

Markovketten und -prozesse: Beispiele und Anwendungen

Johannes Gutenberg-Universität Mainz,
Sommersemester 2019

Matthias Birkner

Version vom 8. August 2019

Kommentare, Korrekturvorschläge, Hinweise auf (Tipp-)Fehler gerne per
Email an birkner@mathematik.uni-mainz.de senden

Inhaltsverzeichnis

1	Optimales Stoppen einer Markov-Kette	2
1.1	Beispiel Auswahlproblem	2
1.1.1	(Um-)Formulierung als Markovkette	2
1.2	Allgemeine Formulierung	4
1.2.1	Die optimale Strategie	7
1.2.2	Zurück zum Auswahlproblem	8
2	Rund um Y.-T. Chen, J. Choi, J.T. Cox, Ann. IHP 52 (2016), 286–322	10
2.1	Zum Moran-Modell	10
2.2	Kingmans Koaleszent	14
2.3	Zur Wright-Fisher-Diffusion	15
2.3.1	Übergangswahrscheinlichkeiten des (Blockzählprozesses des) Kingman-Koaleszenten und der Wright-Fisher-Diffusion	18
2.3.2	Ergänzung: Der „ ∞ -Koaleszent“	21
2.3.3	Ergänzung: Die Skelettkette des Kingman-Koaleszenten	22
2.3.4	Konvergenz des Typenanteilsprozesses	24
2.4	Wähler-Modelle auf endlichen Graphen	28
2.4.1	Einschlägige Martingale	29
2.4.2	Das Hauptergebnis	32
2.5	Zur „Beinahe-Exponentialität“ von Auftreffzeiten	38
2.6	Beispiel-Instanzen für Satz 2.22	42
2.6.1	Moran-Modell	42
2.6.2	Nächste-Nachbar-Irrfahrt auf dem Torus $(\mathbb{Z}/n\mathbb{Z})^d$ in $d = 2$	42
A	Hintergrundmaterial und Ergänzungen	51
A.1	Zeitkontinuierlichen Markovketten (mit endlichem Zustandsraum)	51
A.1.1	Kolmogorovs Vorwärts- und Rückwärtsgleichungen	52
A.2	Zur Skorokhod-Topologie auf dem Raum der càdlàg-Pfade	53

Ich danke Sabine Muth für Ihre Unterstützung beim L^AT_EX-en.

Kapitel 1

Optimales Stoppen einer Markov-Kette

1.1 Beispiel Auswahlproblem

Gegeben sind n paarweise verschiedene reelle Werte $w_1, w_2, \dots, w_n > 0$ (die wir nicht kennen, wir wissen nur die Anzahl - man denke an verschlossene Umschläge).

Wir bekommen die Werte nacheinander in zufälliger Reihenfolge gezeigt und dürfen uns zu jedem Zeitpunkt (ggfs. eingedenk der bereits gesehenen Werte) entscheiden, den gerade neu gezeigten Wert zu behalten, dann endet das Spiel (oder wir behalten gar keinen Wert, dann erhalten wir 0).

Gesucht: Eine Strategie, die mit möglichst großer Wahrscheinlichkeit den größten der n Werte auswählt.

Beispiele einfache Strategien:

- behalte den ersten gesehenen Wert
- behalte den k -ten gesehenen Wert
- betrachte $k/2$ Werte, behalte dann den nächsten gezeigten Wert, der die bisher gesehenen übertrifft.

1.1.1 (Um-)Formulierung als Markovkette

Wir folgen der Darstellung aus Dynkin und Juschkewitsch [DJ, Kapitel III].

Seien Z_1, Z_2, \dots, Z_n die (in dieser Reihenfolge) gezeigten Werte. Wir realisieren dies z.B. folgendermaßen: $Z_i := w_{U_i}$ mit (U_1, \dots, U_n) einer uniformen Permutation von $\{1, \dots, n\}$ (erzeugt etwa durch Ziehen ohne Zurücklegen aus Urne mit n nummerierten Kugeln).

Seien

$$X_0 := 1, \quad X_1 := \inf \{k : Z_k > Z_1\} \quad (\text{bzw.} = \infty, \text{ wenn es kein solches } k \text{ gibt})$$

$$X_i = \inf \{k : Z_k > \max\{Z_1, \dots, Z_{X_{i-1}}\}\}, \quad i \geq 2 \quad (\text{mit Setzung } = \infty, \text{ wenn es kein solches } k \text{ gibt})$$

X_i ist der (Zeit-)Index des i -ten „Rekordwerts“.

Offenbar: Nur die (zufälligen) Zeitpunkte X_0, X_1, X_2, \dots kommen als „Stopp-Momente“ in Frage (wenn man zu einem Zeitpunkt $t \notin \{X_0, X_1, X_2, \dots\}$ aufhört, erhält man sicherlich nicht den größten Wert).

Lemma 1.1. X_0, X_1, \dots, X_n ist eine Markovkette (auf $\{1, 2, \dots, n\} \cup \{\infty\}$).

Beweis. Sei $\pi^{(k)}$ die Permutation von Z_1, \dots, Z_k der Größe nach, d.h. $Z_{\pi^{(k)}(1)} < Z_{\pi^{(k)}(2)} < \dots < Z_{\pi^{(k)}(k)}$; offenbar ist $\pi^{(k)} \sim \text{Unif}(\mathcal{S}_k)$ mit $\mathcal{S}_k = \{\text{Permutationen von } \{1, \dots, k\}\}$.

Gegeben $\pi^{(k)}$ entsteht $\pi^{(k+1)}$ folgendermaßen: das Urbild von $k+1$ unter $\pi^{(k+1)}$, $(\pi^{(k+1)})^{-1}(k+1)$, wird uniform aus $\{1, \dots, k+1\}$ gewählt, die übrigen Werte werden aus $\pi^{(k)}$ unter Beachtung der relativen Reihenfolge übernommen, d.h. für $1 \leq j \leq k+1$ ist

$$\pi^{(k+1)}(j) = \begin{cases} \pi^{(k)}(j), & \text{falls } j < (\pi^{(k+1)})^{-1}(k+1), \\ k+1, & j = (\pi^{(k+1)})^{-1}(k+1), \\ \pi^{(k)}(j-1), & \text{sonst.} \end{cases}$$

Für $\rho \in \mathcal{S}_{k+1}, \rho' \in \mathcal{S}_k$ ist

$$\mathbb{P}(\pi^{(k+1)} = \rho, \pi^{(k)} = \rho') = \mathbb{P}(\pi^{(k+1)} = \rho) = \frac{1}{(k+1)!},$$

wenn ρ' aus ρ durch Weglassen von $k+1$ und entsprechendes „Zusammenschieben der Bilder“ entsteht (ansonsten ist die Wahrscheinlichkeit links = 0).

Somit ist

$$\mathbb{P}(k+1 \text{ ist neuer Rekord} \mid \pi^{(k)} = \rho') = \frac{1}{k+1}$$

(für alle $\rho' \in \mathcal{S}_k$).

Damit finden wir

$$\begin{aligned} & \mathbb{P}(X_0 = a_0, X_1 = a_1, \dots, X_i = a_i) \\ &= \mathbb{P}(a_0 \text{ ist Rekord, } a_0 + 1, \dots, a_1 - 1 \text{ sind keine Rekorde, } a_1 \text{ ist Rekord, } \dots, \\ & \quad a_{i-1} \text{ ist Rekord, } a_{i-1} + 1, \dots, a_i - 1 \text{ sind keine Rekorde, } a_i \text{ ist Rekord}) \\ &= \underbrace{\frac{1}{a_0}}_{\frac{1}{a_0}} \cdot \underbrace{\left(1 - \frac{1}{a_0 + 1}\right)}_{\frac{a_0}{a_0 + 1}} \underbrace{\left(1 - \frac{1}{a_0 + 2}\right)}_{\frac{a_0 + 1}{a_0 + 2}} \dots \underbrace{\left(1 - \frac{1}{a_1 - 1}\right)}_{\frac{a_1 - 2}{a_1 - 1}} \cdot \frac{1}{a_1} \cdot \dots \cdot \frac{1}{a_{i-1}} \cdot \left(1 - \frac{1}{a_{i-1} + 1}\right) \dots \left(1 - \frac{1}{a_i - 1}\right) \cdot \frac{1}{a_i} \\ &= \frac{1}{a_i!} (a_0 \cdot (a_0 + 1) \dots (a_1 - 2)) \cdot 1 \cdot (a_1(a_1 + 1) \dots (a_2 - 2)) \cdot 1 \cdot \dots \cdot (a_{i-1} \cdot (a_{i-1} + 1) \dots (a_i - 2)) \cdot 1 \\ &= \frac{1}{a_0} \cdot \frac{a_0}{a_0 + 1} \cdot \frac{a_0 - 1}{a_0 + 2} \dots \frac{a_1 - 2}{a_1 - 1} \cdot \frac{1}{a_1} \cdot \frac{a_1}{a_1 + 1} \dots \frac{a_2 - 2}{a_2 - 1} \cdot \frac{1}{a_2} \dots \frac{1}{a_1 - 1} \cdot \frac{a_i - 1}{a_i - 1 + 1} \cdot \frac{a_i - 1 + 1}{a_i - 1 + 2} \dots \frac{a_i - 2}{a_i - 1} \cdot \frac{1}{a_i} \\ &= \frac{1}{(a_1 - 1)(a_2 - 1) \dots (a_i - 1) \cdot a_i} \end{aligned}$$

und daher

$$\mathbb{P}(X_{i+1} = a_{i+1} \mid X_0 = a_0, \dots, X_i = a_i) = \frac{a_i}{a_{i+1}(a_{i+1} - 1)} \quad (\text{für } 1 = a_0 < a_1 < \dots < a_i < a_{i+1} \leq n)$$

$$\mathbb{P}(X_{i+1} = \infty \mid X_0 = a_0, \dots, X_i = a_i) = 1 - \sum_{\ell=a_i+1}^n \frac{a_i}{\ell(\ell+1)}$$

Folglich ist (X_0, X_1, \dots, X_n) eine Markovkette mit Übergangsmatrix $P = (p(j, k))_{j, k}$

$$\begin{aligned} p(k, \ell) &= \frac{k}{\ell(\ell-1)} \quad \text{für } 1 \leq k < \ell \leq n, \\ p(k, \infty) &= 1 - \sum_{\ell=k+1}^n \frac{k}{\ell(\ell-1)} = 1 - k \cdot \underbrace{\sum_{\ell=k+1}^n \left(\frac{1}{\ell-1} - \frac{1}{\ell} \right)}_{\frac{1}{k} - \frac{1}{n}} = \frac{k}{n}, \\ p(\infty, \infty) &= 1, \quad p(j, m) = 0 \quad \text{für } m \leq j \end{aligned} \tag{1.1}$$

□

Insbesondere ist

$$f(k) := \mathbb{P}(\text{erhalte max. Wert} \mid \text{Stoppe bei einem Rekord nach } k \text{ Schritten}) = \frac{k}{n} \tag{1.2}$$

(und wir setzen $f(\infty) = 0$, passend zur Interpretation, dass die gesuchte bedingte Wahrscheinlichkeit = 0 ist, wenn man überhaupt nicht stoppt).

1.2 Allgemeine Formulierung

Sei E endlich oder abzählbar, $p = (p(x, y))_{x, y \in E}$ eine stochastische Matrix (irreduzibel, sagen wir) und (X_n) eine p -Markovkette auf E .

Weiter sei eine „Auszahlungsfunktion“

$$f : E \rightarrow [0, \infty)$$

gegeben.

Ein „Spieler“ wählt / definiert eine Stoppzeit τ und erhält dann die Auszahlung $f(X_\tau)$ (auf $\{\tau < \infty\}$), bzw. 0 auf $\{\tau = \infty\}$, d.h. die erwartete Auszahlung bei Start in $X_0 = x$ ist dann

$$\mathbb{E}_x[f(X_\tau)\mathbb{1}_{\{\tau < \infty\}}]$$

Der (erwartete) „Wert des Spiels“ bei Start in $x \in E$ ist

$$v(x) := \sup_{\tau \text{ Stoppzeit}} \mathbb{E}_x[f(X_\tau)\mathbb{1}_{\{\tau < \infty\}}] \tag{1.3}$$

Offenbar ist

$$v(x) \geq \mathbb{E}_x[f(X_0)] = f(x) \quad (\geq 0)$$

(wähle $\tau \equiv 0$).

Proposition 1.2. *Es gilt*

$$v(x) \geq Pv(x) \left(:= \sum_y p(x, y)v(y) \right)$$

für $x \in \mathbb{E}$.

Eine Funktion v mit dieser Eigenschaft heißt *superharmonisch* (bezüglich p). In der Literatur ist auch der Begriff „exzessiv“ gebräuchlich (man fordert dann auch $v(\cdot) \geq 0$, siehe z.B. [DJ]).

Beweis. Sei $x \in \mathbb{E}$, wähle $\varepsilon > 0$.

Zu $y \in E$ gibt (nach Definition von $v(y)$) es eine Stoppzeit $\tau^{(y)}$ mit

$$\mathbb{E}_y[f(X_{\tau^{(y)}})\mathbb{1}_{\{\tau^{(y)} < \infty\}}] \geq v(y) - \varepsilon$$

Die Idee ist nun, die Markovkette den ersten Schritt einfach machen zu lassen und wenn sie in $X_1 = y$ angekommen ist, dann von dort aus die für den Startpunkt y (bis auf ε) optimale Stopp-Regel $\tau^{(y)}$ zu verwenden.

Man kann dies auf verschiedene Weisen formalisieren (die je nach Vorkenntnissen des Lesers unterschiedlich leicht „verdaulich“ sind):

a) Schreibe

$$\tau := 1 + \tau^{(X_1)} \circ \theta,$$

wobei $\theta: E^{\mathbb{N}_0} \rightarrow E^{\mathbb{N}_0}$ den Shift-Operator auf dem Raum der Pfade bezeichnet (siehe z.B. [Kl, Bsp. 20.20]).

b) Alternativ beobachte

$$\tau \text{ Stoppzeit} \Leftrightarrow \{\tau = n\} \stackrel{(\in \mathcal{F}_n)}{=} \{(X_0, X_1, \dots, X_n) \in A_n\} \text{ für eine gewisse Menge } A_n \subset E^{n+1}$$

(Was A_n ist, hängt von der Definition von τ ab, und die A_0, A_1, A_2, \dots müssen geeignete „Kompatibilitätsbedingungen“ erfüllen, denn $\{\tau = n\} \cap \{\tau = m\} = \emptyset$ für $n \neq m$; siehe auch z.B. [LPW, Kap. 17] für Martingalthorie und verwandte Konzepte mit „rein elementaren“ Mitteln.)

Sei $\{\tau^{(y)} = n\} = \{(X_0, \dots, X_n) \in A_{n,y}\}$ für $n \in \mathbb{N}_0$, dann definiere τ via

$$\{\tau = n\} = \bigcup_{y \in E} \left(\{X_1 = y\} \cap \{(X_1, \dots, X_n) \in A_{n-1,y}\} \right), \quad n \in \mathbb{N}$$

(und $\{\tau = 0\} = \emptyset$, $\{\tau = \infty\} = \left(\bigcup_{n \in \mathbb{N}} \{\tau = n\} \right)^c$).

Nach Konstruktion ist $\tau \geq 1$ und

$$v(x) \geq \mathbb{E}_x[f(X_\tau)\mathbb{1}_{\{\tau < \infty\}}] = \sum_y \mathbb{E}_x[f(X_\tau)\mathbb{1}_{\{\tau < \infty, X_1=y\}}]$$

mit der Definition von τ und der Markov-Eigenschaft

$$= \sum_y \mathbb{E}_x[f(X_{\tau^{(y)} \circ \theta_{1+1}})\mathbb{1}_{\{\tau^{(y)} \circ \theta_{1+1} < \infty, X_1=y\}}]$$

$$= \sum_y \underbrace{\mathbb{P}_x(X_1 = y)}_{=p(x,y)} \underbrace{\mathbb{E}_x[\dots | X_1 = y]}_{=\mathbb{E}_y[f(X_{\tau(y)}) \mathbb{1}_{\{\tau(y) < \infty\}}]} \geq \sum_y \mathbb{P}(x, y)v(y) - \varepsilon,$$

$\geq v(y) - \varepsilon$

mit $\varepsilon \downarrow 0$ folgt die Behauptung. □

Beobachtung 1.3. $g : E \rightarrow [0, \infty)$ superharmonisch (d.h. $Pg \leq g$), τ Stoppzeit, so gilt

$$\mathbb{E}_x[g(X_\tau) \mathbb{1}_{\{\tau < \infty\}}] \leq g(x) \quad \text{für alle } x \in E \quad (1.4)$$

Falls $\tau < \infty$ f.s. gilt, so kann man (1.4) auch als $\mathbb{E}_x[g(X_\tau)] \leq g(x)$ schreiben und kann die Forderung $g \geq 0$ durch $\inf_{x \in E} g(x) > -\infty$ ersetzen (betrachte $\tilde{g} = g - \inf g$).

Ganz ohne Beschränktheitsannahmen gilt (1.4) allerdings nicht: Betrachte z.B. $g(x) = x$ und die symmetrische gewöhnliche symmetrische Irrfahrt auf \mathbb{Z} .

Beweis von Beob. 1.3. Wir betrachten zwei Beweise: der erste ist kürzer, verwendet allerdings Sätze aus der Martingaltheorie, der zweite ist elementarer.

1) $S_n^g := g(X_n) - g(X_0)$ ist Supermartingal unter (jedem) \mathbb{P}_x , denn

$$\mathbb{E}_x[g(X_n) | \mathcal{F}_{n-1}] = \mathbb{E}_x[g(X_n) | X_{n-1}] = \sum_{z \in E} p(X_{n-1}, z)g(z) = Pg(X_{n-1}) \leq g(X_{n-1})$$

(wir notieren $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ mit $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ für die „kanonische“ Filtration zur Markovkette (X_n)).

Das Optional-Sampling-Theorem (z.B. [Kl, Satz 10.11]) liefert

$$\mathbb{E}_x[g(X_\tau)] - g(x) = \mathbb{E}_x[S_\tau^g] \leq \mathbb{E}_x[S_0^g] = 0$$

2) Ohne (Super-)Martingalargument: (z.B. wie in [DJ, S. 96f], sozusagen inspiriert vom Riesz'schen Darstellungssatz)

Wähle $\alpha \in (0, 1)$, setze $\varphi(x) := g(x) - \alpha Pg(x) (\geq 0)$, dann ist

$$g(x) = \varphi(x) + \alpha P\varphi(x) + \alpha^2 P^2\varphi(x) + \dots + \alpha^n P^n\varphi(x) + \alpha^{n+1} P^{n+1}g(x)$$

(für jedes n , teleskopierende Summe).

Nun ist

$$0 \leq P^{n+1}g(x) = \underbrace{(P^n(Pg))}_{\leq g}(x) \leq \dots \leq g(x),$$

wegen $|\alpha| < 1$ gilt $\alpha^{n+1} P^{n+1}g(x) \xrightarrow{n \rightarrow \infty} 0$, d.h. (mit $P^0 := I$ und dem Satz von der monotonen Konvergenz)

$$g(x) = \sum_{n=0}^{\infty} \alpha^n P^n\varphi(x) = \mathbb{E}_x\left[\sum_{n=0}^{\infty} \alpha^n \varphi(X_n)\right] \quad (1.5)$$

(g wird als das „ α -Potential“ der (nicht-neg.) Funktion φ dargestellt).

Somit

$$\begin{aligned}
\mathbb{E}_x[\alpha^\tau g(X_\tau)] &= \mathbb{E}_x\left[\sum_{n=0}^{\infty} \alpha^{n+\tau} (P^n \varphi)(X_\tau)\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x\left[\alpha^{n+\tau} \underbrace{(P^n \varphi)(X_\tau)}_{=\mathbb{E}[\varphi(X_{n+\tau})|\mathcal{F}_\tau]}\right] = \sum_{n=0}^{\infty} \mathbb{E}_x[\alpha^{n+\tau} \varphi(X_{n+\tau})] \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x[\alpha^{n+\tau} \varphi(X_{n+\tau})] = \mathbb{E}_x\left[\sum_{m=\tau}^{\infty} \alpha^m \varphi(X_m)\right] \leq g(x)
\end{aligned}$$

(vgl. mit (1.5), alle Summanden sind ≥ 0). Mit $\alpha \uparrow 1$ folgt (1.4). □

Beobachtung 1.4. $v(x)$ ($= \sup_{\tau} \mathbb{E}_x[f(X_\tau) \mathbb{1}_{\{\tau < \infty\}}]$) ist die kleinste exzessive (superharmonische) Majorante von f .

(Es ist $v(x) \geq f(x) \geq 0$ und $v(\cdot)$ ist superharmonisch nach Prop. 1.2; sei $g \geq f$, g exzessiv $\Rightarrow g(x) \geq \mathbb{E}_x[g(X_\tau) \mathbb{1}_{\{\tau < \infty\}}] \geq \mathbb{E}_x[f(X_\tau) \mathbb{1}_{\{\tau < \infty\}}]$ für jede Stoppzeit τ , also $g \geq v$.)

1.2.1 Die optimale Strategie

Sei f die Auszahlungsfunktion und v aus (1.3) der Wert des Spiels (man kann v prinzipiell berechnen via

$$v(x) = \lim_{n \rightarrow \infty} Q^n f(x) \quad \text{mit } Qf(x) := f(x) \vee Pf(x),$$

vgl. [DJ, S. 126]),

$$\Gamma := \{x \in E : v(x) = f(x)\} \quad \text{die „Stützmenge“}$$

($\Gamma \neq \emptyset$ stets, sonst könnte man v „nach unten“ verschieben)

Setze

$$\tau_\Gamma := \inf\{n \geq 0 : X_n \in \Gamma\}$$

Satz 1.5. $|E| < \infty$ und p irreduzibel, so ist $v(x) = \mathbb{E}_x[f(X_{\tau_\Gamma})]$ (und $P_x(\tau_\Gamma < \infty) = 1 \forall x$), d.h. τ_Γ ist optimal.

Beweis. Endlichkeit von E und Irreduzibilität von p garantieren $\tau_\Gamma < \infty$ f.s. für jeden Startpunkt. Setze

$$h(x) := \mathbb{E}_x[f(X_{\tau_\Gamma})],$$

nach Definition ist $h(x) \leq v(x)$ und wegen $v = f$ auf Γ ist $h(x) = \mathbb{E}_x[v(X_{\tau_\Gamma})]$ selbst exzessiv:

Sei $\tau' := \inf\{k \geq 1 : X_k \in \Gamma\}$, so ist

$$Ph(x) = \sum_y p(x, y) h(y) = \sum_y p(x, y) \mathbb{E}_y[v(X_{\tau_\Gamma})] = \mathbb{E}_x[v(X_{\tau'})] \begin{cases} = \mathbb{E}_x[v(X_{\tau_\Gamma})] = h(x), & x \notin \Gamma, \\ \leq v(x) = f(x) = h(x), & x \in \Gamma \end{cases}$$

mit Beobachtung 1.3.

