t, x, y). \end{aligned}$$

Für  $t \rightarrow \infty$  ergibt sich hieraus  $p(t, x, y) \rightarrow \psi(y)$  und damit

$$-\mu(y)\psi(y) + \frac{1}{2}\frac{\partial}{\partial y}\sigma^2(y)\psi(y) = 0,$$

also

$$\psi(y) = \frac{C}{\sigma^2(y)} \exp\left(2 \int_{\eta}^y \frac{\mu(z)}{\sigma^2(z)} dz\right).$$

□

**Beispiel 2.34** (Mutations-Selektions-Gleichgewicht). Wir betrachten die Wright-Fisher-Diffusion mit Mutation und Selektion,

$$dX = \left(\frac{1}{2}\theta_a(1-X) - \frac{1}{2}\theta_A X + \alpha X(1-X)\right)dt + \sqrt{X(1-X)}dW,$$

also  $\mu(x) = \frac{1}{2}\theta_a(1-x) - \frac{1}{2}\theta_A x + \alpha x(1-x)$  und  $\sigma^2(x) = x(1-x)$ . Um die stationäre Verteilung zu berechnen, ist also  $[l, r] = [0, 1]$  und damit

$$\begin{aligned} \psi(x) &= \frac{C}{x(1-x)} \exp\left(2 \int_{\eta}^x \frac{\frac{1}{2}\theta_a(1-y) - \frac{1}{2}\theta_A y + \alpha y(1-y)}{y(1-y)} dy\right) \\ &= \frac{C}{x(1-x)} e^{2\alpha x} \exp\left(\int_{\eta}^x \frac{\theta_a}{y} - \frac{\theta_A}{1-y} dy\right) \\ &= C e^{2\alpha x} x^{\theta_a-1} (1-x)^{\theta_A-1}. \end{aligned}$$

Dies können wir nun mit Bemerkung 2.20 vergleichen, wo wir – zumindest im Fall  $\alpha = 0$  – auf dieselbe Verteilung gekommen sind.

Im Folgenden beschäftigen wir uns mit Diffusionen mit Zustandsraum  $[0, 1]$  (d.h. insbesondere, dass  $\mu(0) \geq 0, \mu(1) \leq 0, \sigma^2(0) = \sigma^2(1) = 0$ ) und setzen

$$T_i := \inf\{t \geq 0 : X_t = i\}, \quad i = 0, 1.$$

Genau genommen beschäftigen wir uns mit den folgenden Fragen:

1. Wie groß ist

$$\mathbb{P}_x[T_1 < T_0]?$$

2. Wie groß ist

$$\mathbb{E}_x[T_0 \wedge T_1]?$$

3. Wie sieht  $X$ , bedingt auf  $\{T_1 < T_0\}$  aus?

**Proposition 2.35** (Treffwahrscheinlichkeiten). *Sei  $X$  eine Diffusion mit Zustandsraum  $[0, 1]$  mit Charakteristiken  $\mu, \sigma^2$  und  $T_0 \wedge T_1 < \infty$ . Dann ist  $S(X)$  für die Skalenfunktion*

$$S(x) = \int_{x_0}^x \exp\left(-2 \int_{\xi}^y \frac{\mu(z)}{\sigma^2(z)} dz\right) dy$$

mit  $\xi, x_0 \in (0, 1)$  ein Martingal. Weiter ist für  $0 < x < 1$

$$\mathbf{P}_x(T_1 < T_0) = \frac{S(x) - S(0)}{S(1) - S(0)}, \quad \mathbf{P}_x(T_0 < T_1) = \frac{S(1) - S(x)}{S(1) - S(0)}.$$

*Beweis.* Wir berechnen

$$GS(x) = S'(x) \left( \mu(x) - \sigma^2(x) \frac{\mu(x)}{\sigma^2} \right) = 0,$$

woraus die erste Behauptung folgt. Weiter gilt für das in  $L^1$  konvergente Martingal  $S(X_{\cdot \wedge T_0 \wedge T_1})$

$$S(x) = \mathbf{E}_x[S(X_{T_0 \wedge T_1})] = S(1) \cdot \mathbf{P}_x(T_1 < T_0) + S(0) \cdot (1 - \mathbf{P}_x(T_1 < T_0)).$$

Auflösen nach  $\mathbf{P}_x(T_1 < T_0)$  bringt die restlichen Ergebnisse.  $\square$

Wir berechnen nun die Fixationswahrscheinlichkeit eines vorteilhaften Typs ( $\alpha > 0$ ) in Abwesenheit von Mutation.

**Proposition 2.36.** *Sei  $X$  Lösung der (2.5) mit  $\theta_a = \theta_A = 0$  und  $\alpha \neq 0$ . Dann gilt*

$$\mathbf{P}_x(T_1 < T_0) = \frac{1 - e^{-2\alpha x}}{1 - e^{-2\alpha}}.$$

*Beweis.* Es ist  $\mu(x)/\sigma^2(x) = \alpha$  und damit ist die Skalenfunktion

$$S(x) = \int_0^x e^{-2\alpha y} dy = \frac{1}{2\alpha} (1 - e^{-2\alpha x}).$$

Daraus berechnen wir

$$\mathbf{P}_x(T_1 < T_0) = \frac{S(x) - S(0)}{S(1) - S(0)} = \frac{1 - e^{-2\alpha x}}{1 - e^{-2\alpha}}.$$

$\square$

**Proposition 2.37.** *Sei  $X$  eine Diffusion mit Zustandsraum  $[0, 1]$  mit Charakteristiken  $\mu, \sigma^2$  und  $T := T_0 \wedge T_1 < \infty$ . Dann ist für  $0 < x < 1$*

$$\mathbf{E}_x \left[ \int_0^T g(X_s) ds \right] = \int_0^1 G(x, y) g(y) dy$$

für

$$G(x, y) = \begin{cases} 2 \frac{S(x) - S(0)}{S(1) - S(0)} \cdot (S(1) - S(y)) m(y), & x \leq y \leq 1, \\ 2 \frac{S(1) - S(x)}{S(1) - S(0)} \cdot (S(y) - S(0)) m(y), & 0 \leq y \leq x, \end{cases}$$

$$m(y) = \frac{1}{\sigma^2(y) S'(y)} = \frac{1}{\sigma^2(y)} \exp \left( 2 \int_0^y \frac{\mu(z)}{\sigma^2(z)} dz \right)$$