Zeige

$$h \geq f \tag{1.6}$$

Auf Γ gilt $h = f$ und somit ist (1.6) dort erfüllt. Wenn (1.6) nicht gälte, so gäbe es ein $a \in E \setminus \Gamma$ mit

$$f(a) - h(a) = \max_{x \in E} (f(x) - h(x)) > 0.$$

Dann erfüllt $\tilde{h}(x) := h(x) + \underbrace{(f(a) - h(a))}_{>0}$: \tilde{h} ist exzessiv und $\tilde{h} \geq f$

$\Rightarrow \tilde{h} \geq v$, also $f(a) = \tilde{h}(a) \geq v(a)$, d.h. $a \in \Gamma$ im Widerspruch zur Annahme.

Somit gilt (1.6), d.h. $h \geq v$, also $h = v$. □

Bemerkung 1.6. 1) Für $|E| = \infty$ gilt Satz 1.5 i.A. nicht (siehe [DJ, Beispiel auf S. 101]):

Betrachte z.B. $E = \mathbb{N}$, $f(1) = 1$, $f(n) = 1 - \frac{1}{n}$ für $n \geq 1$, $p(n, n+1) = 1 - 1/n^2 = 1 - p(n, n-1)$ für $n \geq 2$, $p(1, 1) = 1$.

Es ist $\Gamma = \{1\}$, aber τ_Γ ist nicht optimal: Für $x \geq 2$ ist $\mathbb{P}_x(\tau_\Gamma = \infty) = \prod_{k=x}^{\infty} (1 - k^{-2}) > 0$.

2) Aber: Zu $\varepsilon > 0$ sei $\Gamma_\varepsilon = \{x \in E : v(x) \leq f(x) + \varepsilon\}$ (die „ ε -Stützmenge“) und $\tau_\varepsilon := \inf\{n \geq 0 : X_n \in \Gamma_\varepsilon\}$, dann ist $\mathbb{E}_x[f(X_{\tau_\varepsilon})] \geq v(x) - \varepsilon$ (vgl. [DJ, S. 103]).

1.2.2 Zurück zum Auswahlproblem

Wir erinnern uns an die Markov-Ketten-Formulierung des Auswahlproblems aus Lemma 1.1: Zustandsraum $E = \{1, \dots, n\} \cup \{\infty\}$, Übergangsmatrix p aus (1.1)

$$p(k, \ell) = \frac{k}{\ell(\ell-1)} \text{ für } 1 \leq k < \ell \leq n, \quad p(k, \infty) = \frac{k}{n}, \quad p(\infty, \infty) = 1$$

Die Auszahlungsfunktion aus (1.2) ist

$$f(k) = \frac{k}{n} \text{ für } 1 \leq k \leq n, \quad f(\infty) = 0$$

Gesucht:

Minimales $v : E \rightarrow [0, \infty)$ mit $v \geq f$ und $v \geq Pv$.

Offenbar $v(\infty) = 0$, also muss gelten

$$v(k) \geq \frac{k}{n}, \quad v(k) \geq \sum_{\ell=k+1}^n \frac{k}{\ell(\ell-1)} v(\ell), \quad 1 \leq k \leq n$$

(zudem ist klar: $v(n) = \frac{n}{n} = 1$).

Angenommen, wir kennen $v(\ell)$ für $\ell > k$ schon:

$$\Rightarrow v(k) = \frac{k}{n} \vee \sum_{\ell=k+1}^n \frac{k}{\ell(\ell-1)} v(\ell) \quad (= f(k) \vee Pv(k)),$$

also

$$v(n-1) = \max \left\{ \frac{n-1}{n}, \underbrace{(n-1) \frac{1}{n(n-1)}}_{=\frac{1}{n}} \right\} = \frac{n-1}{n} = f(n-1), \dots,$$

$$v(k) = \max \left\{ \frac{k}{n}, k \underbrace{\sum_{\ell=k+1}^n \frac{1}{\ell(\ell-1)} \cdot \frac{\ell}{n}}_{=\frac{k}{n} \sum_{\ell=k+1}^n \frac{1}{\ell-1} = \frac{k}{n} \left(\frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{n} \right)} \right\} = \frac{k}{n},$$

sofern

$$\frac{1}{k} + \frac{1}{k+1} + \dots + \frac{1}{n} \leq 1 \quad (1.7)$$

gilt.

Sei $k_n :=$ kleinste natürliche Zahl, für die (1.7) gilt. Für $k < k_n$ ist $v(k) > \frac{k}{n}$ (aus obigem, denn $v(\ell) \geq f(\ell)$ stets), somit ist $\Gamma = \{k_n, k_n + 1, k_n + 2, \dots, n\}$. Die optimale Stopp-Strategie lautet also

$$\tau_{\text{opt}} = \inf \{k \geq k_{n-1} : Z_1, \dots, Z_{k-1} < Z_k\}.$$

Für großes n kann man die Bedingung (1.7) approximativ in Integralform formulieren:

$$\frac{1}{k} + \dots + \frac{1}{n} \sim \int_k^n \frac{1}{x} dx = \log n - \log k = \log \frac{n}{k},$$

d.h. es ist $k_n \sim \frac{n}{e}$ für $n \rightarrow \infty$.

Zur Erfolgswahrscheinlichkeit der optimalen Strategie:

$$\begin{aligned} & \mathbb{P}(Z_{\tau_{\text{opt}}} \text{ ist maximaler Wert}) \\ &= \sum_{m=k_n}^n \mathbb{P}(\tau_{\text{opt}} = m, Z_m \text{ ist max. Wert}) \\ &= \sum_{m=k_n}^n \mathbb{P}(\tau_{\text{opt}} = m) \mathbb{P}(Z_m \text{ ist max. Wert} | \tau_{\text{opt}} = m) \\ &= \sum_{m=k_n}^n \frac{k_n - 1}{m(m-1)} \cdot \frac{m}{n} = \frac{k_n - 1}{n} \left(\frac{1}{k_n - 1} + \frac{1}{k_n} + \dots + \frac{1}{n-1} \right), \end{aligned}$$

wobei wir

$$\begin{aligned} \mathbb{P}(\tau_{\text{opt}} = m) &= \mathbb{P} \left((\pi^{(m)})^{-1}(m) = m, (\pi^{(m)})^{-1}(m-1) < k_n \right) = \frac{1}{m} \cdot \frac{k_n - 1}{m-1}, \\ \mathbb{P}(Z_m \text{ ist max. Wert} | \tau_{\text{opt}} = m) &= f(m) = \frac{m}{n} \end{aligned}$$

verwenden.

Mit $k_n \sim \frac{n}{e}$ ergibt sich

$$\mathbb{P}(Z_{\tau_{\text{opt}}} \text{ ist maximaler Wert}) \sim \frac{1}{e} (\approx 0,368) \quad \text{für } n \rightarrow \infty.$$

Kapitel 2

Rund um Y.-T. Chen, J. Choi, J.T. Cox, Ann. IHP 52 (2016), 286–322

2.1 Zum Moran-Modell

Definition 2.1 ((Neutrales 2 Typ-)Moran-Modell¹). Man betrachtet eine Population von konstant N (haploiden) Individuen, jedes Individuum besitzt eine unabhängige, $\text{Exp}(1)$ -verteilte Lebenszeit und wird am Ende seiner Lebenszeit durch den Nachkommen eines rein zufällig aus der Population gezogenen Individuums ersetzt (es gibt nur ein Elter und, sagen wir, man kann durch sein eigenes Kind ersetzt werden).

Wir nehmen zusätzlich an, dass es zwei Typen, sagen wir 0 und 1, gibt, die ohne Mutation vererbt werden.

Sei

$$\xi_t(i) = \text{Typ von Individuum } i \text{ zur Zeit } t. \quad (2.1)$$

Die Dynamik des Prozesses $(\xi_t)_{t \geq 0} = (\xi_t(1), \xi_t(2), \dots, \xi_t(N))_{t \geq 0}$ mit Zustandsraum $E := \{0, 1\}^N$, der über die Typen der Individuen in der Population Buch führt, ist somit folgende: Für $\eta, \eta' \in E$, $\eta \neq \eta'$ ist die Sprungrate von η nach η' ,

$$q_{\eta, \eta'} = \lim_{h \downarrow 0} \frac{\mathbb{P}(\xi_{t+h} = \eta' \mid \xi_t = \eta)}{h} \quad (2.2)$$

gegeben durch

$$q_{\eta, \eta'} = \begin{cases} \frac{X_t}{N}, & \text{falls } \eta' \text{ aus } \eta \text{ entsteht, indem an genau einer Stelle eine 0 zu} \\ & \text{einer 1 wird (d.h. für ein } i \in [N] \text{ gilt } \eta'(i) = 1, \eta(i) = 0 \text{ und} \\ & \eta'(j) = \eta(j) \text{ für } j \neq i), \\ \frac{N - X_t}{N}, & \text{falls } \eta' \text{ aus } \eta \text{ entsteht, indem an genau einer Stelle eine 1 zu} \\ & \text{einer 0 wird (d.h. für ein } i \in [N] \text{ gilt } \eta'(i) = 0, \eta(i) = 1 \text{ und} \\ & \eta'(j) = \eta(j) \text{ für } j \neq i), \\ 0, & \text{sonst.} \end{cases} \quad (2.3)$$

¹Nach Patrick Alfred Pierce Moran, 1917–1988 benannt, von ihm in [M58] eingeführt.

wobei

$$X_t = X(\xi_t) = \sum_{i=1}^N \xi_t(i) \quad (2.4)$$

die Anzahl Typ-1-Individuen zur Zeit t bezeichnet.

Für die Diagonaleinträge gilt

$$q_{\eta,\eta} = - \sum_{\eta' \neq \eta} q_{\eta,\eta'} = - \left(X(\eta) \frac{N - X(\eta)}{N} + (N - X(\eta)) \frac{X(\eta)}{N} \right) = - \frac{2}{N} X(\eta) (N - X(\eta)) \quad (2.5)$$

Beobachtung 2.2. Insbesondere: 0 und N sind (die einzigen) absorbierende Zustände, $(X_t)_{t \geq 0}$ ist ein (beschränktes) Martingal, für die Fixierungswahrscheinlichkeiten gilt

$$\mathbb{P}(\text{Absorption in } N \mid X_0 = x) = \frac{x}{N}$$

Beweis. Allgemein:

$$Qf(x) = \sum_y q_{xy} f(y) = \sum_y q_{xy} (f(y) - f(x))$$

ist der *Generator* (der zugehörigen Übergangshalbgruppe), und

$$M_t^{(f)} := f(X_t) - f(X_0) - \int_0^t Qf(X_s) ds, \quad t \geq 0$$

ist ein Martingal. (Wende dies an auf $f(x) = x$.)

Es gilt $T_{\text{fix}} := \inf \{t \geq 0 : X_t \in \{0, N\}\} < \infty$ und (mit dem optional sampling-Satz)

$$x = \mathbb{E}_x[X_0] = \mathbb{E}_x[X_{T_{\text{fix}}}] = N \cdot \mathbb{P}_x(X_{T_{\text{fix}}} = N) + 0 \cdot \mathbb{P}_x(X_{T_{\text{fix}}} = 0).$$

□

Graphische Konstruktion

Für jedes geordnete Paar (i, j) , $i, j \in \{1, \dots, N\}$, $i \neq j$ sei $(N_t^{(i,j)})_{t \geq 0}$ ein Poissonprozess auf \mathbb{R}_+ mit Rate $\frac{1}{N}$, u.a. für verschiedene Paare. Zu den Sprungzeiten von $(N_t^{(i,j)})_{t \geq 0}$ stirbt Individuum j und wird durch einen Nachkommen von Individuum i ersetzt (s.a. Abb. 2.1).

Bild an der Tafel

Abbildung 2.1: Im Bild: N Kopien der Zeitachse, gerichtete Pfeile zwischen ihnen zu den Sprungzeitpunkten von u.a. Poissonprozessen; das Individuum an der Pfeilspitze stirbt jeweils und wird durch einen Nachkommen des Individuums am Pfeilschaft ersetzt.

Die Dynamik des Prozesses $(\xi_t(1), \xi_t(2), \dots, \xi_t(N))_{t \geq 0}$ aus (2.1), der über die Typen der Individuen in der Population Buch führt (und nicht nur über die Anzahlen) ist somit folgende:

Ersetze zu jedem Sprungzeitpunkt t von $N^{(i,j)}$ den Typ $X_{t-}(j)$ durch $X_t(j) = X_{t(-)}(i)$.

Dies ist wohldefiniert, da unabhängige Poissonprozesse f.s. keine gemeinsamen Sprungzeitpunkte besitzen. Diese Konstruktion ist ein Spezialfall einer sogenannten Harris-Konstruktion², ein in der Theorie der interagierenden Teilchensysteme übliches (und nützliches) Werkzeug.

Bemerkung 2.3 (Ablesern der Genealogie und der Typen aus der graphischen Konstruktion). Für $t > 0$, $i \in [N]$ sei

$$A_s^{(i,t)} = \text{Nr. des Ahnenindividuums zur Zeit } t-s \text{ von Ind. } i \text{ zur Zeit } t \\ (\text{für } 0 \leq s \leq t, \text{ Werte in } [N])$$

Zur Konstruktion von $A^{(i,t)} = (A_s^{(i,t)})_{0 \leq s \leq t}$ verfolgen wir die derzeitige „Zeitachse“ rückwärts und folgen den Pfeilen jeweils in entgegengesetzter Richtung, vgl. auch Abb. 2.1.

In Formeln können wir den Pfad von $A^{(i,t)}$ beispielsweise folgendermaßen fassen (wir schreiben $N^{(j,i)}([a, b])$ für die Anzahl Sprünge des Poissonprozesses $N^{(j,i)}$ im Zeitintervall $[a, b]$):

Sei $T_0^{(i,t)} := 0$, $\tilde{A}_0^{(i,t)} := A_0^{(i,t)} := i$, für $k \in \mathbb{N}$ setzen wir

$$T_k^{(i,t)} := \inf \{ u > T_{k-1}^{(i,t)} : \text{es gibt ein } j \neq i \text{ mit } N^{(j, \tilde{A}_{k-1}^{(i,t)})}([t-u, t - T_{k-1}^{(i,t)}]) = 1 \}$$

bzw. $T_k^{(i,t)} := t$, falls es kein solches u gibt. Falls $T_k^{(i,t)} = t$ gilt, so setzen wir $M^{(i,t)} := k$ und wir brechen die Konstruktion hier ab, andernfalls sei

$$\tilde{A}_k^{(i,t)} \text{ das (f.s.) eindeutig bestimmte } j \text{ mit } N^{(j, \tilde{A}_{k-1}^{(i,t)})}([t - T_k^{(i,t)}, t - T_{k-1}^{(i,t)}]) = 1$$

und wir setzen die Konstruktion fort. Da die endlich vielen Poissonprozesse $N^{(j,i)}$ f.s. keine Häufungspunkte in $[0, t]$ besitzen, bricht die Konstruktion mit Wahrscheinlichkeit 1 nach endlich vielen Schritten ab und wir setzen dann für $0 < s \leq t$

$$A_s^{(i,t)} := \tilde{A}_\ell^{(i,t)} \text{ falls } T_\ell^{(i,t)} \leq s < T_{\ell+1}^{(i,t)} \text{ für } 0 \leq \ell < M^{(i,t)}$$

bzw. $A_t^{(i,t)} := \tilde{A}_{M^{(i,t)}-1}^{(i,t)}$.

$(A_s^{(i,t)})_{0 \leq s \leq t}$ ist eine zeitkontinuierliche Markovkette mit (vollkommen symmetrischen) Sprungraten

$$q_{jk} = \begin{cases} \frac{1}{N}, & k \neq j, \\ -\frac{N-1}{N}, & k = j, \end{cases} \quad (2.6)$$

² nach Theodore Edward Harris, 1919–2005 benannt

man nennt eine solche Kette auch eine (zeitkontinuierliche) „Irrfahrt auf dem vollständigen Graphen V_N der Ordnung N “.

Für $i_1 \neq i_2$ bewegen sich $A^{(i_1,t)}$ und $A^{(i_2,t)}$ unabhängig bis zum „Verschmelzungszeitpunkt“

$$\tau_{i_1,i_2} := \inf\{s \in [0, t] : A_s^{(i_1,t)} = A_s^{(i_2,t)}\},$$

ab dann, d.h. für $u \geq \tau_{i_1,i_2}$, gilt $A_u^{(i_1,t)} = A_u^{(i_2,t)}$.

Für paarweise verschiedene i_1, i_2, \dots, i_n ($\leq N$) bilden

$$A^{(i_1,t)}, \dots, A^{(i_n,t)} \text{ ein System verschmelzender Irrfahrten auf } V_n$$

und mit

$$k \sim_{s,N} \ell : \iff A_s^{(i_k,t)} = A_s^{(i_\ell,t)}, \quad 1 \leq k, \ell \leq n$$

ist

$$\mathcal{R}_s^{(n,N)} := \text{Äquivalenzklassen bezüglich } \sim_{s,N}, \quad s \in [0, t] \quad (2.7)$$

ein (zeittransformierter) Kingman- n -Koaleszent. (siehe Abschnitt 2.2 unten. $(\mathcal{R}_{N_s/2}^{(n,N)})_{s \geq 0}$ wäre wörtlich ein Koaleszent, wenn wir die Zeitachsen in der graphischen Konstruktion „bis $-\infty$ fortsetzten.“)

Aus der Konstruktion ergibt sich folgende (realisierungsweise Form) der „Dualität“:

$$\xi_t(i) = \xi_0(A_t^{(i,t)}) \quad \text{für } 1 \leq i \leq N, t > 0. \quad (2.8)$$

Beweisskizze. Die Tatsache, dass $A^{(i,t)}$ eine zeitkontinuierliche Markovkette ist, folgt anschaulich gesehen aus der Unabhängigkeit der Zuwächse der „treibenden“ Poissonprozesse $N^{(j,k)}$, die symmetrische Form der Sprungratenmatrix (2.6) stammt daher, dass alle Poissonprozesse dieselbe Rate $1/N$ haben. Wenn aktuell $A_s^{(i,t)} = j$, so gibt es für $0 < h \ll 1$ und jedes $j' \neq j$ mit Wahrscheinlichkeit $\approx h/N$ einen Sprung von $N^{(j',j)}$ im Zeitintervall $[t-s-h, t-s)$ und dann springt $A^{(i,t)}$ von j nach j' .

Etwas formaler: Sei

$$\mathcal{F}_u^t := \sigma(N^{(j,k)}([a,b]) : j \neq k, t-u \leq a < b \leq t)$$

die σ -Algebra, die die Informationen über alle Sprünge der $N^{(j,k)}$ zwischen $t-u$ und t enthält. Offenbar kann man $A_s^{(i,t)}$ für $s \leq u$ anhand der Pfade der $N^{(j,k)}$ zwischen $t-u$ und t rekonstruieren (d.h. $A_s^{(i,t)}$ ist \mathcal{F}_u^t -messbar für $s \leq u$) und für $s < t$, $j \neq j' \in [N]$ ist auf dem Ereignis $\{A_s^{(i,t)} = j\}$

$$\begin{aligned} & \frac{1}{h} \mathbb{P}(A_{s+h}^{(i,t)} = j' | \mathcal{F}_s^t) \\ &= \frac{1}{h} \mathbb{P}(N^{(j',j)}([t-s-h, t-s)) = 1, N^{(j'',j)}([t-s-h, t-s)) = 0 \text{ für } j'' \neq j') + \frac{1}{h} R_h \\ &= \frac{1}{h} \cdot e^{-h/N} \frac{h/N}{1!} \cdot (e^{-h/N})^{N-2} + R_h = \frac{1}{N} + o(1) \end{aligned}$$

für $h \downarrow 0$, wobei der Resterm

$$|R_h| \leq \mathbb{P}\left(\sum_{k \neq \ell}^N N^{(k,\ell)}([t-s-h, t-s)) \geq 2\right) = O(h^2)$$

erfüllt. □

Bemerkung. Eine zeitdiskrete Version wurde in u.a. [B16] und [D19] betrachtet, dies war ein (zeitdiskretes) Cannings-Modell, in dem in jeder „Generation“ genau ein Individuum zwei Kinder hat und dafür eines null. Insoweit entspricht eine Generation dort einem Reproduktionsereignis im Moran-Modell, wenn man für die Individuen mit genau einem Kind jeweils Elter und Kind identifiziert.

Wenn $\frac{1}{N}X_0^{(N)} \rightarrow x_0$ gilt, so konvergiert der reskalierte Anteilsprozess

$$(Z_t^{(N)})_{t \geq 0} := \left(\frac{1}{N}X_{Nt/2}^{(N)}\right)_{t \geq 0} \quad (2.9)$$

gegen die (neutrale 2 Typ-)Wright-Fisher-Diffusion (vgl. Def. 2.6 und Bem. 2.7 unten), d.h.

$$(Z_t^{(N)})_{t \geq 0} \xrightarrow[N \rightarrow \infty]{d} (Z_t)_{t \geq 0} \quad (2.10)$$

(Konvergenz in Verteilung auf dem Raum der càdlàg-Pfade, vgl. Anhang A.2). Wir werden dies unten beweisen (siehe Abschnitte 2.3.4 und 2.6.1); für eine entscheidende Heuristik beachte

$$\frac{1}{h} \mathbb{E} \left[\left(\frac{1}{N}X_{N(t+h)/2}^{(N)} - \frac{1}{N}X_{Nt/2}^{(N)} \right)^2 \mid X_{Nt/2}^{(N)} = x \right] = \frac{1}{hN^2} \left(2 \frac{N}{2} h \frac{x(N-x)}{N} + o\left(\frac{N}{2}h\right) \right) = \frac{x}{N} \left(1 - \frac{x}{N} \right) + o(1)$$

für $h \downarrow 0$.

2.2 Kingmans Koaleszent

Sei $\mathcal{E}_n := \{\text{Äquivalenzrelationen auf } [n]\}$, wir notieren $\xi \in \mathcal{E}_n$ etwa durch eine (ungeordnete) Liste der Äquivalenzklassen (z.B. $\xi = \{\{1\}, \{2, 3\}\} \in \mathcal{E}_3$ bedeutet $2 \sim_\xi 3$, $1 \not\sim_\xi 2$, $1 \not\sim_\xi 3$).

Wir schreiben $\xi \leq \eta$, falls

$$i \sim_\xi j \implies i \sim_\eta j \quad \text{gilt,}$$

d.h. η entsteht aus ξ durch Vereinigung einiger Klassen, ggfs. in mehreren Gruppen (z.B. $\{\{1\}, \{2\}, \{3, 4\}, \{5\}, \{6, 7\}, \{8\}\} \leq \{\{1, 2, 6, 7\}, \{3, 4, 5\}, \{8\}\}$).

Definition 2.4. Die zeitkontinuierliche Markovkette $(R_t^{(n)})_{t \geq 0}$ auf \mathcal{E}_n mit Sprungratenmatrix

$$q_{\xi\eta} = \begin{cases} 1 & \text{falls } \eta \text{ aus } \xi \text{ durch Verschmelzung von genau zwei Klassen entsteht,} \\ -\binom{|\xi|}{2} & \text{falls } \eta = \xi, \\ 0 & \text{sonst} \end{cases} \quad (2.11)$$

heißt Kingmans (n -)Koaleszent.

Zumeist betrachten wir den Startzustand $R_0^{(n)} = \{\{1\}, \{2\}, \dots, \{n\}\}$. Wir können den Pfad $(R_t^{(n)})_{t \geq 0}$ als Baum interpretieren, dessen Blätter mit $1, \dots, n$ markiert sind:

Zu den Zeitpunkten $0 = \tau_n^{(n)} < \tau_{n-1}^{(n)} < \dots < \tau_2^{(n)} < \tau_1^{(n)}$, wo $\tau_k^{(n)} := \inf\{t \geq 0 : |R_t^{(n)}| \leq k\}$, verschmelzen jeweils zwei Zweige. Für Stichproben $i, j \in [n]$ können wir den genealogischen Abstand von i und j , $\inf\{t \geq 0 : i \sim_{R_t^{(n)}} j\}$ aus dem Baum ablesen.

vgl. Bild an der Tafel

2.3 Zur Wright-Fisher-Diffusion

Bericht 2.5 (Diffusionsprozesse auf \mathbb{R}). Ein starker Markovprozess $(X_t)_{t \geq 0}$ auf \mathbb{R} mit stetigen Pfaden heißt ein Diffusionsprozess. Die Dynamik eines solchen Prozesses kann man durch seine „infinitesimalen Charakteristiken“ beschreiben:

$$\mathbb{E}[X_{t+h} - X_t | \mathcal{F}_t] = h\mu(X_t) + o(h), \quad \text{Var}[X_{t+h} - X_t | \mathcal{F}_t] = h\sigma^2(X_t) + o(h)$$

für $h \downarrow 0$ mit gewissen Funktionen μ und σ^2 ($\mathcal{F}_t = \sigma(X_s, s \leq t)$ beschreibt die durch Beobachtung des Pfads bis t verfügbare Information). Dann ist (für eine genügend glatte Funktion f auf \mathbb{R})

$$Lf(x) = \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}[f(X_{t+h}) - f(X_t) | X_t = x] = \mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)$$

der sogenannte Generator von X .

Das „kanonischste“ Objekt dieser Klasse, die Brownsche Bewegung $(B_t)_{t \geq 0}$, erfüllt dies mit $\mu(\cdot) \equiv 0$, $\sigma^2(\cdot) \equiv 1$. Man kann zeigen, dass man ein allgemeines X (unter geeigneten Bedingungen an μ und σ^2) als Lösung einer sogenannten stochastischen Differentialgleichung

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t$$

gewinnen kann.

Lesenswerte Einführungen zu diesem Thema finden sich beispielsweise bei Breiman [B, Ch. 16] und bei Kersting und Wakolbinger [KW] [siehe Kap. 3 für die Brownsche Bewegung, Kap. 5 für Markovprozesse und Generatoren].

Definition 2.6. Der starke Markovprozess $Z = (Z_t)_{t \geq 0}$ mit Werten in $[0, 1]$ und stetigen Pfaden, dessen Generator für $f \in C^2([0, 1])$ gegeben ist durch

$$Lf(z) := \lim_{h \downarrow 0} \frac{1}{h} \mathbb{E}_z[f(Z_h) - f(Z_0)] = \frac{1}{2}z(1-z)f''(z), \quad z \in [0, 1]$$

heißt die (neutrale 2-Typ) *Wright-Fisher-Diffusion*.

(Siehe (2.31) unten für die Übergangskerne $\kappa_t(x, \cdot) = \mathcal{L}(Z_t | Z_0 = 0)$ von Z .)

Bemerkung 2.7. 1. Die Wright-Fisher-Diffusion Z ist die eindeutige Lösung des folgenden wohlgestellten Martingalproblems: Für jedes $f \in C^2([0, 1])$ ist der Prozess

$$M_t(f) := f(Z_t) - f(Z_0) - \int_0^t \frac{1}{2}Z_s(1-Z_s)f''(Z_s) ds \quad (2.12)$$

ein stetiges Martingal (bzgl. seiner kanonischen Filtration), wobei $M_0(f) = 0$.

2. Äquivalent könnten wir die Eigenschaft

$$Z \text{ stetiges Martingal mit Werten in } [0, 1] \text{ und } \langle Z \rangle_t = \int_0^t Z_s(1-Z_s) ds \quad (2.13)$$

betrachten.

Die Itô-Formel zeigt, dass Z dann das Martingalproblem mit Generator

$$Lf(z) = \frac{1}{2}z(1-z)f''(z) \quad \text{für } f \in C^2([0,1])$$

löst, denn

$$\begin{aligned} f(Z_t) - f(Z_0) &= \int_0^t f'(Z_s) dZ_s + \frac{1}{2} \int_0^t f''(Z_s) d\langle Z \rangle_s \\ &= \int_0^t f'(Z_s) dZ_s + \frac{1}{2} \int_0^t Z_s(1-Z_s)f''(Z_s) ds, \end{aligned}$$

d.h. $f(Z_t) - f(Z_0) - \int_0^t \frac{1}{2}Z_s(1-Z_s)f''(Z_s) ds = \int_0^t f'(Z_s) dZ_s$ ist ein Martingal.

3. Ein weiterer Blickpunkt: Z ist die (eindeutige, starke) Lösung der stochastischen Differentialgleichung

$$dZ_t = \sqrt{Z_t(1-Z_t)} dB_t,$$

wo (B_t) standard-Brownsche Bewegung.

Das Martingalproblem (2.13) ist eindeutig (via Dualität mit dem Klassenzählprozess des Kingman-Koaleszenten):

Wir zeigen: Sei $Z = (Z_t)_{t \geq 0}$ stetiges Martingal mit Werten in $[0,1]$, $Z_0 = z$ und $\langle Z \rangle_t = \int_0^t Z_s(1-Z_s) ds$ [d.h. $Z_t^2 - Z_0^2 - \int_0^t Z_s(1-Z_s) ds$ ist Martingal], so ist Z die Wright-Fisher-Diffusion.

Mit Itô-Formel ist für $n \in \mathbb{N}$

$$Z_t^n = Z_0^n + \int_0^t nZ_s^{n-1} dZ_s + \frac{1}{2} \int_0^t n(n-1)Z_s^{n-2} d\langle Z \rangle_s,$$

also ist

$$Z_t^n - Z_0^n - \frac{1}{2} \int_0^t n(n-1)(Z_s^{n-1} - Z_s^n) ds = \int_0^t nZ_s^{n-1} dZ_s$$

ein Martingal und somit

$$\begin{aligned} f(n, z, t) &:= \mathbb{E}_z[Z_t^n] = z^n + \binom{n}{2} \int_0^t \mathbb{E}_z[Z_t^{n-1}] - \mathbb{E}_z[Z_t^n] ds \\ &= z^n + \binom{n}{2} \int_0^t f(n-1, z, t) - f(n, z, t) ds \end{aligned}$$

(und $f(1, z, t) = z$ für alle $t \geq 0$, $z \in [0,1]$), d.h. die Familie von Funktionen $f(n, z, t)$ löst ein (lineares) System von Differentialgleichungen

$$\frac{\partial}{\partial t} f(n, z, t) = \binom{n}{2} (f(n-1, z, t) - f(n, z, t)), \quad t \geq 0, n \in \mathbb{N} \quad \text{und} \quad f(n, z, 0) = z^n. \quad (2.14)$$

Dieses System ist eindeutig lösbar (z.B. rekursiv in n).

Sei $(N_t)_{t \geq 0}$ ($N_t = |R_t^{(n)}|$, wenn der Startzustand n Klassen hat) der Blockzählprozess (oder Klassenzählprozess) des Kingman-Koaleszenten (Def. 2.4), wir lesen dort ab, dass $(N_t)_{t \geq 0}$ eine Markovkette auf \mathbb{N} mit Sprungratenmatrix

$$q_{m,m-1} = \binom{m}{2}, \quad q_{m,m} = -\binom{m}{2}, \quad q_{m,\ell} = 0 \text{ falls } \ell \notin \{m-1, m\}$$

ist. Setze

$$g(n, z, t) := \mathbb{E}_n[z^{N_t}]$$

(Satz 2.8 unten beschreibt die Verteilung von N_t , damit lässt sich – wenigstens im Prinzip – die Erzeugendenfunktion von N_t bestimmen).

Gemäß Kolmogorovs Rückwärtsgleichung (vgl. auch Anhang A.1.1) löst $g(n, z, t)$ ebenfalls (2.14):

$$\begin{aligned} \frac{1}{h} \left(\mathbb{E}_n[z^{N_{t+h}}] - \mathbb{E}_n[z^{N_t}] \right) &= \frac{1}{h} \left(\mathbb{E}_n[\mathbb{E}_n[z^{N_{t+h}} | N_h]] - \mathbb{E}_n[z^{N_t}] \right) \\ &= \frac{1}{h} \left(\left(1 - h \binom{n}{2}\right) \mathbb{E}_n[z^{N_t}] + h \binom{n}{2} \mathbb{E}_{n-1}[z^{N_t}] + O(h^2) - \mathbb{E}_n[z^{N_t}] \right) = \binom{n}{2} (\mathbb{E}_{n-1}[z^{N_t}] - \mathbb{E}_n[z^{N_t}]) + O(h) \end{aligned}$$

Da dieses System eindeutig lösbar ist, folgt

$$\mathbb{E}_z[Z_t^n] = \mathbb{E}_n[z^{N_t}], \quad t \geq 0, z \in [0, 1], n \in \mathbb{N}, \quad (2.15)$$

d.h. alle Momente von Z_t , und damit (da Z_t beschränkt ist) auch die Verteilung von Z_t ist festgelegt.

Für die Lösung eines Martingalproblems impliziert Eindeutigkeit der eindimensionalen Verteilungen die Eindeutigkeit der Verteilung des gesamten Prozesses (vgl. z.B. [EK, Thm. 4.4.2] oder [B18, Prop. A.3]).

Bemerkung (Dualität, allgemeiner Fall). Seien $X^{(x)} = (X_t^{(x)})_{t \geq 0}$, $x \in E$ und $Y^{(y)} = (Y_t^{(y)})_{t \geq 0}$, $y \in E'$ Familien von stochastischen Prozessen mit Werten in E bzw. in E' (E , E' seien polnische Räume, sagen wir), es gelte $X_0^{(x)} = x$, $Y_0^{(y)} = y$ f.s. X und Y heißen *dual* mit *Dualitätsfunktion* $H : E \times E' \rightarrow \mathbb{C}$, wenn gilt

$$\mathbb{E}[H(X_t^{(x)}, y)] = \mathbb{E}[H(x, Y_t^{(y)})] \quad \forall t \geq 0, x \in E, y \in E'.$$

(Wir nehmen an, dass H geeignete Messbarkeits- und Beschränktheits/-Wachstumsannahmen erfüllt, so dass die betrachteten Erwartungswerte existieren.)

Insbesondere: Es gebe eine Familie $Y^{(y)}$, $y \in E'$ von Markovprozessen mit Werten in E' und $Y_0^{(y)} = y$ (wie oben) und $H : [0, 1] \times E' \rightarrow \mathbb{C}$ m.b. so dass $\mathbb{E}[|H(x, Y_t^{(y)})|] < \infty$ für $x \in \mathbb{R}$, $y \in E'$, $t \geq 0$ und die Funktionenmenge

$$\{H(\cdot, y), y \in E'\} \quad \text{sei trennend für } \mathcal{M}_1(E),$$

(d.h. $\int_E H(x, y) \mu(dx) = \int_E H(x, y) \nu(dx)$ für alle $y \Rightarrow \mu = \nu$). Weiter gebe es für jedes $x \in [0, 1]$ eine Lösung $X^{(x)}$ des Martingalproblems mit Generator L und Startwert $X_0^{(x)} = x$, die

$$\mathbb{E}[H(X_t^{(x)}, y)] = \mathbb{E}[H(x, Y_t^{(y)})] \quad \forall t \geq 0, y \in E' \quad (2.16)$$

erfüllt. Dann ist das Martingalproblems zum Generator L wohlgestellt, d.h. die Lösung ist eindeutig. (Siehe z.B. [B18, Def. A.6 und Satz A.7] oder [Kl, Kap. 26.3].)

2.3.1 Übergangswahrscheinlichkeiten des (Blockzählprozesses des) Kingman-Koaleszenten und der Wright-Fisher-Diffusion

Sei $(N_t)_{t \geq 0}$ der Blockzählprozess (oder Klassenzählprozess) des Kingman-Koaleszenten, d.h. die zeitkontinuierliche Markovkette auf \mathbb{N} mit Sprungraten

$$q_{m,m-1} = \binom{m}{2}, \quad q_{m,m} = -\binom{m}{2}, \quad q_{m,\ell} = 0 \text{ falls } \ell \notin \{m-1, m\}$$

(in der Notation von Def. 2.4 ist $N_t = |R_t^{(n)}|$, wenn $N_0 = n$).

Satz 2.8. Für $n \in \mathbb{N}$ gilt

$$\mathbb{P}_n(N_t = m) = \sum_{j=m}^n \exp\left(-\binom{j}{2}t\right) \frac{(2j-1)(-1)^{j-m} \binom{m}{(j-1)\uparrow} \binom{n}{j\downarrow}}{m!(j-m)!(n)_{j\uparrow}}, \quad 2 \leq m \leq n, \quad (2.17)$$

$$\mathbb{P}_n(N_t = 1) = 1 - \sum_{j=2}^n \exp\left(-\binom{j}{2}t\right) \frac{(2j-1)(-1)^j \binom{n}{j\downarrow}}{(n)_{j\uparrow}} \quad (2.18)$$

(mit $(x)_{j\downarrow} = x(x-1)\cdots(x-j+1)$, $(x)_{j\uparrow} = x(x+1)\cdots(x+j-1)$ fallende bzw. wachsende Faktorielle).

Lemma 2.9. Seien X_1, X_2, \dots, X_n unabhängig, X_i sei exponentialverteilt mit Parameter λ_i und $\lambda_1 < \lambda_2 < \dots < \lambda_n$. Dann hat $X := X_1 + \dots + X_n$ die Dichte

$$f_X(t) = \sum_{j=1}^n a_j \lambda_j \exp(-\lambda_j t), \quad t \geq 0 \quad \text{mit } a_j := \prod_{k=1, k \neq j}^n \frac{\lambda_k}{\lambda_k - \lambda_j}$$

insbesondere ist

$$\mathbb{P}(X > t) = \sum_{j=1}^n a_j \exp(-\lambda_j t), \quad t \geq 0.$$

Beweisskizze. Die Formel für die Dichte kann man beispielsweise per Induktion durch sukzessive Faltung mit der Exponentialdichte beweisen, für den Induktionsschritt beachten wir

$$\begin{aligned} & \int_0^t \sum_{j=1}^{n-1} \lambda_j \exp(-\lambda_j s) \prod_{k=1, k \neq j}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_j} \times \lambda_n \exp(-\lambda_n(t-s)) ds \\ &= \sum_{j=1}^{n-1} \lambda_j \lambda_n \prod_{k=1, k \neq j}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_j} \times e^{-\lambda_n t} \int_0^t e^{(\lambda_n - \lambda_j)s} ds = \sum_{j=1}^{n-1} \lambda_j \lambda_n \prod_{k=1, k \neq j}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_j} \times \frac{e^{-\lambda_j t} - e^{-\lambda_n t}}{\lambda_n - \lambda_j} \\ &= \sum_{j=1}^{n-1} \lambda_j e^{-\lambda_j t} \prod_{k=1, k \neq j}^n \frac{\lambda_k}{\lambda_k - \lambda_j} - \lambda_n e^{-\lambda_n t} \sum_{j=1}^{n-1} \frac{\lambda_j}{\lambda_n - \lambda_j} \prod_{k=1, k \neq j}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_j}. \end{aligned}$$

Dann verwenden wir die Identität

$$\sum_{j=1}^{n-1} \frac{\lambda_j}{\lambda_n - \lambda_j} \prod_{k=1, k \neq j}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_j} = - \prod_{k=1}^{n-1} \frac{\lambda_k}{\lambda_k - \lambda_n},$$

die (man dividiere beide Seiten durch $\lambda_1\lambda_2\cdots\lambda_{n-1}$ und sortiere Terme) äquivalent ist zu

$$\sum_{j=1}^n \prod_{k=1, k \neq j}^n \frac{1}{\lambda_k - \lambda_j} = 0. \quad (2.19)$$

Sei $\ell_j(x) := \prod_{k=1, k \neq j}^n \frac{x - \lambda_k}{\lambda_j - \lambda_k}$ (das j -te Lagrange-Polynom zu $\lambda_1, \dots, \lambda_n$), $\ell_1(x) + \ell_2(x) + \dots + \ell_n(x)$ ist ein Polynom in x vom Grad $n - 1$, das (mindestens) an den n verschiedenen Stellen $\lambda_1, \dots, \lambda_n$ den Wert 1 annimmt (denn $\ell_j(\lambda_i) = \delta_{ji}$), daher gilt $\ell_1(x) + \ell_2(x) + \dots + \ell_n(x) \equiv 1$. Die linke Seite von (2.19) ist $(-1)^{n-1}$ mal der Koeffizient von x^{n-1} in diesem Polynom.

Alternativ beachte man, dass für $\zeta \in \mathbb{R}$ gilt $\mathbb{E}[e^{i\zeta X_j}] = \int_0^\infty e^{i\zeta x} \lambda_j e^{-\lambda_j x} dx = \frac{\lambda_j}{\lambda_j - i\zeta}$, also $\mathbb{E}[e^{i\zeta X}] = \prod_{j=1}^n \frac{\lambda_j}{\lambda_j - i\zeta} =: \varphi_1(\zeta)$ gilt, während $\int_0^\infty e^{i\zeta x} \sum_{j=1}^n a_j \lambda_j e^{-\lambda_j x} dx = \sum_{j=1}^n \frac{a_j \lambda_j}{\lambda_j - i\zeta} =: \varphi_2(z)$ und es ist $\varphi_2 = \varphi_1$ (φ_2 ist die Partialbruchzerlegungs-Darstellung von φ_1), denn beide sind meromorph auf \mathbb{C} mit jeweils einfachen Polen bei $\zeta = -i\lambda_1, \dots, -i\lambda_n$ und $\lim_{|z| \rightarrow \infty} \varphi_{1/2}(z) = 0$, $\lim_{z \rightarrow -i\lambda_j} \frac{\varphi_1(z)}{\lambda_j - iz} = \lambda_j \prod_{k=1, k \neq j}^n \frac{\lambda_k}{\lambda_k - \lambda_j} = \lim_{z \rightarrow -i\lambda_j} \frac{\varphi_2(z)}{\lambda_j - iz}$.

Die Formel für den Verteilungsschwanz von X ergibt sich durch entsprechendes Integrieren der Dichte. \square

Beweis von Satz 2.8. Für $1 \leq m < n$ ist mit Lemma 2.9 (und S_j u.a., $S_j \sim \text{Exp}(\binom{j}{2})$)

$$\begin{aligned} \mathbb{P}_n(N_t \leq m) &= \mathbb{P}(S_n + S_{n-1} + \dots + S_{m+1} \leq t) \\ &= \int_0^t \sum_{j=m+1}^n \binom{j}{2} \exp(-\binom{j}{2}s) \prod_{\substack{k=m+1, \\ k \neq j}}^n \frac{\binom{k}{2}}{\binom{k}{2} - \binom{j}{2}} ds \\ &= \sum_{j=m+1}^n \left(1 - \exp(-\binom{j}{2}t)\right) \prod_{\substack{k=m+1, \\ k \neq j}}^n \frac{\binom{k}{2}}{\binom{k}{2} - \binom{j}{2}} \end{aligned}$$

(und $\mathbb{P}_n(N_t \leq n) = 1$), somit für $1 < m < n$

$$\begin{aligned} \mathbb{P}_n(N_t = m) &= \mathbb{P}_n(N_t \leq m) - \mathbb{P}_n(N_t \leq m-1) \\ &= \sum_{j=m+1}^n \left(1 - \exp(-\binom{j}{2}t)\right) \left(1 - \frac{\binom{m}{2}}{\binom{m}{2} - \binom{j}{2}}\right) \prod_{\substack{k=m+1, \\ k \neq j}}^n \frac{\binom{k}{2}}{\binom{k}{2} - \binom{j}{2}} \\ &\quad - \left(1 - \exp(-\binom{m}{2}t)\right) \prod_{k=m+1}^n \frac{\binom{k}{2}}{\binom{k}{2} - \binom{m}{2}} \\ &= - \sum_{j=m}^n \left(1 - \exp(-\binom{j}{2}t)\right) \frac{\binom{j}{2}}{\binom{m}{2}} \prod_{\substack{k=m, \\ k \neq j}}^n \frac{\binom{k}{2}}{\binom{k}{2} - \binom{j}{2}} \\ &= - \frac{\prod_{\ell=m}^n \binom{\ell}{2}}{\binom{m}{2}} \sum_{j=m}^n \left(1 - \exp(-\binom{j}{2}t)\right) \prod_{\substack{k=m, \\ k \neq j}}^n \frac{1}{\binom{k}{2} - \binom{j}{2}} \\ &= \frac{\prod_{\ell=m}^n \binom{\ell}{2}}{\binom{m}{2}} \sum_{j=m}^n \exp(-\binom{j}{2}t) \prod_{\substack{k=m, \\ k \neq j}}^n \frac{1}{\binom{k}{2} - \binom{j}{2}} \end{aligned}$$

(in der letzten Zeile nutzen wir aus, dass $\sum_{j=m}^n \prod_{k=m, k \neq j}^n \frac{1}{\binom{k}{2} - \binom{j}{2}} = 0$ gilt, vgl. (2.19) in Lemma 2.9).