Insbesondere ist

$$\mathbf{E}_x[T] = \int_0^1 G(x, y) dy.$$

*Beweis.* Wir definieren

$$w(x) := \mathbf{E}_x \left[ \int_0^T g(X_s) ds \right].$$

Zunächst ist

$$\begin{aligned} w(x) &= \mathbf{E}_x \left[ \int_0^h g(X_s) ds \right] + \mathbf{E}_x [w(X_h)], \\ \mathbf{E}_x \left[ \int_0^h g(X_s) ds \right] &= hg(x) + \mathcal{O}(h^2), \\ \mathbf{E}_x [w(X_h)] &= \mathbf{E}_x [w(x) + (X_h - x)w'(x) \\ &\quad + \frac{1}{2}(X_h - x)^2 w''(x) + \mathcal{O}(h^2)] \\ &= w(x) + h(\mu(x)w'(x) + \frac{1}{2}\sigma^2(x)w''(x) + \mathcal{O}(h)), \end{aligned}$$

also für  $h \rightarrow 0$

$$\mu(x)w'(x) + \frac{1}{2}\sigma^2(x)w''(x) = -g(x), \quad w(0) = w(1) = 0.$$

Genauso schreiben wir

$$\begin{aligned} \exp\left(2 \int_\eta^x \frac{\mu(z)}{\sigma^2(z)} dz\right) \frac{2\mu(x)}{\sigma^2(x)} w'(x) \\ + \exp\left(2 \int_\eta^x \frac{\mu(z)}{\sigma^2(z)} dz\right) w''(x) &= -\frac{2g(x)}{\sigma^2(x)} \exp\left(2 \int_\eta^x \frac{\mu(z)}{\sigma^2(z)} dz\right), \end{aligned}$$

und damit

$$\frac{d}{dx} \left( \exp\left(2 \int_\eta^x \frac{\mu(z)}{\sigma^2(z)} dz\right) w'(x) \right) = -\frac{2g(x)}{\sigma^2(x)} \exp\left(2 \int_\eta^x \frac{\mu(z)}{\sigma^2(z)} dz\right)$$

oder auch

$$\frac{d}{dx} \left( \frac{w'(x)}{S'(x)} \right) = -2m(x)g(x).$$

Diese Gleichung integrieren wir nun und erhalten

$$\begin{aligned} \frac{w'(x)}{S'(x)} &= -2 \int_{x_0}^x m(y)g(y) dy + \beta, \\ w(x) &= -2 \int_0^x S'(\eta) \int_0^\eta m(y)g(y) dy d\eta + \beta \int_0^x S'(\eta) d\eta + \alpha \\ &= -2 \int_0^x (S(x) - S(y))m(y)g(y) dy + \beta(S(x) - S(0)) + \alpha. \end{aligned}$$

Da  $w(0) = 0$ , ist auch  $\alpha = 0$ . Wegen  $w(1) = 0$  ist außerdem

$$\beta = \frac{2}{S(1) - S(0)} \int_0^1 (S(1) - S(y))m(y)g(y) dy.$$

Insgesamt ergibt sich

$$w(x) = 2 \int_x^1 \frac{S(1) - S(y)}{S(1) - S(0)} m(y) g(y) dy + \int_0^x \frac{(S(1) - S(y))(S(x) - S(0)) - (S(1) - S(0))(S(x) - S(y))}{S(1) - S(0)} m(y) g(y) dy$$

und damit die Behauptung.  $\square$

**Bemerkung 2.38.** Die Funktion  $G$  heißt auch Green-Funktion. Die Bedeutung liest man am besten ab, indem man  $g(x) = 1_{[x_1, x_2]}$  einsetzt. Nun ist nämlich

$$\int_{x_1}^{x_2} G(x, y) dy = \text{mittlere Zeit, die vor } T \text{ in } [x_1, x_2] \text{ verbracht wird.}$$

**Proposition 2.39.** Sei  $X$  die Lösung von (2.5) mit  $\theta_a = \theta_A = \alpha = 0$ . Dann gilt

$$\mathbf{E}_x[T] = -2x \log x - 2(1-x) \log(1-x).$$

**Bemerkung 2.40.** Man vergleiche dieses Ergebnis mit Bemerkung 2.16.

*Beweis.* Zunächst berechnen wir die Green-Funktion (mit  $S(x) = x$ )

$$G(x, y) = \begin{cases} 2\frac{x}{y}, & x \leq y \leq 1, \\ 2\frac{1-x}{1-y}, & 0 \leq y \leq x. \end{cases}$$

Daraus folgt sofort

$$\mathbf{E}_x[T] = 2 \left( \int_0^x \frac{1-x}{1-y} dy + \int_x^1 \frac{x}{y} dy \right) = -2(1-x) \log(1-x) - 2x \log x.$$

$\square$

**Proposition 2.41.** Sei  $X$  eine Diffusion mit Zustandsraum  $[0, 1]$  mit Charakteristiken  $\mu, \sigma^2$ . Dann ist  $X$ , bedingt unter  $T_1 < T_0$ , ebenfalls eine Diffusion mit Charakteristiken  $\mu_*$  und  $\sigma_*^2$ , wobei

$$\mu_*(x) = \mu(x) + \frac{S'(x)}{S(x) - S(0)} \sigma^2(x), \quad \sigma_*^2(x) = \sigma^2(x).$$

*Beweis.* Wir wissen bereits aus Proposition 2.35 die Wahrscheinlichkeit  $h(x)$ , bei Start in  $x$  den Rand 1 vor 0 zu treffen. Dies verwenden wir nun und schreiben für den Generator  $G$  von  $X$

$$\begin{aligned} \frac{1}{t} (\mathbf{E}_x[f(X_t) | T_1 < T_0] - \mathbf{E}_x[f(x) | T_1 < T_0]) &= \frac{1}{t} \left( \frac{\mathbf{E}_x[f(X_t) \mathbf{E}_{X_t}[T_1 < T_0]]}{h(x)} - f(x) \right) \\ &= \frac{1}{h(x)} \frac{1}{t} (\mathbf{E}_x[f(X_t) h(X_t)] - f(x) h(x)) \xrightarrow{t \rightarrow 0} \frac{Gf h(x)}{h(x)} \end{aligned}$$

Daraus erhalten wir den Generator der bedingten Diffusion und

$$\begin{aligned} \frac{Gf h(x)}{h(x)} &= \frac{\mu(x)}{h(x)} \frac{d}{dx} f(x) h(x) + \frac{\sigma^2(x)}{2h(x)} \frac{d^2}{dx^2} f(x) h(x) \\ &= \frac{f(x)}{h(x)} Gh(x) + \mu(x) f'(x) + \sigma^2(x) \frac{h'(x)}{h(x)} f'(x) + \frac{1}{2} \sigma^2(x) f''(x). \end{aligned}$$

Mit  $Gh = 0$  folgt damit die Behauptung.  $\square$

**Proposition 2.42.** Sei  $X$  Lösung von (2.5) mit  $\theta_a = \theta_A = \alpha = 0$ . Dann löst  $X$ , bedingt unter  $T_1 < T_0$ , die SDE

$$dX = (1 - X)dt + \sqrt{X(1 - X)}dW.$$

*Beweis.* Wir wenden Proposition 2.41 an und erhalten mit  $S(x) = x$

$$\mu_*(x) = \mu(x) + \frac{1}{x}x(1 - x) = 1 - x, \quad \sigma_*^2(x) = \sigma^2(x).$$

Daraus folgt bereits die Behauptung.  $\square$

**Proposition 2.43.** Sei  $X$  Lösung von (2.5) mit  $\theta_a = \theta_A = 0$  und  $\alpha > 0$ . Dann löst  $X$ , bedingt unter  $T_1 < T_0$ , die SDE

$$dX = \alpha X \coth(\alpha X)(1 - X)dt + \sqrt{X(1 - X)}dW,$$

wobei

$$\coth(x) = \frac{e^x + e^{-x}}{e^x - e^{-x}}.$$

*Beweis.* Wir haben bereits berechnet, dass

$$S(x) = \frac{1 - e^{-2\alpha x}}{2\alpha} \text{ und damit } S'(x) = e^{-2\alpha x}.$$

Wendet man nun Proposition 2.41 an, so erhält man  $\sigma_*^2(x) = x(1 - x)$  und

$$\begin{aligned} \mu_*(x) &= \alpha x(1 - x) + \frac{2\alpha e^{-2\alpha x}}{1 - e^{-2\alpha x}}x(1 - x) = \alpha x(1 - x) \frac{1 - e^{-2\alpha x} + 2e^{-2\alpha x}}{1 - e^{-2\alpha x}} \\ &= \alpha x(1 - x) \frac{e^{\alpha x} + 2e^{-\alpha x}}{e^{\alpha x} - e^{-\alpha x}}. \end{aligned}$$

Daraus folgt bereits die Behauptung.  $\square$

**Bemerkung 2.44** (Entkommen von der 0). Die letzten beiden Propositionen haben eine Gemeinsamkeit in der Nähe von  $x = 0$ . Wir bemerken, dass

$$x \frac{\mu(x)}{\sigma^2(x)} \xrightarrow{x \rightarrow 0} 1$$

in beiden Fällen gilt. (Es ist  $x \coth(x) \xrightarrow{x \rightarrow 0} 1$ .) Dies kann man auch allgemeiner fassen: Erfüllt  $X$  die SDE  $dX = \mu(X)dt + \sigma(X)dW$  mit  $\sigma^2(0) = \sigma^2(1) = 0$  sowie  $\mu(0) \geq 0$  und  $\mu(1) \leq 0$ , so gilt: Ist  $\lim_{x \rightarrow 0} x \frac{\mu(x)}{\sigma^2(x)} \geq 1$ , so ist  $\mathbf{P}_x(T_1 < T_0) = 1$  für alle  $x$ . Wir berechnen nämlich

$$\mathbf{P}_x(T_0 < T_1) = \frac{S(1) - S(x)}{S(1) - S(0)} = \frac{\int_x^1 \exp\left(-2 \int_\xi^y \frac{\mu(z)}{\sigma^2(z)} dz\right) dz}{\int_0^1 \exp\left(-2 \int_\xi^y \frac{\mu(z)}{\sigma^2(z)} dz\right) dz}$$

und

$$\begin{aligned} \int_0^1 \exp\left(-2 \int_\xi^y \frac{\mu(z)}{\sigma^2(z)} dz\right) dz &\geq \int_{0+} \exp\left(2 \int_y^\xi \frac{\mu(z)}{\sigma^2(z)} dz\right) dz \geq \int_{0+} \exp\left(2 \int_y^\xi \frac{1}{y} dz\right) dz \\ &\geq \int_{0+} y^{-2} dy = \infty. \end{aligned}$$

**Proposition 2.45** (Green-Funktion der bedingten Diffusion). *Sei  $X$  eine Diffusion mit Zustandsraum  $[0, 1]$  mit Charakteristiken  $\mu, \sigma^2$ , bedingt auf  $T_1 < T_0$ . Dann ist die Green-Funktion aus Proposition 2.37 gegeben durch*

$$G_*(x, y) = \begin{cases} 2 \frac{S(1) - S(y)}{S(1) - S(0)} \frac{S(y) - S(0)}{\sigma^2(y)S'(y)}, & x \leq y, \\ 2 \frac{(S(1) - S(x))(S(y) - S(0))}{S(1) - S(0)} \frac{S(y) - S(0)}{\sigma^2(y)S'(y)(S(x) - S(0))}, & x \geq y. \end{cases}$$

Insbesondere ist also, für glattes  $g$ ,

$$\mathbf{E}_x \left[ \int_0^T g(X_s) ds \mid T_1 < T_0 \right] = \int_0^1 G_*(x, y) g(y) dy.$$

*Beweis.* Wir müssen die Green-Funktion aus Proposition 2.37 mit  $\sigma_*^2 = \sigma^2$  und  $\mu_*(x) = \mu(x) + \frac{S'(x)}{S(x) - S(0)} \sigma^2(x)$  berechnen. Wir bezeichnen hierfür mit  $S_*$  die Skalenfunktion der bedingten Diffusion, also (wir lassen die unteren Integralgrößen weg, da die Skalenfunktionen nur bis auf additive und multiplikative Konstanten eindeutig ist)

$$\begin{aligned} S_*(x) &= \int^x \exp \left( -2 \int^y \frac{\mu_*(z)}{\sigma_*^2(z)} dz \right) dy = \int^x \exp \left( -2 \int^y \frac{\mu(z)}{\sigma^2(z)} + \frac{S'(z)}{S(z)} dz \right) dy \\ &= \int^x \exp \left( -2 \log S(y) - 2 \int^y \frac{\mu(z)}{\sigma^2(z)} dz \right) dy = \int^x (S(y))^{-2} S'(y) dy \\ &= -(S(y))^{-1} \Big|_{y=x} = -(S(x))^{-1}. \end{aligned}$$

Setzen wir dies nun in die Formel für die Green-Funktion ein, so erhalten wir für  $S(0) \rightarrow 0$

$$\begin{aligned} G_*(x, y) &= \begin{cases} 2 \frac{S_*(x) - S_*(0)}{S_*(1) - S_*(0)} \cdot \frac{S_*(1) - S_*(y)}{\sigma^2(y)S'_*(y)} = 2 \frac{S(x) - S(0)}{S(1) - S(0)} \cdot \frac{S(1) - S(y)}{\sigma^2(y)S'(y)} \frac{S(1)S(0)S(y)^2}{S(x)S(0)S(1)S(y)}, & x \leq y \leq 1, \\ 2 \frac{S_*(1) - S_*(x)}{S_*(1) - S_*(0)} \cdot \frac{S_*(y) - S_*(0)}{\sigma^2(y)S'_*(y)} = 2 \frac{S(1) - S(x)}{S(1) - S(0)} \cdot \frac{S(y) - S(0)}{\sigma^2(y)S'(y)} \frac{S(1)S(0)S(y)^2}{S(1)S(x)S(y)S(0)}, & 0 \leq y \leq x, \end{cases} \\ &= \begin{cases} 2 \frac{S(y) - S(0)}{S(1) - S(0)} \cdot \frac{S(1) - S(y)}{\sigma^2(y)S'(y)}, & x \leq y \leq 1, \\ 2 \frac{S(1) - S(x)}{S(1) - S(0)} \cdot \frac{(S(y) - S(0))^2}{\sigma^2(y)S'(y)S(x) - S(0)}, & 0 \leq y \leq x \end{cases} \end{aligned}$$

und die Behauptung folgt.  $\square$

**Proposition 2.46.** *Sei  $X$  eine Diffusion mit Zustandsraum  $[0, 1]$  und Green-Funktion  $G$ . Dann ist für  $T = T_0 \wedge T_1$ , bedingt unter  $T_1 < T_0$*

$$\mathbf{V}_{0+}^* \left[ \int_0^T g(X_s) ds \right] = 2 \int_0^1 \int_0^y G_*(0, y) G_*(y, z) g(y) g(z) dz dy.$$