Weiter ist

$$\prod_{\ell=m}^n \binom{\ell}{2} = \prod_{\ell=m}^n \frac{\ell(\ell-1)}{2} = 2^{m-n} \frac{n!(n-1)!}{(m-1)!(m-2)!}$$

und

$$\begin{aligned} \prod_{\substack{k=m, \\ k \neq j}}^n \left(\binom{k}{2} - \binom{j}{2} \right) &= \prod_{\substack{k=m, \\ k \neq j}}^n \frac{k(k-1) - j(j-1)}{2} = \prod_{\substack{k=m, \\ k \neq j}}^n \frac{(k+j-1)(k-j)}{2} \\ &= 2^{-(n-m)} \frac{(n+j-1)!}{(2j-1)(m+j-2)!} (-1)^{j-m} (j-m)!(n-j)!. \end{aligned}$$

Insgesamt folgt

$$\begin{aligned} \mathbb{P}_n(N_t = m) &= \sum_{j=m}^n \exp\left(-\binom{j}{2}t\right) \frac{(2j-1)(-1)^{j-m} (m)_{(j-1)\uparrow} (n)_{j\downarrow}}{m!(j-m)!(n)_{j\uparrow}}, \quad 2 \leq m \leq n, \\ \mathbb{P}_n(N_t = 1) &= 1 - \sum_{j=2}^n \exp\left(-\binom{j}{2}t\right) \frac{(2j-1)(-1)^j (n)_{j\downarrow}}{(n)_{j\uparrow}} \end{aligned}$$

(mit $(x)_{j\downarrow} = x(x-1)\cdots(x-j+1)$, $(x)_{j\uparrow} = x(x+1)\cdots(x+j-1)$ fallende bzw. wachsende Faktorielle).

□

Bemerkung 2.10. Alternativ diagonalisiere die Sprungratenmatrix des Klassenzählprozesses des Kingman-Koaleszenten, vgl. [T84, Appendix I]: $\lambda_k = -k(k-1)/2$, $k = 1, \dots, n$ Eigenwerte,

$$\ell_j^{(k)} = \binom{k}{j} (-1)^{k-j} \frac{(j)_{(k-1)\downarrow}}{(k)_{(k-1)\downarrow}} \mathbb{1}(j \leq k),$$

$(\ell_1^{(k)}, \dots, \ell_n^{(k)})$ Links-Eigenvektor zum Eigenwert λ_k ,

$$r_j^{(k)} = \binom{j}{k} \frac{(k)_{k\downarrow}}{(j)_{k\downarrow}} \mathbb{1}(j \geq k),$$

$(r_1^{(k)}, \dots, r_n^{(k)})^T$ Rechts-Eigenvektor.

Bemerkung 2.11. Aus der Struktur der Sprungratenmatrix (2.11) folgt

$$\tau_1^{(n)} = (\tau_{n-1}^{(n)} - \tau_n^{(n)}) + (\tau_{n-2}^{(n)} - \tau_{n-1}^{(n)}) + \cdots + (\tau_1^{(n)} - \tau_2^{(n)}) \stackrel{d}{=} S_n + S_{n-1} + \cdots + S_2, \quad n \geq 2,$$

wobei die S_k unabhängige exponentialverteilte Zufallsvariablen mit Parameter $\binom{k}{2}$ sind, somit

$$\mathbb{E}[\tau_1^{(n)}] = \sum_{k=2}^n E[S_k] = \sum_{k=2}^n \frac{2}{k(k-1)} = 2 \sum_{k=2}^n \left(\frac{1}{k-1} - \frac{1}{k} \right) = 2 \left(1 - \frac{1}{n} \right)$$

und $1 = \mathbb{E}[\tau_1^{(2)}] \leq \mathbb{E}[\tau_1^{(n)}] < \lim_{m \rightarrow \infty} \mathbb{E}[\tau_1^{(m)}] = 2$.

Für das Moran-Modell mit Populationsgröße N bedeutet dies, dass der jüngste gemeinsame Vorfahre der heute lebenden Population im Mittel vor etwa $2\frac{N}{2} = N$ Zeiteinheiten gelebt hat.

Weiter ist

$$\begin{aligned} \text{Var}[\tau_1^{(n)}] &= \sum_{k=2}^n \text{Var}[S_k] = \sum_{k=2}^n \binom{k}{2}^{-2} = \sum_{k=2}^n \frac{4}{k^2(k-1)^2} = \sum_{k=2}^n \left\{ 4\left(\frac{1}{k^2} + \frac{1}{(k-1)^2}\right) + 8\left(\frac{1}{k} - \frac{1}{k-1}\right) \right\} \\ &= \left(8 \sum_{k=1}^{n-1} \frac{1}{k^2} \right) - 4 + \frac{4}{n^2} + \frac{8}{n} - 8 = \left(8 \sum_{k=1}^{n-1} \frac{1}{k^2} \right) - 4\left(1 - \frac{1}{n}\right)\left(3 + \frac{1}{n}\right), \end{aligned}$$

insbesondere

$$1 = \text{Var}[\tau_1^{(2)}] \leq \text{Var}[\tau_1^{(n)}] < \lim_{n \rightarrow \infty} \text{Var}[\tau_1^{(n)}] = 8\frac{\pi^2}{6} - 12 \approx 1.16.$$

Der wesentliche Beitrag zur Gesamtvarianz kommt also von der letzten Verschmelzungszeit S_2 .

Mit $n \rightarrow \infty$ ergibt sich aus (2.17)–(2.18) (vgl. auch den folgenden Abschnitt 2.3.2 zum Start in $N_0 = \infty$)

$$\mathbb{P}_\infty(N_t = m) = \sum_{k=m}^{\infty} \exp\left(-\binom{k}{2}t\right) \frac{(2k-1)(-1)^{k-m} \binom{m}{k-1} \uparrow}{m!(k-m)!}, \quad m \geq 2, \quad (2.20)$$

$$\mathbb{P}_\infty(N_t = 1) = 1 - \sum_{k=2}^{\infty} \exp\left(-\binom{k}{2}t\right) (2k-1)(-1)^k. \quad (2.21)$$

2.3.2 Ergänzung: Der „ ∞ -Koaleszent“

Man kann dem Kingman-Koaleszenten mit $n = \infty$ einen Sinn geben:

1. (Teilstichproben-Konsistenz) Sei $\pi_{n,n-1} : \mathcal{E}_n \rightarrow \mathcal{E}_{n-1}$ die Einschränkung aller Äquivalenzklassen auf $[n-1]$, so gilt

$$\left(\pi_{n,n-1}(R_t^{(n)})\right)_{t \geq 0} \stackrel{d}{=} \left(R_t^{(n-1)}\right)_{t \geq 0}.$$

Dies folgt aus der Form der Sprungraten.

2. (Invarianz der Verteilung unter Permutation der Stichprobennummern) Für eine Permutation σ von $[n]$ und $\xi = \{C_1, \dots, C_a\} \in \mathcal{E}_n$ sei $\sigma(\xi) = \{\sigma(C_1), \dots, \sigma(C_a)\}$ die Äquivalenzrelation, die man erhält, indem man die Elemente der Blöcke von ξ gemäß σ umnummeriert. Es gilt

$$\left(\sigma(R_t^{(n)})\right)_{t \geq 0} \stackrel{d}{=} \left(R_t^{(n)}\right)_{t \geq 0}.$$

Dies folgt aus der Symmetrie der Sprungraten. [Man sagt auch, dass $R_t^{(n)}$ eine austauschbare zufällige Äquivalenzrelation ist.]

3. (unendlich große Stichprobe) Mit obigem und Kolmogorovs Erweiterungssatz ist es möglich, den Kingman-Koaleszenten $(R_t)_{t \geq 0}$ mit Stichprobengröße $n = \infty$ als Markovprozess auf $\mathcal{E} := \{\text{Äquivalenzrelationen auf } \mathbb{N}\}$ mit Startwert $R_0 = \{\{1\}, \{2\}, \dots\}$ zu definieren mit der Eigenschaft $(\pi_{\infty, n}(R_t))_{t \geq 0} \stackrel{d}{=} (\sigma(R_t^{(n)}))_{t \geq 0}$ für jedes $n \in \mathbb{N}$.

Bem. 2.11 zeigt, dass $\mathbb{E}[\tau_1^{(\infty)}] = 2 < \infty$, d.h. auch eine „unendlich große Stichprobe“ findet f.s. in endlicher Zeit ihren ersten gemeinsamen Vorfahren. Obwohl $|R_0| = \infty$ ist, gilt $|R_t| < \infty$ für jedes $t > 0$ fast sicher. [Man sagt auch, dass der Kingman-Koaleszent „aus dem Unendlichen herabsteigt.“]

2.3.3 Ergänzung: Die Skelettkette des Kingman-Koaleszenten

Sei $\xi_i^{(n)}$, $i = n, n-1, \dots, 1$ der Zustand des n -Koaleszenten zum ersten Zeitpunkt, zu dem i Klassen existieren, d.h. $\xi_i^{(n)} = R_{\tau_i^{(n)}}^{(n)}$ (mit $\tau_i^{(n)} := \inf\{t \geq 0: |R_t^{(n)}| \leq i\}$ wie oben). $[(\xi_i^{(n)})_{i=n, n-1, \dots, 1}]$ heißt die *Skelettkette* des Kingman- n -Koaleszenten, sie ist (offenbar) eine Markovkette.]

Proposition 2.12. Für $\xi \in \mathcal{E}_n$ mit i Klassen der Größen $\lambda_1, \dots, \lambda_i \in \mathbb{N}$ (mit $\lambda_1 + \dots + \lambda_i = n$) gilt

$$\mathbb{P}(\xi_i^{(n)} = \xi) = c_{n,i} w(\xi) \quad \text{mit } w(\xi) = \lambda_1! \dots \lambda_i!, \quad c_{n,i} = \frac{i! (n-i)! (i-1)!}{n! (n-1)!}. \quad (2.22)$$

Beispiel. Betrachte $n = 9$, $i = 3$, es ist $c_{9,3} = \frac{3! 6! 2!}{9! 8!} = 1/1\,693\,440$.

λ_i	3-3-3	4-3-2	5-2-2	4-4-1	5-3-1	6-2-1	7-1-1
w	216	288	480	576	720	1440	5040

Wir sehen: die Verteilung hat mehr Gewicht auf „unbalanzierten Aufteilungen.“

Beweis von Prop. 2.12. Rückwärtsinduktion über i : Für $i = n$ gilt $\mathbb{P}(\xi_n^{(n)} = \{\{1\}, \dots, \{n\}\}) = 1$ mit $\lambda_1 = \dots = \lambda_n = 1$, und $c_{n,n} = w(1, \dots, 1) = 1$.

$i \rightarrow i-1$: Es ist

$$\mathbb{P}(\xi_{i-1}^{(n)} = \eta \mid \xi_i^{(n)} = \xi) = \begin{cases} \frac{1}{\binom{i}{2}}, & \text{falls } \eta \text{ aus } \xi \text{ durch Verschmelzung eines Paares von} \\ & \text{Klassen entsteht,} \\ 0, & \text{sonst.} \end{cases}$$

Sei $\eta \in \mathcal{E}_n$, $|\eta| = i-1$, Klassengrößen $\tilde{\lambda}_1, \dots, \tilde{\lambda}_{i-1}$.

$$\begin{aligned} \mathbb{P}(\xi_{i-1}^{(n)} = \eta) &= \frac{2}{i(i-1)} \sum_{\xi: \xi < \eta} \mathbb{P}(\xi_i^{(n)} = \xi) \\ &= \frac{2}{i(i-1)} \sum_{\ell=1}^{i-1} \sum_{m=1}^{\tilde{\lambda}_\ell-1} \frac{1}{2} \binom{\tilde{\lambda}_\ell}{m} c_{n,i} \tilde{\lambda}_1! \dots \tilde{\lambda}_{\ell-1}! m! (\tilde{\lambda}_\ell - m)! \tilde{\lambda}_{\ell+1}! \dots \tilde{\lambda}_{i-1}! \\ &= \frac{c_{n,i}}{i(i-1)} \sum_{\ell=1}^{i-1} \sum_{m=1}^{\tilde{\lambda}_\ell-1} w(\eta) = \frac{c_{n,i} w(\eta)}{i(i-1)} \sum_{\ell=1}^{i-1} \sum_{m=1}^{\tilde{\lambda}_\ell-1} 1 \\ &= \frac{c_{n,i} w(\eta)}{i(i-1)} (n - (i-1)) = c_{n,i-1} w(\eta) \end{aligned}$$

Für das 2. Gleichheitszeichen verwenden wir die Induktionsannahme und zerlegen gemäß der „aufgespaltenen“ Klasse: die ℓ -te Klasse hat $\tilde{\lambda}_\ell$ Elemente, zerlege in 2 Teilmengen d. Größen m und $\tilde{\lambda}_\ell - m$, es gibt $\frac{1}{2} \binom{\tilde{\lambda}_\ell}{m}$ mögliche Wahlen; der Faktor $\frac{1}{2}$ entsteht, weil die Klassen in η als ungeordnet aufgefasst werden. \square

Korollar 2.13. 1. Sei σ eine uniform verteilte Permutation von $\{1, \dots, i\}$, u.a. von $\xi^{(n)}$, $M_i = |C_{i,\sigma(i)}^{(n)}|$ mit $\xi_i^{(n)} = \{C_{i,1}^{(n)}, \dots, C_{i,i}^{(n)}\}$. Dann ist

$$(M_1, \dots, M_i) \text{ uniform verteilt auf } \{(m_1, \dots, m_i) \in \mathbb{N}^i : m_1 + \dots + m_i = n\}.$$

2. Sei $\xi = \{A_1, \dots, A_{i-1}\} \in \mathcal{E}_n$ mit $|A_j| = \lambda_j$. $\mathcal{L}(\xi_i^{(n)} | \xi_{i-1}^{(n)} = \xi)$ kann folgendermaßen beschrieben werden:

- wähle A_j mit W'keit $\frac{\lambda_j - 1}{n - i + 1}$ ($j \in \{1, \dots, i - 1\}$), dann wähle k uniform aus $\{1, \dots, \lambda_j - 1\}$
- spalte A_j uniform in zwei Teile der Größen k und $\lambda_j - k$.

Bemerkung. Für $i = 2$ zeigt Kor. 2.13 die „uniforme Aufspaltung“ der Stichprobe in zwei älteste Familien.

Beweis von Kor. 2.13. 1. Seien m_1, \dots, m_i mit $m_1 + \dots + m_i = n$ gegeben. Nach Prop. 2.12 hat jede Realisierung des (zufällig) geordneten Vektors $(C_{i,\sigma(1)}^{(n)}, \dots, C_{i,\sigma(i)}^{(n)})$, die mit den geforderten Größen m_j verträglich ist, die W'keit $\frac{1}{i!} c_{n,i} m_1! \dots m_i!$, somit

$$\begin{aligned} \mathbb{P}((M_1, \dots, M_i) = (m_1, \dots, m_i)) &= \binom{n}{m_1 \dots m_i} \frac{1}{i!} c_{n,i} m_1! \dots m_i! \\ &= \frac{(n-i)!(i-1)!}{(n-1)!} = \frac{1}{\binom{n-1}{i-1}}, \end{aligned}$$

denn es gibt $\binom{n}{m_1 \dots m_i}$ Partitionen, die bezgl. der Größe der Klassen in Frage kommen, jede hat n. Prop. 2.12 dieselbe W'keit $c_{n,i} m_1! \dots m_i!$,

die W'keit, dass zuf. Perm. σ geg. Ordnung liefert, ergibt nochmals einen Faktor $\frac{1}{i!}$.

(Beachte auch $\#\{(m_1, \dots, m_i) \in \mathbb{N}^i : m_1 + \dots + m_i = n\} = \binom{n-1}{i-1}$: n Kugeln in i (nummerierte) Schachteln legen, so dass keine Schachtel leer ist: $n - 1$ mögl. Plätze für $i - 1$ „Trennwände“.)

2. ξ entstehe aus η durch Aufteilen von A_j in 2 Teile der Größen k und $\lambda_j - k$.

$$\begin{aligned} \mathbb{P}(\xi_i^{(n)} = \xi | \xi_{i-1}^{(n)} = \eta) &= \frac{\mathbb{P}(\xi_{i-1}^{(n)} = \eta | \xi_i^{(n)} = \xi) \mathbb{P}(\xi_i^{(n)} = \xi)}{\mathbb{P}(\xi_{i-1}^{(n)} = \eta)} \\ &= \frac{\frac{1}{\binom{i}{2}} c_{n,i} \lambda_1! \dots \lambda_{j-1}! k! (\lambda_j - k)! \lambda_{j+1}! \dots \lambda_{i-1}!}{c_{n,i-1} \lambda_1! \dots \lambda_{j-1}! \lambda_j! \lambda_{j+1}! \dots \lambda_{i-1}!} = \frac{1}{\binom{i}{2}} \cdot \frac{i(i-1)}{n-i+1} \frac{1}{\binom{\lambda_j}{k}} \\ &= \frac{\lambda_j - 1}{n - i + 1} \cdot \frac{1}{\lambda_j - 1} \cdot 2 \frac{1}{\binom{\lambda_j}{k}} \end{aligned}$$

($\frac{\lambda_j - 1}{n - i + 1} \hat{=}$ Wahl von A_j ; $\frac{1}{\lambda_j - 1} \hat{=}$ Wahl von k ; $2 \frac{1}{\binom{\lambda_j}{k}} \hat{=}$ Wahl der Zerlegung von A_j – beachte: Faktor 2, da die Klassen „ungeordnet“ angegeben werden) \square

Korollar 2.14. *Betrachte eine Teilstichprobe der Grösse n in einem Kingman- m -Koaleszenten, mit $m > n$. Dann erfüllt die Wahrscheinlichkeit des Ereignisses $E_{m,n}$, dass der jüngste gemeinsame Vorfahre (jgV) der n -Stichprobe mit der Wurzel des m -Koaleszenten übereinstimmt,*

$$\mathbb{P}(E_{m,n}) \rightarrow \frac{n-1}{n+1} \quad \text{für } m \rightarrow \infty.$$

Beweis. Wir betrachten $\xi_2^{(n)}$, die erste Aufspaltung des Koaleszenten von der Wurzel aus betrachtet. Diese resultiert in einer Aufspaltung der trivialen Partition $\{\{1, \dots, m\}\}$ in eine Äquivalenzrelation mit genau zwei Klassen der Grössen $m-X$ und X , wobei X nach Kor. 2.13 auf $[m-1]$ uniform verteilt ist. Falls der jgV der n -Stichprobe nicht mit der Wurzel übereinstimmt, so müssen die Ahnenlinien aller n Individuen der Stichprobe alle entweder in dem Block der Grösse $m-X$ oder in dem Block der Grösse X liegen.

Das erste Ereignis hat Wahrscheinlichkeit $\frac{(m-X)_{n\downarrow}}{(m)_{n\downarrow}}$ und das zweite $\frac{(X)_{n\downarrow}}{(m)_{n\downarrow}}$.

Wir erhalten

$$\begin{aligned} \mathbb{P}(E_{mn,}) &= 1 - \mathbb{P}((E_{mn,})^c) = 1 - \sum_{k=1}^{m-1} \left[\frac{(m-k)_{n\downarrow}}{(m)_{n\downarrow}} + \frac{(k)_{n\downarrow}}{(m)_{n\downarrow}} \right] \underbrace{\mathbb{P}(X=k)}_{\frac{1}{m-1}} \\ &\xrightarrow{m \rightarrow \infty} 1 - \int_0^1 (x^n + (1-x)^n) dx \\ &= 1 - \left[\frac{1}{n+1} x^{n+1} \right]_0^1 - \left[\frac{1}{n+1} (1-x)^{n+1} (-1) \right]_0^1 = 1 - \frac{2}{n+1} \end{aligned}$$

für $m \rightarrow \infty$ durch Konvergenz der Riemann-Summe gegen das Riemann-Integral. \square

2.3.4 Konvergenz des Typenanteilsprozesses

Ein Zugang zur Konvergenz des Typenanteils im Moran-Modell gegen die Wright-Fisher-Diffusion via Dualität mit dem Kingman-Koaleszenten

Ein Zugang zur Konvergenz von $(Z^{(N)})_{t \geq 0}$ in (2.10), der zumindest die Konvergenz der endlich-dimensionalen Verteilungen beweist und der ohne den „technischem Apparat“ der Martingalprobleme aus Bem. 2.7 auskommt, die Momente des Grenzprozesses via Stichproben-Interpretation und die Konvergenz der Genealogien (vgl. (2.7) und die Diskussion direkt danach) zu berechnen.

Beobachtung 2.15. Sei $x \in [0, 1]$, im Modell mit Populationsgröße N vergeben wir an die N Individuen der Startgeneration 0 rein zufällig $\lfloor Nx \rfloor$ -mal den Typ 1 und $(N - \lfloor Nx \rfloor)$ -mal den Typ 0, in der Notation (2.1) auf Def. 2.1:

$$((\xi_0(1), \xi_0(2), \dots, \xi_0(N))) \quad \text{ist eine zufällige Permutation von } \underbrace{(1, 1, \dots, 1)}_{\lfloor Nx \rfloor}, \underbrace{(0, \dots, 0)}_{N - \lfloor Nx \rfloor}.$$

Sei $t > 0$ und $n \in \mathbb{N}$ ($n < N$), wir ziehen zum Zeitpunkt $Nt/2$ eine Stichprobe von n Individuen aus der Population. Die Wahrscheinlichkeit, dass alle n den Typ 1 haben ist,

$$\mathbb{E}_{\lfloor \frac{Nx \rfloor}{N}} \left[Z_t^{(N)} \left(Z_t^{(N)} - \frac{1}{N} \right) \dots \left(Z_t^{(N)} - \frac{n-1}{N} \right) \right]$$

$$\begin{aligned}
&= \mathbb{E}_{\lfloor \frac{Nx}{N} \rfloor} \left[\frac{1}{N} X_{Nt/2}^{(N)} \left(\frac{1}{N} X_{Nt/2}^{(N)} - \frac{1}{N} \right) \cdots \left(\frac{1}{N} X_{Nt/2}^{(N)} - \frac{n-1}{N} \right) \right] \\
&= \mathbb{P}_{\lfloor \frac{Nx}{N} \rfloor} \left(\xi_{Nt/2}(J_1) = 1, \xi_{Nt/2}(J_2) = 1, \dots, \xi_{Nt/2}(J_n) = 1 \right) \\
&= \mathbb{P}_{\lfloor \frac{Nx}{N} \rfloor} \left(\xi_0(A_{Nt/2}^{(J_1, Nt/2)}) = 1, \xi_0(A_{Nt/2}^{(J_2, Nt/2)}) = 1, \dots, \xi_0(A_{Nt/2}^{(J_n, Nt/2)}) = 1 \right) \\
&= \mathbb{E} \left[\prod_{i=0}^{|\mathcal{R}_{Nt/2}^{(n, N)}|-1} \frac{\lfloor Nx \rfloor - i}{N - i} \right]
\end{aligned}$$

mit J_1, \dots, J_n einer rein zufälligen Stichprobe vom Umfang n aus $1, \dots, N$ (verwende für das 2. Gleichheitszeichen die Dualitätsformel (2.8) aus Bem. 2.3).