*Beweis.* Wir berechnen zunächst

$$\begin{aligned} \mathbf{E}_{0+}^* \left[ \int_0^T \int_s^T g(X_t) \cdot g(X_s) dt ds \right] &= \mathbf{E}_{0+}^* \left[ \int_0^T g(X_s) \mathbf{E}_{X_s}^* \left[ \int_0^T g(X_t) dt \right] ds \right] \\ &= \int_0^1 G_*(0, y) g(y) \mathbf{E}_y^* \left[ \int_0^T g(X_t) dt \right] dy \\ &= \int_0^1 \int_0^1 G_*(0, y) G_*(y, z) g(y) g(z) dz dy. \end{aligned}$$

Damit folgt

$$\begin{aligned}
\mathbf{V}_{0+}^* \left[ \int_0^T g(X_s) ds \right] &= \mathbf{E}_{0+}^* \left[ \int_0^T \int_0^T g(X_t) g(X_s) dt ds \right] - \left( \int_0^1 G_*(0, y) g(y) dy \right)^2 \\
&= 2\mathbf{E}_{0+}^* \left[ \int_0^T \int_s^T g(X_t) g(X_s) dt ds \right] - 2 \int_0^1 \int_y^1 G_*(0, z) G_*(0, y) g(y) g(z) dz dy \Big)^2 \\
&= 2 \int_0^1 \int_0^1 G_*(0, y) G_*(y, z) g(y) g(z) dz dy - 2 \int_0^1 \int_y^1 G_*(0, z) G_*(0, y) g(y) g(z) dz dy \Big)^2 \\
&= 2 \int_0^1 \int_0^y G_*(0, y) G_*(y, z) g(y) g(z) dz dy,
\end{aligned}$$

da  $G_*(x, z) = G_*(y, z)$  für  $x, y \leq z$ . □

Wir wenden die letzten beiden Propositionen nun an, um die Fixationszeit eines stark vorteilhaften Allels zu untersuchen.

**Proposition 2.47.** *Sei  $X$  die Diffusion mit  $\mu(x) = \alpha x(1-x)$  und  $\sigma^2(x) = x(1-x)$ . Weiter sei  $\mathbf{P}_{0+}^*[\cdot] = \lim_{x \rightarrow 0} \mathbf{P}_x[\cdot | T_1 < T_0]$  die Verteilung von  $X$  gestartet in  $0+$ , bedingt unter  $T_1 < T_0$ . Dann gilt*

$$\frac{\alpha}{\log \alpha} T_1 \xrightarrow{\alpha \rightarrow \infty} L^2(\mathbf{P}_{0+}^*) \quad 2.$$

*Beweis.* Wir müssen zeigen, dass

$$\mathbf{E}_{0+}^* \left[ \left( \frac{\alpha}{2 \log \alpha} T_1 - 1 \right)^2 \right] = \mathbf{V}_{0+}^* \left[ \frac{\alpha}{2 \log \alpha} T_1 \right] + \left( \mathbf{E}_{0+}^* \left[ \frac{\alpha}{2 \log \alpha} T_1 \right] - 1 \right)^2 \xrightarrow{\alpha \rightarrow \infty} 0. \quad (2.6)$$

Zunächst berechnen wir direkt, aus Proposition 2.45

$$G_*(x, y) = \begin{cases} \frac{(1 - e^{-2\alpha(1-y)})(1 - e^{-2\alpha y})}{\alpha y(1-y)(1 - e^{-2\alpha})}, & x \leq y \\ \frac{(e^{-2\alpha x} - e^{-2\alpha})(e^{2\alpha y} - 1)(1 - e^{-2\alpha y})}{\alpha y(1-y)(1 - e^{-2\alpha})(1 - e^{-2\alpha x})}, & x > y. \end{cases} \quad (2.7)$$

Wir beachten, dass  $G_*(x, y)$  fallend in  $x$  ist. Nun schreiben wir, wegen  $\frac{1}{y(1-y)} = \frac{1}{y} + \frac{1}{1-y}$ , für ein beliebiges  $0 < \varepsilon < 1$

$$\begin{aligned}
\left| \mathbf{E}_{0+}^*[T_1] - \frac{2 \log \alpha}{\alpha} \right| &= \frac{1}{\alpha} \left| \int_0^1 \frac{(1 - e^{-2\alpha(1-y)})(1 - e^{-2\alpha y})}{y(1-y)} dy - 2 \log \alpha \right| + o(1) \\
&\leq \frac{2}{\alpha} \left| \int_0^1 \frac{1 - e^{-2\alpha y}}{y} dy - \log \alpha \right| + o(1) \\
&\stackrel{y \rightarrow 2\alpha y}{=} \frac{2}{\alpha} \left| \int_0^{2\alpha} \frac{1 - e^{-y}}{y} dy - \log \alpha \right| + o(1) \\
&= \frac{2}{\alpha} \left| \int_{\alpha^\varepsilon}^{2\alpha} \frac{1}{y} dy - \log \alpha \right| + o(1) = \frac{2}{\alpha} \left| \log \alpha - \log \alpha^\varepsilon - \log \alpha \right| + o(1) \\
&\leq \varepsilon + o(1).
\end{aligned}$$

Daraus folgt die Konvergenz des zweiten Terms in (2.6). Für den ersten Term berechnen wir mittels Proposition 2.46, und für Konstanten  $0 < C < \infty$ ,

$$\begin{aligned}
\mathbf{V}_{0+}^*[T] &= 2 \int_0^1 \int_0^y G_*(0, y) G_*(y, z) dz dy \\
&= \frac{2}{\alpha^2(1 - e^{-2\alpha})} \int_0^1 \int_0^y \frac{(1 - e^{-2\alpha(1-y)})(1 - e^{-2\alpha y})}{y(1-y)} \\
&\quad \frac{(e^{-2\alpha(y-z)} - e^{-2\alpha(1-z)})(1 - e^{-2\alpha z})^2}{z(1-z)(1 - e^{-2\alpha y})} dz dy \\
&= \frac{2}{\alpha^2(1 - e^{-2\alpha})^2} \int_0^1 \int_0^y \frac{(1 - e^{-2\alpha(1-y)})(e^{-2\alpha y} - e^{-2\alpha})}{y(1-y)} \frac{(e^{2\alpha z} - 1)(1 - e^{-2\alpha z})}{z(1-z)} dz dy \\
&\stackrel{y \rightarrow 1-y}{\leq} \frac{2e^{-2\alpha}}{\alpha^2(1 - e^{-2\alpha})^2} \int_0^1 \int_0^{1-y} \frac{e^{2\alpha y} - 1}{y(1-y)} \frac{e^{2\alpha z} - 1}{z(1-z)} dz dy \\
&= \frac{2}{\alpha^2(1 - e^{-2\alpha})^2} \left( \left( e^{-\alpha} \int_0^{1/2} \frac{e^{2\alpha y} - 1}{y(1-y)} dy \right)^2 + 2 \int_0^{1/2} \int_0^y \frac{e^{-2\alpha y} - e^{-2\alpha}}{y(1-y)} \frac{e^{2\alpha z} - 1}{z(1-z)} dz dy \right),
\end{aligned}$$