Da $(\mathcal{R}_{Ns/2}^{(n, N)})_{0 \leq s \leq t}$ ein Kingman-Koaleszent (mit „Laufzeit“ t) ist, folgt

$$\lim_{N \rightarrow \infty} \mathbb{E}_{\lfloor \frac{Nx}{N} \rfloor} [(Z_t^{(N)})^n] = \mathbb{E}[x^{|R_t^{(n)}|}]$$

Für $x \in [0, 1]$, $t > 0$ wird daher ein W 'maß $\kappa_t(x, dy)$ eindeutig (das Momentenproblem auf dem kompakten Intervall $[0, 1]$ ist eindeutig lösbar, vgl. z.B. [Kl, Kor. 15.32] oder [B17, Kor. 4.30] durch die Forderung

$$\int_{[0,1]} y^n \kappa_t(x, dy) = \mathbb{E}[x^{|R_t^{(n)}|}] \quad \text{für } n \in \mathbb{N} \quad (2.23)$$

festgelegt und es gilt

$$\mathcal{L}(Z_t^{(N)} | Z_0^{(N)} = \frac{1}{N} \lfloor Nx \rfloor) \xrightarrow{N \rightarrow \infty} \kappa_t(x, \cdot) \quad (2.24)$$

(Wir werden in Abschnitt 2.6.1 unten zeigen, dass $\kappa_t(x, dy) = \mathbb{P}_x(Z_t \in dy)$ mit $(Z_t)_{t \geq 0}$ aus Def. 2.6 und Bem. 2.7, d.h. κ_t ist der Übergangskern der Wright-Fisher-Diffusion.)

Somit ist

$$\mathbb{P}_z(Z_t = 1) = \lim_{n \rightarrow \infty} \mathbb{E}_z[Z_t^n] = \lim_{n \rightarrow \infty} \sum_{j=1}^n \mathbb{P}_n(N_t = j) z^j = \sum_{j=1}^{\infty} \mathbb{P}_{\infty}(N_t = j) z^j \quad (2.25)$$

und analog

$$\mathbb{P}_z(Z_t = 0) = \sum_{j=1}^{\infty} \mathbb{P}_{\infty}(N_t = j) (1 - z)^j. \quad (2.26)$$

Um aus (2.23) auch die Dichte im Inneren $(0, 1)$ des Einheitsintervalls zu bestimmen, benötigen wir noch einen weiteren Gedanken:

Für $a, b > 0$ sei

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} y^{a-1} (1-y)^{b-1}, \quad y \in (0, 1)$$

die Beta- (a, b) -Dichte, für $Y_{a,b} \sim \text{Beta}(a, b)$ ist

$$\mathbb{E}[Y_{a,b}] = a/(a+b) \quad \text{und} \quad \text{Var}[Y_{a,b}] = ab/((a+b)^2(a+b+1)).$$

Für stetiges $f : [0, 1] \rightarrow \mathbb{R}$ gilt somit

$$\mathbb{E}[f(Y_{a,b})] \rightarrow f(y) \quad \text{wenn } a, b \rightarrow \infty \text{ mit } \frac{a}{a+b} \rightarrow y \in (0, 1)$$

(verwende z.B. die Chebychev-Ungleichung um zu sehen, dass $Y_{a,b}$ dann sehr eng um seinen Erwartungswert konzentriert ist).

Für $n_1, n_2 \in \mathbb{N}$ gilt

$$\begin{aligned} \mathbb{E}_z \left[\binom{n_1 + n_2}{n_1} Z_t^{n_1} (1 - Z_t)^{n_2} \right] \\ = \sum_{j=2}^{n_1+n_2} \mathbb{P}_{n_1+n_2}(N_t = j) \sum_{k=1}^{j-1} \binom{j}{k} z^k (1-z)^{j-k} \frac{\binom{n_1-1}{k-1} \binom{n_2-1}{j-k-1}}{\binom{n_1+n_2-1}{j-1}}. \end{aligned} \quad (2.27)$$

Der Erwartungswert auf der linken Seite ist die Wahrscheinlichkeit, in einer zufälligen Stichprobe der Größe $n_1 + n_2$ aus der Population zur Zeit t den Typ 1 n_1 -mal und den Typ 0 n_2 -mal zu sehen.

Für die Gleichung betrachten wir die folgende alternative Berechnung dieser Wahrscheinlichkeit, die auf die rechte Seite führt:

- Auf $\{N_t^{(n_1+n_2)} = j\}$ ($2 \leq j \leq n_1 + n_2$) denken wir uns die j Linien, die die Ahnen der Stichprobe zur Zeit 0 bilden, (zufällig) nummeriert.
- Eine gegebene Teilmenge K_1 der Größe k (mit $1 \leq k < j$) dieser j Linien erhält Typ 1, die übrigen Typ 2: dies hat W'keit $z^k (1-z)^{j-k}$
(und: es gibt $\binom{j}{k}$ verschiedene solche Teilmengen).
- Die W'keit, dass die k Linien aus K_1 genau n_1 Stichproben repräsentieren (und demnach repräsentieren die $j-k$ Linien aus $[j] \setminus K_1$ genau n_2 Stichproben), ist

$$\frac{|\{(m_1, \dots, m_j) \in \mathbb{N}^j : \sum_{i \in K_1} m_i = n_1, m_1 + \dots + m_j = n_1 + n_2\}|}{|\{(m_1, \dots, m_j) \in \mathbb{N}^j : m_1 + \dots + m_j = n_1 + n_2\}|} = \frac{\binom{n_1-1}{k-1} \binom{n_2-1}{j-k-1}}{\binom{n_1+n_2-1}{j-1}},$$

vergleiche Kor. 2.13.

Nehmen wir an, $\mathbb{P}_x(Z_t \in \cdot)$ besitzt Dichte $f_t(x, y)$ im Inneren $(0, 1)$.

Sei nun $y \in (0, 1)$, wir setzen $n_1 := \lfloor ny \rfloor$, $n_2 := n - n_1$, dann ist (setze $a = n_1 + 1$, $b = n_2 + 1$ in (2.27) oben)

$$\begin{aligned} \int_0^1 \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} z^{a-1} (1-z)^{b-1} f_t(x, z) dz &= \mathbb{E}_x \left[\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} Z_t^{a-1} (1-Z_t)^{b-1} \right] \\ &= (n+1) \mathbb{E}_x \left[\binom{n_1 + n_2}{n_1} Z_t^{n_1} (1-Z_t)^{n_2} \right] \end{aligned}$$

$$= \sum_{j=2}^{n_1+n_2} \mathbb{P}_{n_1+n_2}(N_t = j) \sum_{k=1}^{j-1} \binom{j}{k} x^k (1-x)^{j-k} (n+1) \frac{\binom{n_1-1}{k-1} \binom{n_2-1}{j-k-1}}{\binom{n_1+n_2-1}{j-1}} \quad (2.28)$$

[beachte $\Gamma(a+b) = \Gamma(n_1+n_2+2) = (n+1)!$

Es ist

$$(n+1) \frac{\binom{n_1-1}{k-1} \binom{n_2-1}{j-k-1}}{\binom{n_1+n_2-1}{j-1}} = \frac{(j-1)!}{(k-1)!(j-k-1)!} (n+1) \frac{([\!ny] - 1)_{(k-1)\downarrow} (n - [\!ny] - 1)_{(j-k-1)\downarrow}}{(n-1)_{(j-1)\downarrow}} \\ \xrightarrow{n \rightarrow \infty} \frac{(j-1)!}{(k-1)!(j-k-1)!} y^{k-1} (1-y)^{j-k-1} \quad (2.29)$$

Demnach finden wir insgesamt

Beobachtung 2.16. $\kappa_t(x, \cdot)$ für $t > 0$ besitzt auf $(0, 1)$ die Dichte

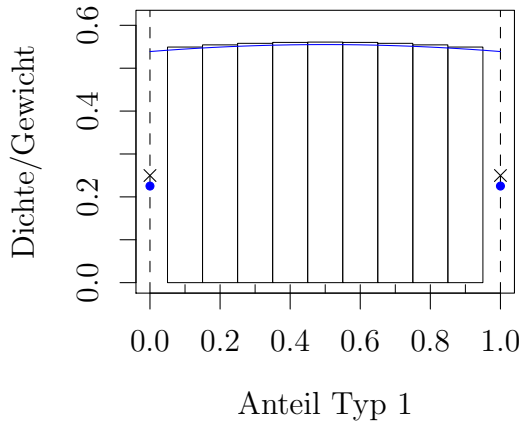
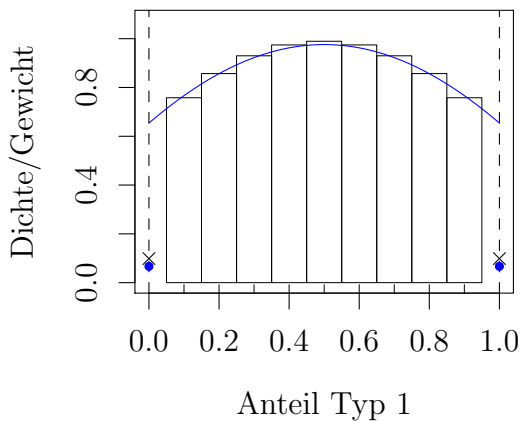
$$f_t(x, y) = \sum_{j=2}^{\infty} \mathbb{P}_{\infty}(N_t = j) \sum_{k=1}^{j-1} \binom{j}{k} x^k (1-x)^{j-k} \frac{(j-1)!}{(k-1)!(j-k-1)!} y^{k-1} (1-y)^{j-k-1}, \quad x, y \in (0, 1) \quad (2.30)$$

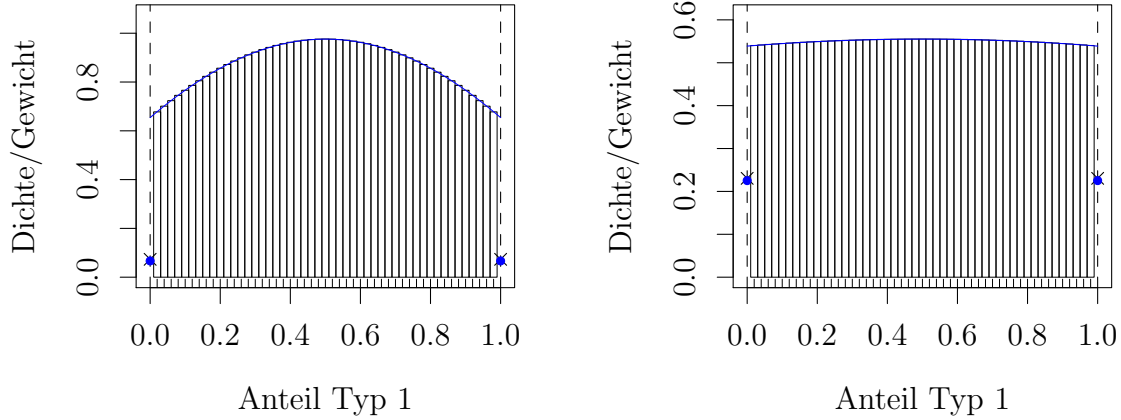
[also eine Mischung aus $\text{Beta}(k, j-k)$ -Dichten] und

$$\kappa_t(x, dy) = \mathbb{P}_x(Z_t = 1) \delta_1(dy) + \mathbb{P}_x(Z_t = 0) \delta_0(dy) + f_t(x, y) \mathbb{1}_{(0,1)}(y) dy. \quad (2.31)$$

mit $\mathbb{P}_x(Z_t = 1)$ aus (2.25) und $\mathbb{P}_x(Z_t = 0)$ aus (2.25). (Die Annahme, dass $\mathbb{P}_x(Z_t \in \cdot)$ tatsächlich eine stetige Dichte im Inneren $(0, 1)$ besitzt, können wir nachträglich durch Berechnung der Laplace-Transformierten rechtfertigen, vgl. [T84].)

Numerische Illustrationen





Vergleich der Übergangswahrscheinlichkeiten des Moran-Modells zum Zeitpunkt $Nt/2$ (schwarze Balken, normiert als Wahrscheinlichkeitshistogramm; schwarze Kreuze = Gewichte der Zustände 0 und 1) und Übergangsverteilung der Wright-Fisher-Diffusion (blaue Kurve = Dichte in $(0, 1)$; blaue Punkte = Gewichte der Atome auf 0 und auf 1), Startzustand $z_0 = 0.5$; linke Spalte: $t = 0.5$, rechte Spalte: $t = 1.0$; obere Zeile: $N = 10$, untere Zeile: $N = 50$

2.4 Wähler-Modelle auf endlichen Graphen

Szenario: E endliche Menge, $(q(x, y))_{x, y \in E}$ Q -Matrix einer (irreduzibelen) zeitkontinuierlichen Markovkette auf E (und wir interpretieren: es gibt eine Kante von x nach y , wenn $q(x, y) > 0$). $(q_t)_{t \geq 0} = (q_t(x, y); x, y \in E, t \geq 0)$ die zugehörige Übergangshalbgruppe. Sei $\pi = (\pi(x))_{x \in E}$ die eindeutige stationäre Verteilung zu q (d.h. $\sum_z \pi(z)q(z, y) = 0$ für alle $y \in E$).

$$d_E(t) = \max_{x \in E} \|q_t(x, \cdot) - \pi\|_{\text{TV}} \quad (2.32)$$

(mit $\|\mu - \nu\|_{\text{TV}} = \frac{1}{2} \sum_{x \in E} |\mu(x) - \nu(x)| = \max_{A \subset E} |\mu(A) - \nu(A)|$),

$$\pi_{\text{diag}} = \sum_{x \in E_n} (\pi(x))^2, \quad \mathbf{t}_{\text{mix}} = \inf \left\{ t \geq 0 : d_E(t) \leq \frac{1}{2e} \right\}, \quad \mathbf{t}_{\text{meet}} = \mathbb{E}[M_{U, U'}] \quad (2.33)$$

wobei $M_{U, U'}$ die Zeit bezeichnet, bis sich zwei unabhängige q -Irrfahrer mit unabhängigen Startpositionen $U, U' \sim \pi$ zum ersten Mal treffen (siehe (2.43) unten)

Wähler-Modell auf E : Zustandsraum $\{0, 1\}^E$, wir notieren die Zustände als $\xi = (\xi(x))_{x \in E} \in \{0, 1\}^E$. Wir schreiben $\widehat{\xi} = (\widehat{\xi}(x))_{x \in E}$ mit $\widehat{\xi}(x) = 1 - \xi(x)$. Für $\xi \in \{0, 1\}^E$, $x \in E$ entsteht

$$\xi^x \in \{0, 1\}^E \quad \text{mit} \quad \xi^x(z) = \begin{cases} \xi(z), & z \neq x, \\ 1 - \xi(z), & z = x \end{cases}$$

aus ξ , indem die „Meinung“ an Position x „geflippt“ wird.

Dynamik von $(\xi_t)_{t \geq 0}$:

$$c(\xi, \xi^x) = \sum_y (\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y))q(x, y)$$

(Interpretation: Der Wähler bei x kopiert mit Rate $q(x, y)$ die Meinung des Wählers bei y), Generator

$$Lf(\xi) = \sum_{x \in E} c(\xi, \xi^x)(f(\xi^x) - f(\xi)) \quad (2.34)$$

Es gibt eine graphische Konstruktion (siehe auch [L10, Section 4.3], [L85, Chapter V], [L99, Part. II]).

2.4.1 Einschlägige Martingale

Für $\xi \in \{0, 1\}^E$ seien

$$p_1(\xi) = \sum_{x \in E} \pi(x)\xi(x), \quad p_0(\xi) = \sum_{x \in E} \pi(x)\widehat{\xi}(x) = 1 - p_1(\xi) \quad (2.35)$$

$$p_{10}(\xi) = \sum_{x, y \in E} \bar{\nu}(x, y)\xi(x)\widehat{\xi}(y), \quad p_{01}(\xi) = \sum_{x, y \in E} \bar{\nu}(x, y)\widehat{\xi}(x)\xi(y) \quad (2.36)$$

mit

$$\nu(x, y) = \pi(x)^2 q(x, y) \mathbb{1}_{x \neq y}, \quad \nu(\mathbb{1}) = \sum_{x, y \in E} \nu(x, y), \quad \bar{\nu}(x, y) = \nu(x, y) / \nu(\mathbb{1}) \quad (2.37)$$

Lemma 2.17 ([CCC16, Proposition 3.1]). *Folgende Prozesse sind Martingale:*

$$(p_1(\xi_t))_{t \geq 0} \quad (2.38)$$

$$(p_1(\xi_t)p_0(\xi_t) + \nu(\mathbb{1}) \int_0^t (p_{10}(\xi_s) + p_{01}(\xi_s)) ds)_{t \geq 0} \quad (2.39)$$

$$((p_1(\xi_t))^2 - \nu(\mathbb{1}) \int_0^t (p_{10}(\xi_s) + p_{01}(\xi_s)) ds)_{t \geq 0} \quad (2.40)$$

Beweisskizze. Für (2.38) beobachte

$$p_1(\xi^x) - p_1(\xi) = (\widehat{\xi}(x) - \xi(x))\pi(x)$$

und daher

$$\begin{aligned} Lp_1(\xi) &= \sum_{x \in E} c(\xi, \xi^x)(\widehat{\xi}(x) - \xi(x))\pi(x) = \sum_{x, y \in E} (\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y))q(x, y)(\widehat{\xi}(x) - \xi(x))\pi(x) \\ &= \sum_{x, y \in E} (\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y))q(x, y)\pi(x) - 2 \sum_{x, y \in E} \xi(x)\widehat{\xi}(y)q(x, y)\pi(x) \\ &= \sum_{x, y \in E} \widehat{\xi}(x)\xi(y)q(x, y)\pi(x) - \sum_{x, y \in E} \xi(x)\widehat{\xi}(y)q(x, y)\pi(x) = 0 \end{aligned}$$

Für (2.39) beobachten wir zunächst, dass

$$p_1(\xi^x)p_0(\xi^x) - p_1(\xi)p_0(\xi) = (p_1(\xi) + (1 - 2\xi(x))\pi(x))(p_0(\xi) + (2\xi(x) - 1)\pi(x)) - p_1(\xi)p_0(\xi)$$

$$= \pi(x)(\widehat{\xi}(x) - \xi(x))p_0(\xi) + \pi(x)(\xi(x) - \widehat{\xi}(x))p_1(\xi) - \pi(x)^2$$

und somit

$$\begin{aligned} (L(p_1 p_0))(\xi) &= p_0(\xi) \underbrace{Lp_1(\xi)}_{=0} + p_1(\xi) \underbrace{Lp_0(\xi)}_{=0} - \sum_{x,y \in E} (\xi(x)\widehat{\xi}(y) + \widehat{\xi}(x)\xi(y))q(x,y)\pi(x)^2 \\ &= -\nu(\mathbb{1})(p_{10}(\xi) + p_{01}(\xi)) \end{aligned}$$

(2.40) folgt aus (2.39) zusammen mit $p_0(\xi) = 1 - p_1(\xi)$ und (2.38). \square

Korollar 2.18 ([CCC16, Corollary 3.2]). *Für $\gamma > 0$ ist die (prävisibile) quadratische Variation des Martingals $(p_1(\xi_{\gamma t}))_{t \geq 0}$ gegeben durch*

$$\langle p_1(\xi_{\gamma \cdot}) \rangle_t = \gamma \nu(\mathbb{1}) \int_0^t (p_{10}(\xi_{\gamma s}) + p_{01}(\xi_{\gamma s})) ds, \quad t \geq 0 \quad (2.41)$$

Beweis. Dies folgt aus (2.40) und der Definition der prävisiblen quadratische Variation. \square

Dualität: $(\widehat{X}_t^x, t \geq 0, x \in E)$ verschmelzende q -Markovketten mit $\widehat{X}_0^x = x$, so gilt

$$\mathbb{E}_\eta \left[\prod_{x \in F} \xi_t(x) \right] = \mathbb{E} \left[\prod_{x \in F} \eta(\widehat{X}_t^x) \right] \quad (2.42)$$

für jedes $F \subset E$. (Man kann (2.42) z.B. direkt aus der graphischen Konstruktion ablesen, oder aus einer Generator-Rechnung erhalten.)

Seien $(X_t^x, t \geq 0, x \in E)$ unabhängige q -Markovketten mit $X_0^x = x$,

$$M_{x,y} = \inf \{ t \geq 0 : X_t^x = X_t^y \} \quad (2.43)$$

$U, U' \sim \pi$ u.a., $(V, V') \sim \bar{\nu}$ (somit auch in Formeln $\mathbf{t}_{\text{meet}} = \mathbb{E}[M_{U,U'}]$).

Lemma 2.19 ([CCC16, Proposition 3.3]). *Für $\gamma > 0, t > 0$ und $\xi \in \{0,1\}^E$ gilt*

$$\mathbb{E}_\xi [p_1(\xi_{\gamma t})p_0(\xi_{\gamma t})] = \mathbb{E} [\xi(X_{\gamma t}^U)\widehat{\xi}(X_{\gamma t}^{U'}); M_{U,U'} > \gamma t] \quad (2.44)$$

$$\mathbb{E}_\xi [p_{10}(\xi_{\gamma t})] = \mathbb{E} [\xi(X_{\gamma t}^V)\widehat{\xi}(X_{\gamma t}^{V'}); M_{V,V'} > \gamma t] \quad (2.45)$$

$$\mathbb{E}_\xi [p_{01}(\xi_{\gamma t})] = \mathbb{E} [\widehat{\xi}(X_{\gamma t}^V)\xi(X_{\gamma t}^{V'}); M_{V,V'} > \gamma t] \quad (2.46)$$

Beweisskizze. Verwende die Definitionen von p_0, p_1, p_{10}, p_{01} , Linearität des Erwartungswerts und die Dualitätsformel (2.42). \square

Korollar 2.20 ([CCC16, Formel (3.13) und Corollary 3.4]). *Für $\gamma, t > 0$ gilt*

$$\sup_{\xi \in \{0,1\}^E} \mathbb{E}_\xi [p_{10}(\xi_{\gamma t}) + p_{01}(\xi_{\gamma t})] \leq 2\mathbb{P}(M_{V,V'} > \gamma t) \quad (2.47)$$

und

$$\mathbb{P}(M_{U,U'} > \gamma t) = 1 - \pi_{\text{diag}} - 2\gamma \nu(\mathbb{1}) \int_0^t \mathbb{P}(M_{V,V'} > \gamma s) ds \quad (2.48)$$

Zudem ist

$$\mathbb{E}[M_{V,V'}] = \int_0^\infty \mathbb{P}(M_{V,V'} > s) ds = \frac{1 - \pi_{\text{diag}}}{2\nu(\mathbb{1})} \quad (2.49)$$

Beweisskizze. (2.47) folgt aus (2.45)–(2.46).

Für (2.48) betrachte o.E. $\gamma = 1$ und verwende (mit $\xi_0(x)$, $x \in E$ u.i.v., $\text{Ber}(u)$ -verteilt, $0 < u < 1$)

$$\mathbb{E}[p_1(\xi_t)p_0(\xi_t)] = \mathbb{E}[p_1(\xi_0)p_0(\xi_0)] - \nu(\mathbb{1}) \int_0^t \mathbb{E}[p_{10}(\xi_s) + p_{01}(\xi_s)] ds$$

(verwende (2.39) aus Lemma 2.17). Mit den Dualitätsformeln aus Lemma 2.19 ergibt sich

$$u(1-u)\mathbb{P}(M_{U,U'} > t) = u(1-u)(1 - \pi_{\text{diag}}) - \nu(\mathbb{1}) \int_0^t 2\mathbb{P}(M_{V,V'} > s) ds$$

Division durch $u(1-u)$ liefert (2.48).