wobei wir die Fläche  $\{(y, z) : z \leq 1 - y\}$  in  $\{(y, z) : y, z \leq 1/2\}$ ,  $\{(y, z) : y \leq 1/2, 1/2 \leq z \leq 1 - y\}$  und  $\{(y, z) : y \geq 1/2, z \leq 1 - y\}$  aufgeteilt haben und die Symmetrie des Integranden ausgenutzt haben. Hieraus sehen wir

$$\mathbf{V}_{0+}^*[T] \leq \frac{2}{\alpha^2} \left( C + 8 \int_0^{1/2} \int_0^y \frac{e^{-2\alpha y} - e^{-2\alpha}}{y} \frac{e^{2\alpha z} - 1}{z} dz dy \right)$$

und die Konvergenz der Varianz folgt, da

$$\begin{aligned}
&\int_0^{1/2} \int_0^y \frac{e^{-2\alpha y} - e^{-2\alpha}}{y} \frac{e^{2\alpha z} - 1}{z} dz dy \stackrel{y \rightarrow 2\alpha y, z \rightarrow 2\alpha z}{=} \int_0^\alpha \frac{e^{-y} - e^{-2\alpha}}{y} \int_0^y \frac{e^z - 1}{z} dz dy \\
&= \int_1^\alpha \frac{e^{-y} - e^{-2\alpha}}{y} \int_1^y \frac{e^z - 1}{z} dz dy + C' \leq C \int_1^{2\alpha} \frac{1}{y^2} dy + C' = \mathcal{O}(1),
\end{aligned}$$

für Konstanten  $0 < C, C' < \infty$ , wobei wir das Integral  $\int_1^y dy$  in  $\int_1^{y/2} dy + \int_{y/2}^y dy$  aufgeteilt haben, um '≤' einzusehen.  $\square$

### 3 Neurowissenschaft

Wir stelle nun Modelle vor, die das Verhalten von Neuronen im Gehirn abbilden. Zunächst müssen wir hierfür über die zugrundeliegende liegende Biologie reden. Das Nervensystem besteht aus elektrisch anregbaren Zellen, die man Neuronen nennt. Diese verarbeiten Informationen (wobei nicht klar ist, wie sie das genau tun). Jedes Neuron erhält elektrischen Input von benachbarten (oder verbundenen) Nervenzellen über Dendriten. Dieser Input wird im Kern des Neurons gebündelt und führt zu einer Spannungsdifferenz über die Membran der Zelle hinweg, das Membranpotential. Erreicht das Membranpotential einen gewissen Wert, so entlädt es sich entlang des Axons, von wo aus es zu anderen Neuronen gelangt. Das zentrale Objekt, das wir hier modellieren wollen, ist genau dieses Membranpotential, sowie die Zeitpunkte, zu denen es sich entlädt.

### 3.1 Modelle eines einzelnen Neurons

Das Input des Membranpotentials kann elektrischer oder chemischer Natur (via Ionen) sein. Wir gehen davon aus, dass ein Neuron, das wir betrachten, über Dendriten zu verschiedenen Neuronen Verbindungen hat, die das Membranpotential entweder erhöhen oder erniedrigen können. Ein entsprechendes Modell für ein einzelnes Neuron wäre das folgende.

**Definition 3.1** (Stein's Modell für synaptischen Input). *Das Membranpotential  $V$  ist ein Markov-Prozess mit Generator*

$$Gf(v) = -\frac{1}{\tau}f'(v) + \lambda_+(f(v + a_+) - f(v)) + \lambda_-(f(v - a_-) - f(v))$$

für glatte  $f$ . Mit anderen Worten ist  $V$  stückweise deterministisch, und zwar zwischen den Sprungpunkten von Poisson-Prozessen mit Rate  $\lambda_+$  und  $\lambda_-$ . An solchen Poisson-Punkten springt  $V$  um  $a_+$  nach oben oder um  $a_-$  nach unten.

Im Grenzwert vieler Inputs berechnen wir nun eine Approximation.

**Proposition 3.2** (Ornstein-Uhlenbeck-Grenzwert des Stein'schen Modells). *Sei*

$$\lambda_+ = N\kappa_+, \quad \lambda_- = N\kappa_-, \quad a_+ = b_+/\sqrt{N}, \quad a_- = b_-/\sqrt{N},$$

so dass

$$\sqrt{N}(\kappa_+b_+ - \kappa_-b_-) \xrightarrow{N \rightarrow \infty} \mu.$$

Weiter setzen wir

$$\sigma^2 := \kappa_+b_+^2 + \kappa_-b_-^2.$$

Sei  $V = V^N$  das Membranpotential im Stein'schen Modell, so gilt  $V^N \xrightarrow{N \rightarrow \infty}_{fdd} X$ , wobei

$$dX = \left( \mu - \frac{1}{\tau}X \right) dt + \sqrt{\sigma^2} dW.$$

*Beweis.* Wir schreiben direkt mit dem Generator  $G^N$  für  $V^N$

$$\begin{aligned} G^N f(v) &= -\frac{1}{\tau}f'(v) + N\kappa_+(f(v + b_+/\sqrt{N}) - f(v)) + N\kappa_-(f(v - b_-/\sqrt{N}) - f(v)) \\ &= -\frac{1}{\tau}f'(v) + N\kappa_+ \left( f'(v) \frac{b_+}{\sqrt{N}} + \frac{1}{2}f''(v) \frac{b_+^2}{N} \right) + N\kappa_- \left( -f'(v) \frac{b_-}{\sqrt{N}} + \frac{1}{2}f''(v) \frac{b_-^2}{N} \right) + o(1) \\ &\xrightarrow{N \rightarrow \infty} -\frac{1}{\tau}f'(v) + \mu f'(v) + \frac{1}{2}\sigma^2 f''(v). \end{aligned}$$

Daraus folgt die Behauptung. □

### Der Ornstein-Uhlenbeck-Prozess

Da der Ornstein-Uhlenbeck-Prozess nun häufiger auftauchen wird, studieren wir ihn zunächst etwas genauer.

**Proposition 3.3** (Ein-dimensionaler Ornstein-Uhlenbeck-Prozess). *Sei  $\mu \in \mathbb{R}$  und  $\sigma^2 > 0$ . Die Lösung der SDE*

$$dX = -\mu(x_0 - X)dt + \sigma dW$$

heißt ein-dimensionaler Ornstein-Uhlenbeck-Prozess mit Zentrum  $x_0$ .

**Bemerkung 3.4.** Ist  $X$  ein Ornstein-Uhlenbeck-Prozess mit Zentrum  $x_0$ , so ist  $Y = X - x_0$  ein Ornstein-Uhlenbeck-Prozess mit Zentrum 0, also Lösung von

$$dY = -\mu Y dt + \sigma dW. \quad (3.1)$$

**Lemma 3.5** (Der Ornstein-Uhlenbeck-Prozess ist normalverteilt). *Sei  $Y$  die Lösung von (3.1) mit  $Y_0 = y_0$ . Dann ist  $Y_t \sim N\left(e^{-\mu t} y_0, \frac{\sigma^2}{\mu}(1 - e^{-2\mu t})\right)$ . Insbesondere gilt  $Y_t \xrightarrow{t \rightarrow \infty} Y_\infty \sim N\left(0, \frac{\sigma^2}{\mu}\right)$ , falls  $\mu > 0$ .*

**Bemerkung 3.6.** Mit Hilfe der stochastischen Integration kann man auch einsehen, dass der Prozess

$$Y_t = e^{-\mu t} \left( y_0 + \sigma \int_0^t e^{\mu s} dW_s \right)$$

die SDE (3.1) löst. Insbesondere ist für diesen Prozess

$$\begin{aligned} \mathbf{E}[Y_t] &= e^{-\mu t} y_0, \\ \mathbf{V}[Y_t] &= \mathbf{E}\left[\left(e^{-\mu t} \sigma \int_0^t e^{\mu s} dW_s\right)^2\right] = \sigma^2 e^{-2\mu t} \int_0^t e^{2\mu s} ds = \frac{\sigma^2}{2\mu}(1 - e^{-2\mu t}). \end{aligned}$$

*Beweis.* Für  $\lambda \in \mathbb{R}$  ist

$$\left( e^{\lambda Y_t} - \int_0^t -\mu \lambda Y_s e^{\lambda Y_s} + \frac{1}{2} \sigma^2 \lambda^2 e^{\lambda Y_s} ds \right)_{t \geq 0}$$

ein Martingal. Insbesondere ist

$$\frac{\partial}{\partial t} \mathbf{E}[e^{\lambda Y_t}] = \mathbf{E}\left[-\mu \lambda Y_t e^{\lambda Y_t} + \frac{1}{2} \sigma^2 \lambda^2 e^{\lambda Y_t}\right] = -\mu \lambda \frac{\partial}{\partial \lambda} \mathbf{E}[e^{\lambda Y_t}] + \frac{1}{2} \sigma^2 \lambda^2 \mathbf{E}[e^{\lambda Y_t}].$$

Wählt man nun  $\lambda(t) = \lambda_0 e^{\mu t}$  und setzt  $h(t, \lambda) := \mathbf{E}[e^{\lambda Y_t}]$ , bedeutet das

$$\frac{d}{dt} h(t, \lambda(t)) = \frac{\partial}{\partial t} h(t, \lambda(t)) + \mu \lambda(t) \frac{\partial}{\partial \lambda} h(t, \lambda(t)) = \frac{1}{2} \sigma^2 \lambda^2(t) h(t, \lambda(t)).$$

Diese Differentialgleichung lässt sich lösen, indem man Trennung der Variablen verwendet und

$$\log(h(t, \lambda(t))) - \log(h(0, \lambda_0)) = \int_{h(0, \lambda_0)}^{h(t, \lambda(t))} \frac{dh}{h} = \int_0^t \sigma^2 \lambda_0^2 e^{2\mu s} ds = \frac{\sigma^2}{4\mu} \lambda_0^2 (e^{2\mu t} - 1)$$

schreibt. Wählt man  $\lambda_0 = \lambda e^{-\mu t}$ , folgt

$$\log \mathbf{E}[e^{\lambda Y_t}] = \log h(t, \lambda) = \lambda e^{-\mu t} y_0 + \frac{\sigma^2}{4\mu} \lambda^2 (1 - e^{-2\mu t}).$$

Mit anderen Worten<sup>7</sup> ist  $Y_t \sim N\left(e^{-\mu t} y_0, \frac{\sigma^2}{2\mu} (1 - e^{-2\mu t})\right)$ .  $\square$

**Proposition 3.7** (Treffzeiten des Ornstein-Uhlenbeck-Prozesses). *Sei  $\mu > 0$  und  $Y$  die Lösung von (3.1), sowie  $T_y := \inf\{t \geq 0 : Y_t = y\}$  die Treffzeit von  $y$ . Dann gilt für  $\lambda y > 0$*

$$\mathbf{E}_{y_0} \left[ \exp \left( -\lambda y e^{\mu T_y} - \lambda^2 \frac{\sigma^2}{4\mu} e^{2\mu T_y} \right) \right] = e^{\lambda y_0 - \lambda^2 \frac{\sigma^2}{4\mu}}.$$

*Beweis.* Wir suchen  $t \mapsto \alpha(t), \beta(t)$ , so dass  $M = (M_t)_{t \geq 0}$  mit

$$M_t = \exp(\alpha(t)X_t + \beta(t))$$

ein Martingal ist. Wir berechnen

$$\mathbf{E}[\exp(\alpha(t)X_t + \beta(t)) | \mathcal{F}_s] = \exp(\alpha(s)X_s + \beta(s)) \mathbf{E}_{X_s}[\exp(\alpha(t)X_t - \alpha(s)X_s + \beta(t) - \beta(s))],$$

also sind  $\alpha$  und  $\beta$  so zu bestimmen, dass

$$\mathbf{E}_x[e^{\alpha(t)X_{t-s}}] = \exp\left(\alpha(t)e^{-\mu(t-s)}x + \frac{\sigma^2}{4\mu}\alpha(t)^2(1 - e^{-2\mu(t-s)})\right) = e^{\alpha(s)x - (\beta(t) - \beta(s))}.$$

Hierzu wählen wir

$$\alpha(t) = \lambda e^{\mu t}, \quad \beta(t) = -\frac{\sigma^2}{4\mu} \lambda^2 (e^{2\mu t} - 1).$$

Da  $(M_{t \wedge T_y})_{t \geq 0}$  ein Martingal ist, folgt

$$\begin{aligned} e^{\lambda y_0} &= \mathbf{E} \left[ \exp \left( \lambda e^{\mu T_y \wedge t} X_{T_y \wedge t} - \frac{\sigma^2}{4\mu} \lambda^2 (e^{2\mu T_y \wedge t} - 1) \right) \right] \\ &\xrightarrow{t \rightarrow \infty} \mathbf{E} \left[ \exp \left( \lambda e^{\mu T_y} y - \frac{\sigma^2}{4\mu} \lambda^2 (e^{2\mu T_y} - 1) \right) \right]. \end{aligned}$$

Daraus folgt die Behauptung.  $\square$

Wie beschrieben entwickelt sich das Membranpotential, bis es eine gewisse Grenze übersteigt, und entlädt sich anschließend entlang seines Axons. Das Membranpotential selbst wird dann auf einen Anfangswert zurückgesetzt. Da wir das Membranpotential durch einen Ornstein-Uhlenbeck-Prozess beschrieben haben, wäre also die Verteilung der Treffzeit aus Proposition 3.7, also die Zeit bis zur nächsten Entladung des Potentials, interessant. Die ganze Verteilung können wir allerdings nicht berechnen, sondern nur – mit Hilfe der Theorie eindimensionaler Diffusionen – Momente der Verteilung (siehe Übung). Wir wollen nun jedoch zumindest in einem einfacheren Fall die Verteilungen der Treffzeit für einen handhabbaren Fall durchrechnen.

<sup>7</sup>Für  $Z \sim N(\mu, \sigma^2)$  ist bekanntlich  $\mathbf{E}[e^{\lambda Z}] = e^{t\mu + \frac{1}{2}t^2\sigma^2}$ .

### Treffzeiten der Brown'schen Bewegung

**Proposition 3.8.** Sei  $X$  eine Brown'sche Bewegung mit  $X_0 = 0$ . Dann ist

$$f(t) = \frac{x}{\sqrt{2\pi t^3}} \exp\left(-\frac{x^2}{2t}\right) 1_{[0,\infty)}(t)$$

die Dichte von  $T_x := \inf\{t \geq 0 : X_t = x\}$ .