(2.49) folgt mit $t \rightarrow \infty$ aus (2.48) (denn angesichts der Irreduzibilität von q ist $M_{U,U'} < \infty$ fast sicher). \square

Lemma 2.21 ([CCC16, Proposition 6.1, (1)]). *Für $0 \leq s < t$ gilt*

$$\begin{aligned} \sup_{\xi \in \{0,1\}^E} |\mathbb{E}_\xi [p_{10}(\xi_t)] - \mathbb{P}(M_{V,V'} > s)p_1(\xi)p_0(\xi)| \\ \leq \mathbb{P}(M_{V,V'} \in (s, t]) + 4\mathbb{P}(M_{V,V'} > s) d_E(t-s) \end{aligned} \quad (2.50)$$

(und analog für $p_{01}(\xi_t)$).

Beweis. Es ist für $\xi \in \{0,1\}^E$

$$|(q_t \xi)(x) - p_1(\xi)| \leq 2d_E(t) \quad (2.51)$$

und mit (2.45) gilt für $s < t$

$$\mathbb{E}_\xi [p_{10}(\xi_t)] = \mathbb{E} [\xi(X_t^V) \widehat{\xi}(X_t^{V'}); M_{V,V'} > s] + \varepsilon_1(\xi; s, t) \quad (2.52)$$

wobei

$$|\varepsilon_1(\xi; s, t)| \leq \mathbb{P}(M_{V,V'} \in (s, t]) \quad (2.53)$$

Die Markoveigenschaft von $(X^V, X^{V'})$ angewendet zur Zeit s liefert

$$\begin{aligned} & \mathbb{E} [\xi(X_t^V) \widehat{\xi}(X_t^{V'}); M_{V,V'} > s] \\ &= \mathbb{E} \left[\mathbb{E} [\xi(X_{t-s}^V) \widehat{\xi}(X_{t-s}^{V'}) | (v, v') = (X_s^V, X_s^{V'})]; M_{V,V'} > s \right] = \mathbb{E} [q_{t-s} \xi(X_s^V) q_{t-s} \widehat{\xi}(X_s^{V'}); M_{V,V'} > s] \\ &= \mathbb{P}(M_{V,V'} > s) p_1(\xi) p_0(\xi) + \mathbb{E} [q_{t-s} \xi(X_s^V) q_{t-s} \widehat{\xi}(X_s^{V'}) - p_1(\xi) p_0(\xi); M_{V,V'} > s] \end{aligned} \quad (2.54)$$

Schließlich ist mit (2.51)

$$\begin{aligned} & \mathbb{E} [|q_{t-s} \xi(X_s^V) q_{t-s} \widehat{\xi}(X_s^{V'}) - p_1(\xi) p_0(\xi)|; M_{V,V'} > s] \\ & \leq \mathbb{E} [p_0(\xi) |q_{t-s} \xi(X_s^V) - p_1(\xi)|; M_{V,V'} > s] + \mathbb{E} [q_{t-s} \xi(X_s^V) |q_{t-s} \widehat{\xi}(X_s^{V'}) - p_0(\xi)|; M_{V,V'} > s] \\ & \leq 4d_E(t-s) \mathbb{P}(M_{V,V'} > s) \end{aligned} \quad (2.55)$$

Einsetzen von (2.53), (2.54) und (2.55) in (2.52) liefert (2.50). \square

2.4.2 Das Hauptergebnis

Szenario: Wir betrachten eine Schar von Markovketten. $\xi^{(n)}$, E_n , $q^{(n)}$, $\pi^{(n)}$, $\nu^{(n)}$, $\bar{\nu}^{(n)}$, $\pi_{\text{diag}}^{(n)} = \sum_{x \in E_n} (\pi^{(n)}(x))^2$, $\mathbf{t}_{\text{mix}}^{(n)}$, $\mathbf{t}_{\text{meet}}^{(n)}$, $\mathbb{P}^{(n)}$, $\mathbb{E}^{(n)}$, etc.

(Startbedingung: Wir betrachten (nur) $\xi_0^{(n)} \sim \text{Ber}(u)$ u.i.v. mit einem $u \in [0, 1]$; siehe [CCC16, Ende von Section 5, S. 306] für Diskussion allgemeinerer Startverteilungen.)

Satz 2.22 (eine Teilaussage von [CCC16, Thm. 2.2]). *Es gelte*

$$\lim_{n \rightarrow \infty} \pi_{\text{diag}}^{(n)} = 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \frac{\mathbf{t}_{\text{mix}}^{(n)}}{\mathbf{t}_{\text{meet}}^{(n)}} = 0 \quad (2.56)$$

und $\xi_0(x) \sim \text{Ber}_u$ seien u.i.v. mit $u \in [0, 1]$. Dann gilt mit $\gamma_n = \mathbf{t}_{\text{meet}}^{(n)}$

$$\left(p_1(\xi_{\gamma_n t}^{(n)}) \right)_{t \geq 0} \xrightarrow[n \rightarrow \infty]{d} (Z_t)_{t \geq 0} \quad (2.57)$$

Siehe Abschnitt 2.6 unten für Beispiel-Instanzen.

(Zwischen-)Ziel (vgl. [CCC16, Prop. 5.3]): Für $T > 0$ gilt

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left| \int_0^T \left(\gamma_n \nu_n(\mathbb{1}) [p_{10}(\xi_{\gamma_n t}^{(n)}) + p_{01}(\xi_{\gamma_n t}^{(n)})] - p_1(\xi_{\gamma_n t}^{(n)}) p_0(\xi_{\gamma_n t}^{(n)}) \right) dt \right| \right] = 0 \quad (2.58)$$

Tatsächlich gilt die Konvergenz in (2.58) auch in L^p für jedes $p \geq 1$, siehe [CCC16, Prop. 5.3].

Mittels (2.58) zeigen wir, dass jeder Limespunkt von $(p_1(\xi_{\gamma_n t}^{(n)}))_{t \geq 0}$ das (wohlgestellte) Martingalproblem (2.12) löst (und somit die Wright-Fisher-Diffusion ist).

Lemma 2.23 (vgl. [CCC16, Lemma 6.2]). *Es gibt eine Folge (s'_n) mit $s'_n = o(\gamma_n)$,*

$$\lim_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V, V'} > s'_n) = 1 \quad (2.59)$$

und

$$\varepsilon_n := \sup_{\xi \in E_n} \left| \gamma_n \nu_n(\mathbb{1}) \mathbb{E}_{\xi}^{(n)} [p_{10}(\xi_{s'_n}) + p_{01}(\xi_{s'_n})] - p_1(\xi) p_0(\xi) \right| \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.60)$$

Beweisskizze. Betrachte die Produktkette auf $E_n \times E_n$ mit Übergangshalbgruppe $\tilde{q}_t^{(n)} = q_t^{(n)} \otimes q_t^{(n)}$, startend in ihrer stationären Verteilung $\pi^{(n)} \otimes \pi^{(n)}$; die Auftreffzeit auf der Diagonale $D_n = \{(x, x) : x \in E_n\} \subset E_n \times E_n$ hat (nach Konstruktion) dieselbe Verteilung wie $M_{U, U'}^{(n)}$. Weiterhin gilt

$$\max_{(x, x') \in E_n \times E_n} \|\tilde{q}_t^{(n)} - \pi^{(n)} \otimes \pi^{(n)}\|_{\text{TV}} \leq 2d_{E_n}(t)$$

Annahme (2.56) zusammen mit Satz 2.24 ([A82, Thm. 1.4]) und Korollar 2.25 liefert dann

$$\frac{M_{U, U'}}{\gamma_n} \xrightarrow[n \rightarrow \infty]{d} \text{Exp}(1) \quad (2.61)$$

(siehe Abschnitt 2.5 unten für Intuition/Details), d.h.

$$\mathbb{P}^{(n)}(M_{U,U'} > \gamma_n t) \xrightarrow[n \rightarrow \infty]{} e^{-t} \quad \text{für alle } t \geq 0 \quad (2.62)$$

Mit (2.48) aus Korollar 2.20 und Annahme 2.56 folgt daraus

$$2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n s) ds \xrightarrow[n \rightarrow \infty]{} 1 - e^{-t} \quad \text{für alle } t \geq 0 \quad (2.63)$$

Wie in [CCC16, Cor. 4.2] gewinnen wir aus (2.63)

$$2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n t) \xrightarrow[n \rightarrow \infty]{} e^{-t} \quad \text{für alle } t > 0 \quad (2.64)$$

(2.64) ist intuitiv plausibel, etwas mehr Details:

Die Funktion $f_n(t) := 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n t)$ ist stetig und fallend, und sie erfüllt $f_n(t) \leq 1/t$ (verwende (2.49) und die Markov-Ungleichung).

Mit [CCC16, S. 300]: Wähle $a > 0$, $G_n(t) := 1 - a f_n(t)$ ist Verteilungsfunktion auf eines (Sub-)Wahrscheinlichkeitsmaßes auf $[a, \infty)$. Für jede konvergente Teilfolge (die es gemäß Auswahlssatz von Helly-Bray stets gibt) mit $G_{n_k}(\cdot) \rightarrow G(\cdot)$ (in allen Stetigkeitspunkten von G) gilt

$$\int_s^t G_{n_k}(u) du \xrightarrow[k \rightarrow \infty]{} \int_s^t G(u) du \quad \text{für } a < s < t$$

mit dominierter Konvergenz (beachte: als Verteilungsfunktion G hat höchstens abzählbar viele Unstetigkeitsstellen). (2.63) zeigt

$$\int_s^t G(u) du = (t - s) + a(e^{-t} - e^{-s})$$

und daher $G(t) = 1 - ae^{-t}$ für jeden Stetigkeitspunkt von G . Da G nicht-fallend ist, gilt dies für alle $t \in (a, \infty)$. Also ist G stetig auf (a, ∞) und $\lim_{k \rightarrow \infty} f_{n_k}(t) = e^{-t}$ für $t > a$. Da der Grenzwert nicht von der Teilfolge abhängt, gilt $\lim_{n \rightarrow \infty} f_n(t) = e^{-t}$ für $t > a$. (Schließlich $a \downarrow 0$.)

Wir können somit eine Folge (s_n) mit $s_n = o(\gamma_n)$ und

$$\lim_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > s_n) = 1 \quad (2.65)$$

wählen (für jede Nullfolge $\varepsilon_n \rightarrow 0$ gilt $\liminf_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \varepsilon_n \gamma_n) \geq 1$ gemäß (2.64), andererseits gilt für $\varepsilon_n := \inf\{\varepsilon > 0 : 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \varepsilon \gamma_n)\} \leq 1$ nach Konstruktion $\limsup_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \varepsilon_n \gamma_n) \leq 1$ und wiederum wegen (2.64) ist $\limsup_{n \rightarrow \infty} \varepsilon_n \leq t$ für jedes $t > 0$, d.h. $(\varepsilon_n)_n$ ist eine Nullfolge).

Wähle $u_n \rightarrow \infty$ mit $\lim_{n \rightarrow \infty} u_n \mathbf{t}_{\text{mix}}^{(n)} / \gamma_n = 0$, setze $s'_n = s_n \vee u_n \mathbf{t}_{\text{mix}}^{(n)}$. Somit

$$\delta_n = s'_n / \gamma_n \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.66)$$

und

$$d_{E_n}(s'_n) \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.67)$$

(denn $s'_n/\mathbf{t}_{\text{mix}}^{(n)} \rightarrow \infty$; erinnere $d_E(\cdot)$ aus (2.32) und siehe (2.90) unten oder [LPW, Section 4.5]).

Wie im Beweis von [CCC16, Prop. 4.3] gilt dann auch

$$\lim_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > s'_n) = 1 \quad (2.68)$$

(Vergleich mit (2.64) zeigt $\liminf_{n \rightarrow \infty} \dots \geq e^{-t}$ für jedes $t > 0$, d.h. $\liminf_{n \rightarrow \infty} \dots \geq 1$; (2.65) liefert $\limsup_{n \rightarrow \infty} \dots \leq 1$).

Indem wir dasselbe Argument nochmals auf $s''_n = 2s'_n$ anwenden, folgt auch

$$\lim_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} \in (s'_n, 2s'_n]) = 0 \quad (2.69)$$

Mit Lemma 2.21 finden wir für jedes $\xi \in E_n$

$$\begin{aligned} & \left| \gamma_n \nu_n(\mathbb{1}) \mathbb{E}_\xi [p_{10}(\xi_{s'_n}) + p_{01}(\xi_{s'_n})] - p_1(\xi)p_0(\xi) \right| \\ & \leq 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} \in (s'_n, 2s'_n]) + 8\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > s'_n) d_{E_n}(s'_n) \\ & \quad + \left| 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > s'_n) - 1 \right| p_1(\xi)p_0(\xi) \end{aligned}$$

mit (2.67), (2.68) und (2.69) folgt (2.60). \square

Beweis von (2.58) für $p = 1$. Sei $\delta_n = s'_n/\gamma_n \rightarrow 0$ wie in (2.66) im Beweis von Lemma 2.23. Zunächst gilt mit (2.47) aus Korollar 2.20

$$\begin{aligned} \mathbb{E}^{(n)} \left[\int_0^{2\delta_n} \left| \gamma_n \nu(\mathbb{1}) [p_{10}(\xi_{\gamma_n t}^{(n)}) + p_{01}(\xi_{\gamma_n t}^{(n)})] - p_1(\xi_{\gamma_n t}^{(n)}) p_0(\xi_{\gamma_n t}^{(n)}) \right| ds \right] \\ \leq 2\gamma_n \nu_n(\mathbb{1}) \int_0^{2\delta_n} \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n s) ds + 2\delta_n \xrightarrow[n \rightarrow \infty]{} 0 \end{aligned}$$

Es genügt also zu zeigen, dass

$$I_n(T) := \mathbb{E}^{(n)} \left[\int_{2\delta_n}^T \left| \gamma_n \nu(\mathbb{1}) [p_{10}(\xi_{\gamma_n t}^{(n)}) + p_{01}(\xi_{\gamma_n t}^{(n)})] - p_1(\xi_{\gamma_n t}^{(n)}) p_0(\xi_{\gamma_n t}^{(n)}) \right| ds \right] \xrightarrow[n \rightarrow \infty]{} 0 \quad (2.70)$$

gilt.

Schreibe $\bar{p}(\xi) = p_{10}(\xi) + p_{01}(\xi)$ und (mit $\mathcal{F}_t^n = \sigma(\xi_{\gamma_n u} : u \leq t)$)

$$H_n(s) = \gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) - \mathbb{E}^{(n)} \left[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_{s-2\delta_n}^n \right] \quad (2.71)$$

Mit Dreiecksungleichung und Cauchy-Schwarz-Ungleichung ist

$$\begin{aligned} I_n(T) & \leq \left(\mathbb{E}^{(n)} \left[\left(\int_{2\delta_n}^T H_n(s) ds \right)^2 \right] \right)^{1/2} \\ & \quad + \mathbb{E}^{(n)} \left[\int_{2\delta_n}^T \left| \mathbb{E}^{(n)} \left[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_{s-2\delta_n}^n \right] - p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)}) \right| ds \right] \\ & \quad + \mathbb{E}^{(n)} \left[\left| \int_{2\delta_n}^T p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)}) - p_1(\xi_{\gamma_n s}) p_0(\xi_{\gamma_n s}) ds \right| \right] \end{aligned}$$

$$=: (I_{n,1}(T))^{1/2} + I_{n,2}(T) + I_{n,3}(T) \quad (2.72)$$

Zum ersten Term auf der rechten Seite von (2.72):

$$I_{n,1}(T) = 2\mathbb{E}^{(n)} \left[\iint_{2\delta_n \leq r \leq s \leq T} H_n(r) H_n(s) \mathbb{1}_{r > s - 2\delta_n} ds dr \right]$$

Zur Rechtfertigung des Indikators $\mathbb{1}_{r > s - 2\delta_n}$ beachte: Für $2\delta_n \leq r < s - 2\delta_n$ ist

$$\mathbb{E}^{(n)}[H_n(r)H_n(s)] = \mathbb{E}^{(n)} \left[H_n(r) \underbrace{\mathbb{E}^{(n)}[H_n(s) \mid \mathcal{F}_r^n]}_{=0} \right] = 0$$

Mit der Definition von $H_n(s)$ aus (2.71) ergibt sich

$$\begin{aligned} \frac{I_{n,1}(T)}{2(\gamma_n \nu_n(\mathbb{1}))^2} &= \mathbb{E}^{(n)} \left[\int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \bar{p}(\xi_{\gamma_n r}) \bar{p}(\xi_{\gamma_n s}) \right] \\ &\quad - \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \mathbb{E}^{(n)} \left[\mathbb{E}^{(n)}[\bar{p}(\xi_{\gamma_n r}) \mid \mathcal{F}_{r-2\delta_n}^n] \bar{p}(\xi_{\gamma_n s}) \right] \\ &\quad - \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \mathbb{E}^{(n)} \left[\bar{p}(\xi_{\gamma_n r}) \mathbb{E}^{(n)}[\bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_{s-2\delta_n}^n] \right] \\ &\quad + \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \mathbb{E}^{(n)} \left[\mathbb{E}^{(n)}[\bar{p}(\xi_{\gamma_n r}) \mid \mathcal{F}_{r-2\delta_n}^n] \times \mathbb{E}^{(n)}[\bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_{s-2\delta_n}^n] \right] \\ &=: I'_{n,1}(T) - I'_{n,2}(T) - I'_{n,3}(T) + I'_{n,4}(T) \end{aligned}$$

Beobachtungen:

$$\lim_{n \rightarrow \infty} 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) = 1 \quad (2.73)$$

(verwende (2.64)),

$$K_t := \sup_{n \in \mathbb{N}} 2\gamma_n \nu_n(\mathbb{1}) \int_0^t \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n s) ds < \infty \quad (2.74)$$

(dies folgt z.B. aus (2.49) in Korollar 2.20); für $r < s$ gilt gemäß der Markov-Eigenschaft und (2.47) aus Korollar 2.20

$$\mathbb{E}^{(n)}[\bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_r^n] = \mathbb{E}_{\xi_{\gamma_n r}}^{(n)}[\bar{p}(\xi_{\gamma_n(s-r)})] \leq 2\mathbb{P}^{(n)}(M_{V,V'} > \gamma_n(s-r)) \quad (2.75)$$

Damit ergibt sich

$$\begin{aligned} &(\gamma_n \nu_n(\mathbb{1}))^2 I'_{n,1}(T) \\ &= \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \mathbb{E}^{(n)} \left[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n r}) \mathbb{E}^{(n)}[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_r^n] \right] \\ &\leq \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \mathbb{E}^{(n)}[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n r})] \times 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n(s-r)) \\ &\leq 2\gamma_n \nu_n(\mathbb{1}) \int_{2\delta_n}^T \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n r) dr \times 2\gamma_n \nu_n(\mathbb{1}) \int_0^{2\delta_n} \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n s) ds \\ &\leq K_T \times 2\gamma_n \nu_n(\mathbb{1}) \int_0^{2\delta_n} \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n s) ds \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Weiter ist (verwende analog (2.75) und (2.47) aus Korollar 2.20)

$$\begin{aligned}
0 &\leq (\gamma_n \nu_n(\mathbb{1}))^2 I'_{n,2}(T) \\
&\leq \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \times 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n s) \\
&\leq \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \left(2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \right)^2 \\
&\leq 2\delta_n T \times \left(2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \right)^2 \xrightarrow{n \rightarrow \infty} 0
\end{aligned}$$

Ähnlich ist

$$\begin{aligned}
0 &\leq (\gamma_n \nu_n(\mathbb{1}))^2 I'_{n,3}(T) \\
&\leq \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \times 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n r) \\
&\leq 2\delta_n \times 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \times \int_0^T 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > \gamma_n r) dr \\
&=\leq 2\delta_n \times 2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \times K_T \xrightarrow{n \rightarrow \infty} 0
\end{aligned}$$

Schließlich

$$\begin{aligned}
0 &\leq (\gamma_n \nu_n(\mathbb{1}))^2 I'_{n,4}(T) \\
&\leq \int_{2\delta_n}^T dr \int_r^{T \wedge (r+2\delta_n)} ds \left(2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \right)^2 \\
&\leq 2\delta_n T \left(2\gamma_n \nu_n(\mathbb{1}) \mathbb{P}^{(n)}(M_{V,V'} > 2\gamma_n \delta_n) \right)^2 \xrightarrow{n \rightarrow \infty} 0
\end{aligned}$$

insgesamt also $I_{n,1}(T) \rightarrow 0$ mit $n \rightarrow \infty$.

Mit ε_n aus (2.60) in Lemma 2.23 ist

$$\begin{aligned}
I_{n,2}(T) &= \mathbb{E}^{(n)} \left[\int_{2\delta_n}^T \left| \mathbb{E}^{(n)} \left[\gamma_n \nu_n(\mathbb{1}) \bar{p}(\xi_{\gamma_n s}) \mid \mathcal{F}_{s-2\delta_n}^n \right] - p_1(\xi_{\gamma_n(s-2\delta_n)}) p_0(\xi_{\gamma_n(s-2\delta_n)}) \right| ds \right] \\
&\leq \int_{2\delta_n}^T \varepsilon_n ds \leq \varepsilon_n T \xrightarrow{n \rightarrow \infty} 0
\end{aligned}$$

Zudem ist (da $0 \leq p_1(\cdot) p_0(\cdot) = p_1(\cdot)(1 - \varepsilon_n) \leq 1/4$ stets)

$$I_{n,3}(T) \leq \int_0^{2\delta_n} \mathbb{E}^{(n)} [p_1(\xi_{\gamma_n u}) p_0(\xi_{\gamma_n u})] ds + \int_{T-2\delta_n}^T \mathbb{E}^{(n)} [p_1(\xi_{\gamma_n u}) p_0(\xi_{\gamma_n u})] ds \leq \delta_n \xrightarrow{n \rightarrow \infty} 0$$

Insgesamt folgt (2.70). □

Beweis von Satz 2.22. 1. Zeige (siehe [CCC16, Beweis von Theorem 5.1]):

$$\text{Die Folge } (Y_n(t))_{t \geq 0} := (p_1(\xi_{\gamma_n t}^{(n)}))_{t \geq 0}, n \in \mathbb{N} \text{ ist } C\text{-straff} \tag{2.76}$$

Das Martingal Y_n hat nach Korollar 2.18 die (prävisible) quadratische Variation

$$\langle Y_n \rangle_t = \gamma_n \nu_n(\mathbb{1}) \int_0^t \left(p_{10}(\xi_{\gamma_n s}^{(n)}) + p_{01}(\xi_{\gamma_n s}^{(n)}) \right) ds \quad (2.77)$$

und wir zeigen zunächst:

$$\text{Die Folge } (\langle Y_n \rangle_t)_{t \geq 0}, n \in \mathbb{N} \text{ ist } C\text{-straff} \quad (2.78)$$

und verwenden dazu Aldous' Straffheitskriterium (z.B. [JS, Thm. VI.4.5] oder [EK, Ch. 3, Thm. 8.6]):

(i) Für $n \in \mathbb{N}, K > 0, T > 0$ ist (denn $t \mapsto \langle Y_n \rangle_t$ ist monoton wachsend)

$$\begin{aligned} \mathbb{P}^{(n)}\left(\sup_{t \leq T} \langle Y_n \rangle_t > K\right) &= \mathbb{P}^{(n)}(\langle Y_n \rangle_T > K) \\ &\leq \frac{1}{K} \mathbb{E}^{(n)}[\langle Y_n \rangle_T] = \frac{\gamma_n \nu_n(\mathbb{1})}{K} \int_0^T \mathbb{E}^{(n)}[p_{10}(\xi_{\gamma_n s}^{(n)}) + p_{01}(\xi_{\gamma_n s}^{(n)})] ds \\ &\leq \frac{2\gamma_n \nu_n(\mathbb{1})}{K} \int_0^T \mathbb{P}^{(n)}(M_{V, V'}^{(n)} > \gamma_n s) ds \leq \frac{1}{K} \end{aligned}$$

(für das erste „ \leq “ verwende, dass $t \mapsto \langle Y_n \rangle_t$ monoton wachsend ist, dann die Markov-Ungleichung, (2.77) und schließlich (2.47) sowie (2.49) aus Korollar 2.20).