**Bemerkung 3.9.** Im Beweis werden wir die Cauchy-Schlömilch-Transformation verwenden: Für eine messbare, nicht-negative Funktion  $f$  und  $a, b > 0$  gilt

$$\int_0^\infty f((at - b/t)^2) dt = \frac{1}{a} \int_0^\infty f(y^2) dy.$$

In der Tat: Es gilt

$$\begin{aligned} \int_0^\infty f((at - b/t)^2) dt &\stackrel{x=b/(at)}{=} \frac{b}{a} \int_0^\infty \frac{1}{x^2} f((b/x - ax)^2) dx = \frac{b}{a} \int_0^\infty \frac{1}{t^2} f((at - b/t)^2) dt \\ &= \frac{1}{2} \int_0^\infty \left(1 + \frac{b}{a} \frac{1}{t^2}\right) f((at - b/t)^2) dt \\ &\stackrel{y=at-b/t}{=} \frac{1}{2a} \int_{-\infty}^\infty f(y^2) dy = \frac{1}{a} \int_0^\infty f(y^2) dy \end{aligned}$$

*Beweis.* Wir gehen folgendermaßen vor: Zunächst ist  $M = (\exp(\sqrt{2\lambda}X_t - \lambda t))_{t \geq 0}$  ein Martingal. Hierfür können wir das Optional Sampling Theorem verwenden, um  $\mathbf{E}[e^{-\lambda T_x}] = e^{-\sqrt{2\lambda}x}$  zu berechnen. Am Ende zeigen wir, dass die angegebene Dichte dieselbe Laplace-Transformierte besitzt.

Das Optional Sampling Theorem zeigt, dass  $M_{t \wedge T_x}$  ein in  $L^1$  konvergentes Martingal ist. Daraus folgt

$$1 = \mathbf{E}[\exp(\sqrt{2\lambda}X_{t \wedge T_x} - \lambda t \wedge T_x)] \xrightarrow{t \rightarrow \infty} \mathbf{E}[\exp(\sqrt{2\lambda}x - \lambda T_x)].$$

Für die Laplace-Transformierte berechnen wir, mit ein paar Tricks,

$$\begin{aligned} &\frac{x}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{t^3}} \exp\left(-\lambda t - \frac{x^2}{2t}\right) dt \\ &= -x \sqrt{\frac{2}{\pi}} \frac{d}{da} \int_0^\infty \frac{1}{\sqrt{t}} \exp\left(-\lambda t - \frac{a}{2t}\right) dt \Big|_{a=x^2} \\ &\stackrel{y=\sqrt{t}}{=} -2x \sqrt{\frac{2}{\pi}} \frac{d}{da} \int_0^\infty \exp\left(-\lambda y^2 - \frac{a}{2y^2}\right) dy \Big|_{a=x^2} \\ &= -2x \sqrt{\frac{2}{\pi}} \frac{d}{da} \exp(-\sqrt{2a\lambda}) \int_0^\infty \exp\left(-\left(\sqrt{\lambda}y - \sqrt{\frac{a}{2}} \frac{1}{y}\right)^2\right) dy \Big|_{a=x^2} \quad (3.2) \\ &= -2x \sqrt{\frac{2}{\pi\lambda}} \frac{d}{da} \exp(-\sqrt{2a\lambda}) \int_0^\infty e^{-y^2} dy \Big|_{a=x^2} \\ &= -x \sqrt{\frac{2}{\lambda}} \frac{d}{da} \exp(-\sqrt{2a\lambda}) \Big|_{a=x^2} \\ &= x \sqrt{\frac{2}{\lambda}} \frac{2\lambda}{2\sqrt{2a\lambda}} e^{-\sqrt{2a\lambda}} \Big|_{a=x^2} = e^{-\sqrt{2\lambda}x}, \end{aligned}$$

und die Aussage ist gezeigt.  $\square$

Ein Neuron, das durch ein Membranpotential  $V$  modelliert wird, und beim Treffen von  $v_{th}$  zur Zeit  $T_{th}$  nach  $v_r$  springt, heißt auch *Leaky integrate-and-fire neuron*. Nehmen wir

$$dV = (I - \theta V)dt + \sigma dW,$$

so ist es *leaky* wegen des Terms  $-\theta V$ , es integriert dabei den Input auf  $I dt + \sigma dW$  auf, und wegen des Sprungs im Membranpotential feuert es.

### 3.2 Modelle mehrerer gekoppelter Neuronen

Im Gehirn sind viele Neuronen gekoppelt. Dabei feuern sie in abhängiger Art und Weise, wobei die Abhängigkeiten einfach dadurch erklärt werden, dass der Output (Feuern) eines Neurons den Input für verbundene Neuronen darstellt. Um Netzwerke aus  $N$  solcher Neuronen zu betrachten, eignet sich die Gleichung

$$dV^i = g(V_t^i)dt + \frac{\alpha}{N} \sum_{j=1}^N \sum_{k=1}^{\infty} \delta_{T_k^j}(t) + \sigma dW^i, \quad i = 1, \dots, N. \quad (3.3)$$

Hier ist  $g$  eine Lipschitz-stetige Funktion,  $\alpha$  ein Kopplungsparameter und  $T_1^i < T_2^i < T_3^i < \dots$  die Folge von Treffzeiten (oder Überschreitungszeiten) von  $v_{th}$  des Prozesses  $V^i$ . Zu diesen Zeiten springt  $V^i$  nach  $v_r$ . Zu beachten ist hier, dass alle Interaktionen zwischen Neuronen identisch modelliert werden.

Will man in (3.3) den Grenzwert  $N \rightarrow \infty$  betrachten, so stellt man zunächst fest, dass die Gleichung von der Form

$$dV^i = g(V_t^i)dt + \frac{1}{N} \sum_{j=1}^N f(V_t^j)dt + \sigma dW^i \quad (3.4)$$

für eine geeignete Funktion  $f$  ist. Da  $V^1, \dots, V^N$  beinahe unabhängig sind, wird für  $N \rightarrow \infty$  der Grenzwert

$$dV = \left( g(V_t) + \int f(x) \mu_t(dx) \right) dt + \sigma dW \quad (3.5)$$

lauten, wobei  $V_t \sim \mu_t$ . Interessanterweise hängt hier also ein Parameter der stochastischen Differentialgleichung (die Drift) nicht vom Zustand des Prozesses  $V_t$  ab, sondern von der ganzen Verteilung von  $V_t$ . Dies beleuchten wir zunächst etwas genauer.

**Proposition 3.10.** *Sei  $f$  Lipschitz-stetig und  $X_0 = x_0$ . Dann hat*

$$dX = \left( g(X_t) + \int f(x) \mu_t(dx) \right) dt + \sigma dW \quad (3.6)$$

eine eindeutige Lösung mit  $X_t \sim \mu_t$ .

*Beweis.* Für die Eindeutigkeit seien  $X$  und  $Y$  Lösungen von (3.6) mit  $X_0 = Y_0$ . Dann ist

$$\mathbf{E}[|X_t - Y_t|] = \int_0^t \mathbf{E}[|g(X_s) - g(Y_s)|] + |\mathbf{E}[f(X_s) - f(Y_s)]| ds \leq L \int_0^t \mathbf{E}[|X_s - Y_s|] ds$$

für ein geeignetes  $L$ . Mit dem Gronwall-Lemma<sup>8</sup> folgt nun  $X = Y$ .