(ii) Für $\delta > 0$ und $(\mathcal{F}_t^{(n)})$ -Stopppzeiten $S \leq T \leq S + \delta$ ist

$$\begin{aligned} \mathbb{E}^{(n)}[|\langle Y_n \rangle_T - \langle Y_n \rangle_S|] &= \mathbb{E}^{(n)}[\langle Y_n \rangle_T - \langle Y_n \rangle_S] = \gamma_n \nu_n(\mathbb{1}) \mathbb{E}^{(n)}\left[\int_S^T (p_{10}(\xi_{\gamma_n u}^{(n)}) + p_{01}(\xi_{\gamma_n u}^{(n)})) du\right] \\ &\leq \gamma_n \nu_n(\mathbb{1}) \mathbb{E}^{(n)}\left[\mathbb{E}_{\xi_{\gamma_n S}^{(n)}}^{(n)}\left[\int_0^\delta (p_{10}(\xi_{\gamma_n u}^{(n)}) + p_{01}(\xi_{\gamma_n u}^{(n)})) du\right]\right] \\ &\leq 2\gamma_n \nu_n(\mathbb{1}) \int_0^\delta \mathbb{P}^{(n)}(M_{V, V'}^{(n)} > \gamma_n u) du \xrightarrow{n \rightarrow \infty} 1 - e^{-\delta} \end{aligned}$$

(für das erste „ \leq “ verwende die starke Markov-Eigenschaft zur Zeit S , für das zweite „ \leq “ verwende (2.47) aus Korollar 2.20 und für die Konvergenz (2.63)).

(2.78) folgt aus (i) und (ii) (die C -Straffheit folgt aus der Tatsache, dass jedes $\langle Y_n \rangle$ stetige Pfade hat).

Da die $[0, 1]$ -wertige Folge $Y_n(0)$, $n \in N$ automatisch straff ist und jedes Y_n ein Martingal ist, folgt aus (2.78) zusammen mit [JS, Thm. VI.4.13], dass die Folge Y_n straff ist. Weiterhin hat Y_n höchstens Sprünge der Größe $\max_{x \in E_n} \pi^{(n)}(x) \leq (\pi_{\text{diag}}^{(n)})^{1/2} \rightarrow 0$ für $n \rightarrow \infty$ nach Annahme (2.56). Somit gilt (2.76).

2. Zeige:

$$\text{Jeder Häufungspunkt von } (p_1(\xi_{\gamma_n t}^{(n)}))_{t \geq 0} \text{ löst das Martingalproblem} \quad (2.79)$$

der Wright-Fisher-Diffusion ((2.13) aus Bem. 2.7).

Sei $n_k \nearrow \infty$, so dass $Y_{n_k} \rightarrow^d Z$ für $k \rightarrow \infty$, wo Z ein $[0, 1]$ -wertiger stochastischer Prozess mit stetigen Pfaden ist (gemäß 1. gibt es eine solche Teilfolge). Da alle Y_n gleichmäßig (durch 1) beschränkte Martingale sind, ist auch Z ein Martingal (bezüglich der von ihm selbst erzeugten Filtration).

Wir zeigen:

$$Z_t^2 - Z_0^2 - \int_0^t Z_u(1 - Z_u) du, \quad t \geq 0 \quad \text{ist ein Martingal} \quad (2.80)$$

Seien $0 \leq s < t$, $m \in \mathbb{N}$, $0 \leq s_1 < s_2 < \dots < s_m \leq s$, $\varphi_1, \dots, \varphi_m \in C([0, 1], \mathbb{R})$, so ist

$$\begin{aligned} & \mathbb{E}\left[\left(Z_t^2 - Z_s^2 - \int_s^t Z_u(1 - Z_u) du\right) \times \prod_{j=1}^m \varphi_j(Z_{s_j})\right] \\ &= \lim_{k \rightarrow \infty} \mathbb{E}\left[\left(Y_{n_k}(t)^2 - Y_{n_k}(s)^2 - \int_s^t Y_{n_k}(u)(1 - Y_{n_k}(u)) du\right) \times \prod_{j=1}^m \varphi_j(Y_{n_k}(s_j))\right] \end{aligned}$$

(verwende C -Straffheit). Nach Definition (der quadratischen Variation) ist für jedes n

$$\begin{aligned} 0 &= \mathbb{E}\left[\left(Y_n(t)^2 - Y_n(s)^2 - ((Y_n)_t - (Y_n)_s)\right) \times \prod_{j=1}^m \varphi_j(Y_n(s_j))\right] \\ &= \mathbb{E}\left[\left(Y_n(t)^2 - Y_n(s)^2 - \gamma_n \nu_n(\mathbb{1}) \int_s^t (p_{10}(\xi_{\gamma_n u}) + p_{01}(\xi_{\gamma_n u})) du\right) \times \prod_{j=1}^m \varphi_j(Y_n(s_j))\right] \end{aligned}$$

(verwende Korollar 2.18 für das 2. „=“), zusammen mit (2.58) folgt (2.80):

$$\begin{aligned} & \left| \mathbb{E}\left[\left(Z_t^2 - Z_s^2 - \int_s^t Z_u(1 - Z_u) du\right) \times \prod_{j=1}^m \varphi_j(Z_{s_j})\right] \right| \\ & \leq \limsup_{k \rightarrow \infty} \mathbb{E}\left[\left| \int_0^T (\gamma_{n_k} \nu_{n_k}(\mathbb{1}) [p_{10}(\xi_{\gamma_{n_k} t}^{(n_k)}) + p_{01}(\xi_{\gamma_{n_k} t}^{(n_k)})] - p_1(\xi_{\gamma_{n_k} t}^{(n_k)}) p_0(\xi_{\gamma_{n_k} t}^{(n_k)})) dt \right| \times \prod_{j=1}^m |\varphi_j(Y_n(s_j))| \right] \\ & \leq \left(\prod_{j=1}^m \|\varphi_j\|_\infty \right) \lim_{n \rightarrow \infty} \mathbb{E}\left[\left| \int_0^T (\gamma_n \nu_n(\mathbb{1}) [p_{10}(\xi_{\gamma_n t}^{(n)}) + p_{01}(\xi_{\gamma_n t}^{(n)})] - p_1(\xi_{\gamma_n t}^{(n)}) p_0(\xi_{\gamma_n t}^{(n)})) dt \right| \right] = 0 \end{aligned}$$

□

2.5 Zur „Beinahe-Exponentialität“ von Auftreffzeiten

Wir benötigen:

Satz 2.24 (D. Aldous, [A82, Thm. 1.4 (a)]). $(X_t)_{t \geq 0}$ irreduzible zeitkontinuierliche Markovkette auf E ($|E| < \infty$) mit stationärer Verteilung π , $A \subset E$, $T_A := \inf\{t \geq 0 : X_t \in A\}$ die Auftreffzeit in A , t_{mix} die Mischzeit.

$$\sup_{t \geq 0} \left| \mathbb{P}_\pi(T_A > t) - \exp(-t/E_\pi[T_A]) \right| \leq \Delta \quad (2.81)$$

mit

$$\Delta := 15 \frac{t_{mix}}{\mathbb{E}_\pi[T_A]} \left(2 + \log^+ \left(\frac{\mathbb{E}_\pi[T_A]}{t_{mix}} \right) \right) \quad (2.82)$$

(wobei $\log^+(x) = \log(x) \vee 0$).

Korollar 2.25. *Betrachte eine Folge von Markovketten $X^{(n)}$ auf E_n , $A_n \subset E_n$, mit zugehöriger stationärer Verteilung $\pi^{(n)}$, etc. wie in Abschnitt 2.4.2. Falls gilt (wie in Annahme (2.56) in Satz 2.22)*

$$\lim_{n \rightarrow \infty} \frac{\mathbf{t}_{\text{mix}}^{(n)}}{\mathbf{t}_{\text{meet}}^{(n)}} = 0 \quad (2.83)$$

so folgt (siehe (2.43) und die Formeln danach für die Definition von $M_{U,U'}$)

$$\mathcal{L}_{\pi^{(n)}}(M_{U,U'}/\mathbf{t}_{\text{mix}}^{(n)}) \xrightarrow[n \rightarrow \infty]{} \text{Exp}(1) \quad (2.84)$$

Beobachtungen/Vorbemerkungen:

1. Die *quasi-stationäre Verteilung* von X (Formulierung: „bezüglich T_A “?) ist

$$\rho = \lim_{t \rightarrow \infty} \mathbb{P}_\pi(X_t \in \cdot \mid T_A > t) \quad (2.85)$$

(wir nehmen [zunächst] Existenz an; diese ist sichergestellt, wenn X irreduzibel auf $E \setminus A$ ist, siehe [A82, Rem. 2.18] für den allgemeinen Fall). Dann ist

$$\rho = \mathbb{P}_\rho(X_t \in \cdot \mid T_A > t) \quad \text{für jedes } t \geq 0 \quad (2.86)$$

insbesondere ist

$$\mathbb{P}_\rho(T_A > t) = \exp(-t/\mathbb{E}_\rho[T_A]) \quad (2.87)$$

(und die Idee des Beweises von Satz 2.24 ist zu zeigen, dass $\|\pi - \rho\|_{\text{TV}}$ klein ist).

2. Für eine Zufallsvariable Y und ein Ereignis B mit $\mathbb{P}(B) > 0$ gilt

$$\|\mathbb{P}(Y \in \cdot) - \mathbb{P}(Y \in \cdot \mid B)\|_{\text{TV}} \leq 2(1 - \mathbb{P}(B)) \quad (2.88)$$

(denn für Ereignisse A, B mit $\mathbb{P}(B) > 0$ ist $|\mathbb{P}(A) - \mathbb{P}(A|B)| = |\mathbb{P}(B)\mathbb{P}(A|B) + \mathbb{P}(B^c)\mathbb{P}(A|B^c) - \mathbb{P}(A|B)| \leq \mathbb{P}(A|B)|\mathbb{P}(B) - 1| + \mathbb{P}(B^c)\mathbb{P}(A|B^c) \leq 1 - \mathbb{P}(B) + \mathbb{P}(B^c) = 2 - 2\mathbb{P}(B)$, wende dies an auf Ereignisse der Form $A = \{Y \in C\}$).

3. Für jede Startverteilung μ gilt

$$\|\mathbb{P}_\mu(X_t \in \cdot) - \pi\|_{\text{TV}} \leq \exp(1 - t/\mathbf{t}_{\text{mix}}) \quad (2.89)$$

denn $d_E(t) = \max_{x \in E} \|q_t(x, \cdot) - \pi\|_{\text{TV}}$ (aus (2.32)) ist nicht-wachsend in t und erfüllt $d_E(s+t) \leq 2d_E(s)d_E(t)$ für $s, t > 0$ (verwende jeweils die Markoveigenschaft). Nach Definition ist $d_E(\mathbf{t}_{\text{mix}}) \leq 1/(2e)$, somit induktiv

$$d_E(k\mathbf{t}_{\text{mix}}) \leq \frac{1}{2}e^{-k} \quad \text{für } k \in \mathbb{N} \quad (2.90)$$

was (2.89) impliziert.

4. Es ist

$$\max_{x \in E} \mathbb{E}_x[T_A] \leq \frac{\mathbf{t}_{\text{mix}} + \mathbb{E}_\pi[T_A]}{1 - 1/(2e)} \quad (2.91)$$

(für $x \in E$ setze $\mu := \mathbb{P}_x(X_{\mathbf{t}_{\text{mix}}} \in \cdot)$, so ist $\mathbb{E}_x[T_A] \leq \mathbf{t}_{\text{mix}} + \mathbb{E}_\mu[T_A]$, andererseits ist $\|\mu - \pi\|_{\text{TV}} \leq 1/(2e)$ und daher $\mathbb{E}_\mu[T_A] \leq \mathbb{E}_\pi[T_A] + (2e)^{-1} \max_{y \in E} \mathbb{E}_y[T_A]$, also gilt

$$\max_{x \in E} \mathbb{E}_x[T_A] \leq \mathbf{t}_{\text{mix}} + \mathbb{E}_\pi[T_A] + \frac{1}{2e} \max_{y \in E} \mathbb{E}_y[T_A],$$

d.h. (2.91)).

Beweis von Satz 2.24. Wir zeigen zunächst

$$\sup_{t \geq 0} \left| \mathbb{P}_\pi(T_A > t) - \exp(-t/\mathbb{E}_\pi[T_A]) \right| \leq 3\|\rho - \pi\|_{\text{TV}} \quad (2.92)$$

Es ist für $t > 0$ mit (2.87)

$$\begin{aligned} \left| \mathbb{P}_\pi(T_A > t) - \exp(-t/\mathbb{E}_\pi[T_A]) \right| &\leq \|\rho - \pi\|_{\text{TV}} + \sup_{s \geq 0} \left| \exp(-s/\mathbb{E}_\rho[T_A]) - \exp(-s/\mathbb{E}_\pi[T_A]) \right| \\ &\leq \|\rho - \pi\|_{\text{TV}} + \left| \frac{\mathbb{E}_\rho[T_A]}{\mathbb{E}_\pi[T_A]} - 1 \right| \end{aligned} \quad (2.93)$$

(denn für $f(s) := e^{-as} - e^{-bs}$ gilt $\max_{s \geq 0} |f(s)| = |f(s_*)|$ mit $s_* = (\log(a \vee b) - \log(a \wedge b))/(a \vee b - a \wedge b)$ und $|f(s_*)| = ((a \wedge b)/(a \vee b))^{(a \wedge b)/(a \vee b - a \wedge b)} (1 - (a \wedge b)/(a \vee b)) \leq 1 - (a \wedge b)/(a \vee b) \leq |1 - \frac{a}{b}|$).

Betrachte den Fall $a < b$, dann ist $f(s) \geq 0$ für $s \geq 0$ mit $0 = f(0) = \lim_{s \rightarrow \infty} f(s)$ und $f'(s) = -ae^{-as} + be^{-bs} = e^{-as}(be^{-(b-a)s} - a)$ hat s_* als einzige Nullstelle, an der dann notwendigerweise ein Maximum vorliegen muss.

Im Fall $a > b$ hat $-f(s) \geq 0$ an der Stelle s_* sein Maximum. Weiter ist für $a, b > 0$ stets $0 \leq 1 - (a \wedge b)/(a \vee b) \leq |1 - a/b|$; für $a \leq b$ ist nichts zu beweisen, für $a > b$ gilt $1 - b/a < a/b - 1 \iff 1 - \frac{1}{x} < x - 1$ für $x = a/b > 1$ [was wegen Konvexität von $x \mapsto \frac{1}{x}$ erfüllt ist: $\frac{1}{x} > 2 - x$].

Weiter ist (verwende (2.91) für das 2. „ \leq “, $|\mathbb{E}_\mu[Y] - \mathbb{E}_\nu[Y]| \leq \|\mu - \nu\|_{\text{TV}} \max_x \mathbb{E}_x[Y]$ für das 1. „ \leq “)

$$\begin{aligned} \left| \mathbb{E}_\rho[T_A] - \mathbb{E}_\pi[T_A] \right| &\leq \|\rho - \pi\|_{\text{TV}} \max_{x \in E} \mathbb{E}_x[T_A] \\ &\leq \|\rho - \pi\|_{\text{TV}} \frac{\mathbf{t}_{\text{mix}} + \mathbb{E}_\pi[T_A]}{1 - 1/(2e)} \end{aligned} \quad (2.94)$$

und somit

$$\left| \frac{\mathbb{E}_\rho[T_A]}{\mathbb{E}_\pi[T_A]} - 1 \right| \leq \frac{\mathbf{t}_{\text{mix}}/\mathbb{E}_\pi[T_A] + 1}{1 - 1/(2e)} \|\rho - \pi\|_{\text{TV}} \leq \frac{1 + 1/3}{1 - 1/(2e)} \|\rho - \pi\|_{\text{TV}} \quad (2.95)$$

sofern

$$\frac{\mathbf{t}_{\text{mix}}}{\mathbb{E}_\pi[T_A]} \leq \frac{1}{3}$$

gilt (was wir o.E. annehmen dürfen, da ansonsten sich ansonsten in (2.82) ergibt $\Delta > 1$, so dass die Aussage (2.81) von Satz 2.24 trivialerweise gilt).

Einsetzen von (2.95) in (2.93) liefert (2.92).

Setze

$$\alpha := \log(\mathbb{E}_\pi[T_A]/\mathbf{t}_{\text{mix}}), \quad s := (1 + \alpha)\mathbf{t}_{\text{mix}}, \quad (2.96)$$

$$\delta := e^{-\alpha} + \frac{s}{\mathbb{E}_\pi[T_A]} = \frac{\mathbf{t}_{\text{mix}}}{\mathbb{E}_\pi[T_A]} (2 + \log(\mathbb{E}_\pi[T_A]/\mathbf{t}_{\text{mix}})) \quad (2.97)$$

wir zeigen

$$\|\rho - \pi\|_{\text{TV}} \leq 5\delta \quad (2.98)$$

Nach Definition von s in (2.96) gilt zusammen mit (2.89)

$$\|\mathbb{P}_\mu(X_s \in \cdot) - \pi\|_{\text{TV}} \leq e^{-\alpha} \quad \text{für jede Startverteilung } \mu \quad (2.99)$$

(insbesondere für $\mu = \mathbb{P}_\lambda(X_s \in \cdot \mid T_A \geq s)$) und daher (zusammen mit der Markoveigenschaft)

$$\mathbb{P}_\pi(T_A \leq s) - e^{-\alpha} \leq \mathbb{P}_\lambda(s \leq T_A \leq 2s \mid T_A \geq s) \leq \mathbb{P}_\pi(T_A \leq s) + e^{-\alpha} \quad (2.100)$$

für jede Startverteilung λ , was

$$\mathbb{P}_\lambda(T_A > 2s) \leq 1 - \mathbb{P}_\pi(T_A \leq s) + e^{-\alpha} \quad \text{für jede Startverteilung } \lambda \quad (2.101)$$

impliziert (wir nehmen an, dass die rechte Seite von (2.101) < 1 ist, ansonsten ist (2.102) trivialerweise erfüllt). Induktiv folgt (für jedes λ)

$$\mathbb{P}_\lambda(T_A > 2ns) \leq (1 - \mathbb{P}_\pi(T_A \leq s) + e^{-\alpha})^n, \quad n \in \mathbb{N}$$

und daher

$$\mathbb{E}_\pi[T_A] \leq 2s \sum_{n=0}^{\infty} \mathbb{P}_\pi(T_A > 2ns) \leq \frac{2s}{\mathbb{P}_\pi(T_A \leq s) - e^{-\alpha}}$$

Umstellen liefert

$$\mathbb{P}_\pi(T_A \leq s) \leq e^{-\alpha} + \frac{2s}{\mathbb{E}_\pi[T_A]} \quad (2.102)$$

Schließlich seien

$$\rho_0 = \pi, \quad \rho_n = \mathbb{P}_\pi(X_{ns} \in \cdot \mid T_A \geq ns), \quad n \in \mathbb{N}$$

damit ist (zusammen mit (2.100), (2.102) und der Definition von δ aus (2.96))

$$\mathbb{P}_{\rho_{n+1}}(T_A < s) = \mathbb{P}_{\rho_n}(s \leq T_A < 2s \mid T_A \geq s) \leq 2\delta \quad (2.103)$$

Insgesamt ergibt sich

$$\|\rho_{n+1} - \pi\|_{\text{TV}} = \|\mathbb{P}_{\rho_n}(X_s \in \cdot \mid T_A \geq s) - \pi\|_{\text{TV}} \leq e^{-\alpha} + 2\mathbb{P}_{\rho_n}(T_A < s) \leq 5\delta \quad (2.104)$$

(verwende (2.99) und (2.88) für das erste „ \leq “, (2.103) für das zweite „ \leq “).

Mit $n \rightarrow \infty$ gilt $\rho_n \Rightarrow \rho$ (vgl. (2.85)), womit (2.98) folgt.

(2.92) zusammen mit (2.98) und der Definition von δ aus (2.97) liefert (2.81). \square

2.6 Beispiel-Instanzen für Satz 2.22

2.6.1 Moran-Modell

Das Moran-Modell aus Definition 2.1 passt in den Rahmen von Satz 2.22 mit $E_n = \{1, \dots, n\}$, $P = (p(x, y))_{x, y \in E_n}$, $p(x, y) = \frac{1}{n}$.

Es ist

$$\pi^{(n)}(x) = \frac{1}{n}, \quad \mathbb{P}_x(X_t = y) = (1 - e^{-t})\frac{1}{n} + e^{-t}\delta_{xy}, \quad x, y \in E_n \quad (2.105)$$

und somit $\mathbf{t}_{\text{mix}}^{(n)} = O(1)$, weiter gilt für unabhängige Kopien $(X_t), (X'_t)$ einer zeitkontinuierlichen Markovkette mit Generatormatrix $Q = P - I$ und jeweils Startverteilung $\pi^{(n)}$ für $M := \inf\{t \geq 0 : X_t = X'_t\}$

$$\begin{aligned} \mathbf{t}_{\text{meet}}^{(n)} = \mathbb{E}[M] &= \int_0^\infty \mathbb{P}(M > t) dt = \frac{n-1}{n} \int_0^\infty \sum_{j=0}^\infty e^{-2t} \frac{(2t)^j}{j!} (1 - 1/n)^j dt \\ &= \frac{n-1}{n} \int_0^\infty e^{-2t/n} dt = \frac{n-1}{n} \frac{n}{2} \underset{n \rightarrow \infty}{\sim} \frac{n}{2} \end{aligned} \quad (2.106)$$

Somit zeigt Satz 2.22, dass der reskalierte Typenanteilsprozess des (neutralen 2-Typ-)Moran-Modells $\left(Z_t^{(n)}\right)_t = \left(\frac{1}{n}X_{nt/2}^{(n)}\right)_t$ gegen die Wright-Fisher-Diffusion konvergiert, d.h. (2.10) gilt.