Wir zeigen die Existenz nur für  $g = 0$ , für allgemeine  $g$  folgt sie analog. Sei  $L$  die Lipschitz-Konstante von  $f$ . Wir verwenden eine Picard-Iteration und starten mit

$$X_t^0 = x_0 + \sigma W_t,$$

und iterativ

$$X_t^{n+1} = x_0 + \int_0^t \mathbf{E}[f(X_s^n)] ds + \sigma W_t, \quad n = 0, 1, 2, \dots$$

Damit ist

$$X_t^{n+1} - X_t^n = \int_0^t \mathbf{E}[f(X_s^n) - f(X_s^{n-1})] ds,$$

also

$$\begin{aligned} \mathbf{E}[\sup_{s \leq t} |X_t^{n+1} - X_t^n|] &\leq \int_0^t \mathbf{E}[|f(X_s^n) - f(X_s^{n-1})|] ds \leq L \int_0^t \mathbf{E}[|X_s^n - X_s^{n-1}|] ds \\ &\leq L^2 \int_0^t \int_0^s \mathbf{E}[|X_r^{n-1} - X_r^{n-2}|] dr ds = L^2 \int_0^t \int_r^t \mathbf{E}[|X_r^{n-1} - X_r^{n-2}|] ds dr \\ &= L^2 \int_0^t (t-s) \mathbf{E}[|X_s^{n-1} - X_s^{n-2}|] ds \\ &\leq L^3 \int_0^t (t-s) \int_0^s \mathbf{E}[|X_r^{n-2} - X_r^{n-3}|] dr ds \\ &= L^3 \int_0^t \mathbf{E}[|X_r^{n-2} - X_r^{n-3}|] \int_r^t (t-s) ds dr \\ &= \frac{L^3}{2} \int_0^t (t-s)^2 \mathbf{E}[|X_s^{n-2} - X_s^{n-3}|] ds \\ &\leq \dots \\ &\leq \frac{L^n}{(n-1)!} \int_0^t (t-s)^{n-1} \mathbf{E}[|X_s^1 - X_s^0|] ds \end{aligned}$$

und damit, für ein geeignetes  $C > 0$

$$\mathbf{E} \left[ \sum_{n=0}^{\infty} \sup_{s \leq t} |X_t^{n+1} - X_t^n| \right] \leq C e^{L(t-s)}.$$

Wegen der Vollständigkeit von  $L^1$  gibt es also ein  $X_t$  mit  $X^n \xrightarrow{n \rightarrow \infty} X$  uniform auf Kompakta. Die Konvergenz liefert außerdem

$$\begin{aligned} X_t &= \lim_{n \rightarrow \infty} X_t^{n+1} = x_0 + \int_0^t \mathbf{E}[f(X_s^n)] ds + \sigma W_t \\ &= x_0 + \int_0^t \mathbf{E}[f(X_s)] ds + \sigma W_t, \end{aligned}$$

also löst  $X$  die Gleichung (3.6). □

<sup>8</sup>Das Gronwall-Lemma besagt: Sei  $f$  so, dass  $f(t) \leq a(t) + \int_0^t b(s)f(s) ds$  für  $a, b \geq 0$ . Dann gilt  $f(t) \leq a(t) + \int_0^t a(s)b(s)e^{\int_s^t b(r) dr} ds$ .

**Proposition 3.11.** Seien  $f, g$  Lipschitz-stetig,  $X^N = (X^{1,N}, \dots, X^{N,N})$  Lösung von

$$dX^{i,N} = \left( g(X_t^{i,N}) + \frac{1}{N} \sum_{j=1}^N f(X_t^{j,N}) \right) dt + \sigma dW^i, \quad i = 1, \dots, N \quad (3.7)$$

und  $X = (X^1, X^2, \dots)$  so, dass  $X^1, X^2, \dots$  unabhängig sind und

$$dX^i = \left( g(X_t^i) + \int f(x) \mu_t(dx) \right) dt + \sigma dW^i, \quad i = 1, 2, \dots \quad (3.8)$$

mit  $X_0^{i,N} = X_0^i$ ,  $i = 1, \dots, N$ . Dann gilt  $X_t^{i,N} \xrightarrow{N \rightarrow \infty}_{f_s} X_t^i$ .

*Beweis.* Es gilt für ein  $C > 0$

$$\begin{aligned} \max_{1 \leq i \leq N} |X_t^{i,N} - X_t^i| &= \max_{1 \leq i \leq N} \left| \int_0^t g(X_s^{i,N}) - g(X_s^i) + \frac{1}{N} \sum_{j=1}^N f(X_s^{j,N}) - \int f(x) \mu_s(dx) ds \right| \\ &\leq C \left( \int_0^t \max_{1 \leq i \leq N} |X_s^{i,N} - X_s^i| + \frac{1}{N} \sum_{j=1}^N |X_s^{j,N} - X_s^j| \right) + \left| \frac{1}{N} \sum_{j=1}^N f(X_s^j) - \mathbf{E}[f(X_s^i)] \right| ds \\ &\leq C \left( \int_0^t \max_{1 \leq i \leq N} |X_s^{i,N} - X_s^i| + \max_{1 \leq j \leq N} |X_s^{j,N} - X_s^j| \right) + \left| \frac{1}{N} \sum_{j=1}^N f(X_s^j) - \mathbf{E}[f(X_s^i)] \right| ds. \end{aligned}$$

Mit dem Gronwall-Lemma folgt nun

$$\begin{aligned} \max_{1 \leq i \leq N} |X_t^{i,N} - X_t^i| &\leq \int_0^t \left| \frac{1}{N} \sum_{j=1}^N f(X_s^j) - \mathbf{E}[f(X_s^i)] \right| ds \\ &\quad + \int_0^t e^{2C(t-s)} \int_0^s \left| \frac{1}{N} \sum_{j=1}^N f(X_r^j) - \mathbf{E}[f(X_r^i)] \right| dr ds \end{aligned}$$

Da  $X^1, X^2, \dots$  unabhängig sind, folgt mit dem Gesetz großer Zahlen und majorisierter Konvergenz

$$\int_0^t \left| \frac{1}{N} \sum_{j=1}^N f(X_s^j) - \mathbf{E}[f(X_s^i)] \right| ds \xrightarrow{N \rightarrow \infty}_{f_s} 0$$

und damit die Behauptung.  $\square$

**Bemerkung 3.12.** In der Anwendung – siehe (3.3) – verwenden wir

$$f(V_t^j) = \alpha \sum_{k=1}^{\infty} \delta_{T_k^j}(t),$$

wobei  $T_k^j$  der  $k$ -te Sprungzeitpunkt von  $V^j$  ist. Offenbar ist diese Funktion nicht Lipschitz-stetig, so wie wir es in den letzten beiden Propositionen gefordert ist. Demnach können wir nun nicht einmal schließen, dass das Modell aus (3.3) überhaupt eine eindeutige Lösung besitzt, oder ob es konvergiert. Dies ist momentan ein aktueller Forschungsgegenstand.