Bemerkung. Für $\xi \in \{0, 1\}^{E_n}$ ist $p_1(\xi) = \frac{1}{n} \sum_{i=1}^n \xi(i)$, weiter ist hier $\nu^{(n)}(x, y) = \mathbb{1}_{\{x \neq y\}} n^{-3}$, $\nu^{(1)}(\mathbb{1}) = (n-1)/n^2$, $\bar{\nu}^{(n)}(x, y) = \mathbb{1}_{\{x \neq y\}} \frac{1}{n(n-1)}$,

$$p_{10}(\xi) = \frac{1}{n(n-1)} \sum_{\substack{x, y=1 \\ x \neq y}}^n \xi(x)(1 - \xi(y)) = p_{01}(\xi)$$

und mit $\gamma_n = n/2$ ergibt sich

$$\begin{aligned} \gamma_n \nu^{(1)}(\mathbb{1})(p_{10}(\xi) + p_{01}(\xi)) &= \frac{n-1}{n} \frac{1}{n(n-1)} \sum_{\substack{x, y=1 \\ x \neq y}}^n \xi(x)(1 - \xi(y)) \\ &= \frac{1}{n^2} \sum_{x, y=1}^n \xi(x)(1 - \xi(y)) = p_1(\xi)p_0(\xi) \end{aligned}$$

d.h. die linke Seite von (2.58) ist $\equiv 0$. (Insoweit ist die Anwendung von Satz 2.22 auf das Moran-Modell “overkill.”)

2.6.2 Nächste-Nachbar-Irrfahrt auf dem Torus $(\mathbb{Z}/n\mathbb{Z})^d$ in $d = 2$

Sei $E_n = (\mathbb{Z}/n\mathbb{Z})^2 = \{0, 1, \dots, n\}^2$, die Übergangsmatrix

$$p(x, y) = \begin{cases} \frac{1}{4}, & \text{falls } (x_1 - x_2 \equiv \pm 1 \pmod n \text{ und } y_1 = y_2) \\ & \text{oder } (x_1 = x_2 \text{ und } y_1 - y_2 \equiv \pm 1 \pmod n) \\ 0, & \text{sonst,} \end{cases} \quad (2.107)$$

für $x = (x_1, x_2), y = (y_1, y_2) \in E_n$. Es ist dann $\pi^{(n)}(\cdot) \equiv 1/n^2$, $\pi_{\text{diag}}^{(n)} = 1/n$.

Satz 2.26 (siehe [CCC16, Section 8.1.1, nach Cox [C89]]). *Für die symmetrische nächste-Nachbar-Irrfahrt auf dem zweidimensionalen Torus $(\mathbb{Z}/n\mathbb{Z})^2$ gilt*

$$\mathbf{t}_{\text{mix}}^{(n)} = O(n^2) \quad (2.108)$$

und

$$\mathbf{t}_{\text{meet}}^{(n)} \sim \gamma_n := \frac{1}{\pi} n^2 \log n \quad (2.109)$$

Insbesondere ist (2.56) erfüllt und demnach konvergiert gemäß Satz 2.22 der reskalierte Typenanteilsprozess gegen die Wright-Fisher-Diffusion.

Bemerkung. Satz 2.22 greift nicht (und gilt auch nicht) für den eindimensionalen Fall $E_n = \mathbb{Z}/n\mathbb{Z}$, denn dort wachsen $\mathbf{t}_{\text{mix}}^{(n)}$ und $\mathbf{t}_{\text{meet}}^{(n)}$ beide proportional zu n^2 . Das Limesobjekt wäre ein (Funktional eines) Brownian web auf dem Kreis. Siehe auch [C89, Section 7] und [SSS17].

Spektralzerlegung der Übergangsmatrix der symmetrischen gewöhnlichen Irrfahrt auf $\mathbb{Z}/n\mathbb{Z}$

Wir betrachten zunächst den eindimensionalen Fall (vgl. [Kl, Kap. 18.4, speziell Bsp. 18.24]): Sei nun $P = (p(x, y))_{x, y \in \mathbb{Z}/n\mathbb{Z}}$ mit

$$p(x, y) = \begin{cases} \frac{1}{2}, & \text{falls } x - y \equiv +1 \pmod{n} \text{ oder } x - y \equiv -1 \pmod{n} \\ 0, & \text{sonst} \end{cases}$$

die Übergangsmatrix der symmetrischen gewöhnlichen Irrfahrt auf $\mathbb{Z}/n\mathbb{Z}$. Für $k \in \{0, 1, \dots, n-1\}$ sei $\theta_k := e^{2\pi i k/n}$,

$$\begin{aligned} v^{(k)} &= \frac{1}{\sqrt{n}} (\theta_k^0, \theta_k^1, \dots, \theta_k^{n-1})^T \\ \lambda_k &= \frac{1}{2} \theta_k + \frac{1}{2} \bar{\theta}_k = \cos\left(\frac{2\pi k}{n}\right) \quad \left(= \cos\left(\frac{2\pi}{n} - \frac{2\pi k}{n}\right) = \cos\left(\frac{2\pi(n-k)}{n}\right) = \lambda_{n-k} \right) \end{aligned} \quad (2.110)$$

$v^{(k)}$ ist (Rechts-)Eigenvektor von P zum Eigenwert λ_k :

$$(Pv^{(k)})_x = \frac{1}{2} (\theta_k^{x-1} + \theta_k^{x+1}) = \frac{1}{2\sqrt{n}} (\theta_k^{-1} + \theta_k^1) \theta_k^x = \lambda_k \frac{\theta_k^x}{\sqrt{n}} = \lambda_k v_x^{(k)},$$

zudem ist für $0 \leq k, \ell \leq n-1$ (wir setzen $\langle a, b \rangle = \sum_{x=0}^{n-1} \overline{a_x} b_x$)

$$\langle v^{(k)}, v^{(\ell)} \rangle = \frac{1}{n} \sum_{x=0}^{n-1} \overline{e^{2\pi i k x/n}} e^{2\pi i \ell x/n} = \frac{1}{n} \sum_{x=0}^{n-1} e^{2\pi i (\ell-k)x/n} = \delta_{k\ell}$$

d.h. $v^{(0)}, v^{(1)}, \dots, v^{(n-1)}$ bilden ein ONS von (Rechts-)Eigenvektoren von P . (Da P symmetrisch ist, sind $(v^{(0)})^T, (v^{(1)})^T, \dots, (v^{(n-1)})^T$ ein ONS von Links-Eigenvektoren von P zu denselben Eigenwerten; die Eigenwerte sind reell mit Betrag ≤ 1 ; falls n gerade ist, so

hat die [zeitdiskrete] Kette Periode 2 und ist -1 ein Eigenwert; falls n ungerade ist, haben alle Eigenwerte $\neq 1$ Betrag < 1 ; die Spektrallücke ist $\lambda_0 - \lambda_1 = 1 - \cos(2\pi/n) \sim_{n \rightarrow \infty} 2\pi^2/n^2$.)

$$V = \begin{pmatrix} v^{(0)} & v^{(1)} & \dots & v^{(n-1)} \\ | & | & & | \\ | & | & & | \\ \dots & \dots & & \dots \\ | & | & & | \end{pmatrix} \quad \text{mit Einträgen } V(x, y) = v_x^{(y)} = \frac{1}{\sqrt{n}} \theta_y^x = \frac{1}{\sqrt{n}} e^{2\pi i x y / n}$$

ist unitär ($\overline{V}^T V = I$) und

$$P = V \begin{pmatrix} \lambda_0 & 0 & \dots & 0 \\ 0 & \lambda_1 & 0 & \dots \\ \vdots & 0 & \lambda_2 & 0 \\ & & & \ddots \\ 0 & \dots & 0 & \lambda_{n-1} \end{pmatrix} \overline{V}^T = V \Lambda \overline{V}^T \quad (2.111)$$

somit für $t \geq 0$

$$e^{t(P-I)} = \sum_{m=0}^{\infty} \frac{t^m}{m!} (P-I)^m = V \begin{pmatrix} e^{t(\lambda_0-1)} & 0 & \dots & 0 \\ 0 & e^{t(\lambda_1-1)} & 0 & \dots \\ \vdots & 0 & e^{t(\lambda_2-1)} & 0 \\ & & & \ddots \\ 0 & \dots & 0 & e^{t(\lambda_{n-1}-1)} \end{pmatrix} \overline{V}^T \quad (2.112)$$

Wir erhalten für die (zeitkontinuierliche) symmetrische gewöhnliche Irrfahrt auf $\mathbb{Z}/n\mathbb{Z}$

$$\begin{aligned} \mathbb{P}_x(X_t = y) &= \sum_{z=0}^{n-1} V(x, z) e^{t(\lambda_z-1)} \overline{V}^T(z, y) = \frac{1}{n} \sum_{z=0}^{n-1} e^{2\pi i x z / n} e^{t(\lambda_z-1)} e^{-2\pi i z y / n} \\ &= \frac{1}{n} \sum_{z=0}^{n-1} e^{t(\lambda_z-1)} e^{2\pi i z(x-y)/n} = \frac{1}{n} \sum_{z=0}^{n-1} e^{t(\lambda_z-1)} \cos(2\pi z(x-y)/n) \\ &= \frac{1}{n} + \frac{1}{n} \sum_{z=1}^{n-1} e^{t(\lambda_z-1)} \cos(2\pi z(x-y)/n) \end{aligned} \quad (2.113)$$

Insbesondere ist die stationäre Verteilung $\pi^{(n)}$ uniform auf $\mathbb{Z}/n\mathbb{Z}$ und

$$\begin{aligned} \|\mathbb{P}_x(X_t = \cdot) - \pi^{(n)}\|_{\text{TV}} &= \frac{1}{2n} \sum_{y=0}^{n-1} \sum_{z=1}^{n-1} e^{t(\lambda_z-1)} |\cos(2\pi z(x-y)/n)| \\ &\leq \frac{1}{2} \sum_{z=1}^{n-1} e^{t(\lambda_z-1)} = \frac{1}{2} \sum_{z=1}^{n-1} \exp(-t(1 - \cos(2\pi z/n))) \\ &\leq \sum_{z=1}^{\lfloor n/2 \rfloor} \exp(-t(1 - \cos(2\pi z/n))) \leq \sum_{z=1}^{\lfloor n/2 \rfloor} \exp(-8t \frac{z^2}{n^2}) \end{aligned} \quad (2.114)$$

(verwende $\cos(x) \leq 1 - \frac{2}{\pi^2} x^2$ für $0 \leq x \leq \pi$). Wir finden $\mathbf{t}_{\text{mix}}^{(n)} = O(n^2)$ (siehe auch [LPW, Thm. 5.6] für einen alternativen Zugang mittels Kopplung).

Zur erwarteten Treffzeit (allgemeine Eigenschaften)

Allgemein: $|E| < \infty$, $(X_t)_{t \geq 0}$ irreduzible Markovkette auf E mit Sprungratenmatrix Q , Übergangsmatrix $p_t(x, y) = (e^{tQ})_{x,y}$. Sei $0 \in E$ ein ausgezeichneteter Zustand, $T_0 := \inf\{t \geq 0 : X_t = 0\}$ und

$$F(x, \lambda) := \mathbb{E}_x [e^{-\lambda T_0}], \quad \lambda \geq 0, x \in E \quad (2.115)$$

Zerlegung nach dem ersten Besuch in 0 zusammen mit der starken Markoveigenschaft zeigt

$$p_t(x, 0) = \int_0^t p_{t-u}(0, 0) \mathbb{P}_x(T_0 \in du) \quad (2.116)$$

und somit

$$\begin{aligned} G(x, \lambda) &:= \int_0^\infty e^{-\lambda t} p_t(x, 0) dt = \int_0^\infty e^{-\lambda t} \int_0^t p_{t-u}(0, 0) \mathbb{P}_x(T_0 \in du) dt \\ &= \int_0^\infty \int_0^\infty \mathbb{1}_{\{u \leq t\}} e^{-\lambda(t-u)} p_{t-u}(0, 0) dt e^{-\lambda u} \mathbb{P}_x(T_0 \in du) \\ &= G(0, \lambda) \int_0^\infty e^{-\lambda u} \mathbb{P}_x(T_0 \in du) = G(0, \lambda) F(x, \lambda) \end{aligned} \quad (2.117)$$

d.h.

$$F(x, \lambda) = \frac{G(x, \lambda)}{G(0, \lambda)} = \frac{\int_0^\infty e^{-\lambda t} p_t(x, 0) dt}{\int_0^\infty e^{-\lambda t} p_t(0, 0) dt} \quad (2.118)$$

Somit

$$\begin{aligned} \mathbb{E}_x[T_0] &= -\frac{\partial}{\partial \lambda} F(x, \lambda) \Big|_{\lambda=0^+} = \lim_{\lambda \downarrow 0} \frac{1}{\lambda} \left(1 - \frac{G(x, \lambda)}{G(0, \lambda)} \right) \\ &= \lim_{\lambda \downarrow 0} \frac{\int_0^\infty e^{-\lambda t} (p_t(0, 0) - p_t(x, 0)) dt}{\int_0^\infty \lambda e^{-\lambda t} p_t(0, 0) dt} = \frac{1}{\pi(0)} \int_0^\infty (p_t(0, 0) - p_t(x, 0)) dt \end{aligned} \quad (2.119)$$

(beachte: $\mathbb{E}_x[T_0] < \infty$, denn T_0 hat exponentiell abfallenden Verteilungsschwanz und ist daher eine integrierbare Majorante).

Bemerkung. 1. Das Integral in (2.119) ist (stets) endlich, denn $|p_t(0, 0) - p_t(x, 0)| \leq 4d_E(t)$ konvergiert exponentiell in $t \rightarrow \infty$ gegen 0.

2. Die Funktion $f : E \rightarrow [0, \infty)$, $f(x) = \mathbb{E}_x[T_0]$ erfüllt (zerlege gemäß dem ersten Sprung von x aus)

$$f(x) = \frac{1}{-q(x, x)} + \sum_{y \neq x} \frac{q(x, y)}{-q(x, x)} f(y) \quad \iff \quad Qf(x) = -1, \quad x \in E \setminus \{0\} \quad (2.120)$$

(und $f(0) = 0$).

(2.120) ist ein (diskretes) Poissonproblem, und (2.119) ist eine (Integral-)Darstellung der Lösung von (2.120).

$\mathbf{t}_{\text{mix}}^{(n)}$ und $\mathbf{t}_{\text{meet}}^{(n)}$ für den 2-dimensionalen Torus $(\mathbb{Z}/n\mathbb{Z})^2$: Beweis von Satz 2.26

Sei $X_t = (X_t^{(1)}, X_t^{(2)})_{t \geq 0}$ zeitkontinuierliche Irrfahrt auf $(\mathbb{Z}/n\mathbb{Z})^2$ mit Gesamtsprungrate 1 und diskreten Sprüngen gemäß (2.107). Dann sind $X^{(1)}$ und $X^{(2)}$ unabhängige symmetrische gewöhnliche Irrfahrten auf $\mathbb{Z}/n\mathbb{Z}$ jeweils mit Gesamtsprungrate $1/2$, somit ist mit (2.113)

$$\begin{aligned} \mathbb{P}_{(x_1, x_2)}(X_t = (y_1, y_2)) &= \mathbb{P}_{x_1}(X_{t/2}^{(1)} = y_1) \mathbb{P}_{x_2}(X_{t/2}^{(2)} = y_2) \\ &= \frac{1}{n^2} \sum_{z_1, z_2=0}^{n-1} e^{t(\lambda_{z_1} + \lambda_{z_2} - 2)/2} \cos(2\pi z_1(x_1 - y_1)/n) \cos(2\pi z_2(x_2 - y_2)/n) \end{aligned} \quad (2.121)$$

Mit (2.114) folgt auch

$$\mathbf{t}_{\text{mix}}^{(n)} = O(n^2) \quad (2.122)$$

denn

$$\begin{aligned} & \frac{1}{2} \sum_{y_1, y_2=0}^{n-1} \left| \mathbb{P}_{x_1}(X_{t/2}^{(1)} = y_1) \mathbb{P}_{x_2}(X_{t/2}^{(2)} = y_2) - \frac{1}{n^2} \right| \\ & \leq \frac{1}{2} \sum_{y_1, y_2=0}^{n-1} \left| \mathbb{P}_{x_1}(X_{t/2}^{(1)} = y_1) \right| \left| \mathbb{P}_{x_2}(X_{t/2}^{(2)} = y_2) - \frac{1}{n} \right| + \frac{1}{2} \sum_{y_1, y_2=0}^{n-1} \frac{1}{n} \left| \mathbb{P}_{x_1}(X_{t/2}^{(1)} = y_1) - \frac{1}{n} \right| \\ & = \|\mathbb{P}_{x_1}(X_{t/2}^{(1)} \in \cdot) - \pi\|_{\text{TV}} + \|\mathbb{P}_{x_2}(X_{t/2}^{(2)} \in \cdot) - \pi\|_{\text{TV}} \end{aligned}$$

Sind X, X' unabhängige Kopien der Irrfahrt auf $E_n = (\mathbb{Z}/n\mathbb{Z})^2$, jeweils mit Startverteilung $\pi^{(n)}$ (= uniforme Verteilung auf E_n), so ist $\tilde{X}_t = X_t - X'_t$ ebenfalls symmetrische gewöhnliche Irrfahrt auf E_n (mit Gesamtsprungrate 2) und Startverteilung $\pi^{(n)}$, daher ist (setze (2.121) in (2.119) ein)

$$\begin{aligned} \mathbf{t}_{\text{meet}}^{(n)} &= \sum_{x \in E_n} \pi^{(n)}(x) \mathbb{E}_x \left[\frac{T_0}{2} \right] \\ &= \frac{1}{n^2} \sum_{x_1, x_2=0}^{n-1} \frac{1}{2\pi^{(n)}(0)} \int_0^\infty \frac{1}{n^2} \sum_{z_1, z_2=0}^{n-1} e^{t(\lambda_{z_1} + \lambda_{z_2} - 2)/2} (1 - \cos(2\pi z_1 x_1/n) \cos(2\pi z_2 x_2/n)) dt \\ &= \sum_{z_1, z_2=0}^{n-1} \int_0^\infty e^{t(\lambda_{z_1} + \lambda_{z_2} - 2)/2} \frac{1}{n^2} \sum_{x_1, x_2=0}^{n-1} (1 - \cos(2\pi z_1 x_1/n) \cos(2\pi z_2 x_2/n)) \frac{dt}{2} \end{aligned} \quad (2.123)$$

Für $z_1 \neq 0$ ist $\sum_{x_1=0}^{n-1} \cos(2\pi z_1 x_1/n) = 0$ (und analog falls $z_2 \neq 0$), somit

$$\begin{aligned} \mathbf{t}_{\text{meet}}^{(n)} &= \sum_{\substack{z_1, z_2=0 \\ z_1+z_2>0}}^{n-1} \int_0^\infty e^{-s(2-\lambda_{z_1}+\lambda_{z_2})} ds = \sum_{\substack{z_1, z_2=0 \\ z_1+z_2>0}}^{n-1} \frac{1}{2-\lambda_{z_1}+\lambda_{z_2}} = 2 \sum_{z=1}^{n-1} \frac{1}{1-\lambda_z} + \sum_{z_1, z_2=1}^{n-1} \frac{1}{2-\lambda_{z_1}+\lambda_{z_2}} \\ &= \sum_{z=1}^{\lfloor n/2 \rfloor} \frac{2}{1-\lambda_z} + \sum_{z_1, z_2=1}^{\lfloor n/2 \rfloor} \frac{4}{2-\lambda_{z_1}+\lambda_{z_2}} \end{aligned}$$

$$= \sum_{z=1}^{\lfloor n/2 \rfloor} \frac{2}{1 - \cos(2\pi z/n)} + \sum_{z_1, z_2=1}^{\lfloor n/2 \rfloor} \frac{4}{2 - \cos(2\pi z_1/n) - \cos(2\pi z_2/n)} \quad (2.124)$$

(erinnere (2.110)). Somit ist

$$\begin{aligned} \mathbf{t}_{\text{meet}}^{(n)} &\underset{n \rightarrow \infty}{\sim} \sum_{\substack{z_1, z_2=1 \\ z_1^2 + z_2^2 \geq a_n}}^{\lfloor n/2 \rfloor} \frac{4}{2\pi^2(z_1^2 + z_2^2)/n^2} = \frac{2n^2}{\pi^2} \sum_{\substack{z_1, z_2=1 \\ z_1^2 + z_2^2 \geq a_n}}^{\lfloor n/2 \rfloor} \frac{1}{z_1^2 + z_2^2} \\ &\sim \frac{2n^2}{\pi^2} \int_1^{n/2} \int_1^{n/2} \mathbb{1}_{\{z_1^2 + z_2^2 \geq a_n\}} \frac{1}{z_1^2 + z_2^2} dz_1 dz_2 \\ &= \frac{2n^2}{\pi^2} \int_{\sqrt{a_n}}^{n/2} \int_0^{\pi/2} \frac{1}{r^2} d\varphi r dr + \frac{2n^2}{\pi^2} \int_{n/2}^{n/\sqrt{2}} \int_{\arccos(n/r)}^{\pi/2 - \arccos(n/r)} d\varphi \frac{dr}{r} \\ &= \frac{n^2}{\pi} (\log(n/2) - \log(\sqrt{a_n})) + \frac{2n^2}{\pi^2} \int_{n/2}^{n/\sqrt{2}} \frac{\pi/2 - 2 \arccos(n/r)}{r} dr \sim \frac{1}{\pi} n^2 \log(n) \end{aligned} \quad (2.125)$$

z.B. mit Wahl $a_n = (\log(n))^{1/2}$ (wir benötigen nur $a_n \rightarrow \infty$ und $a_n = o(\log(n))$), dann ist auch $\log(a_n) = o(\log(n))$; der vernachlässigte Teil der Doppelsumme trägt höchstens

$$2 \sum_{z_1=1, z_2=0}^{\sqrt{a_n}} \frac{4}{2 - \cos(2\pi z_1/n) - \cos(2\pi z_2/n)} \leq Cn^2 \sum_{z_1, z_2=0}^{\sqrt{a_n}} 1 = cn^2 a_n = o(n^2 \log(n))$$

bei; weiter beachte $0 \leq \arccos(n/r) \leq \pi/4$ für $n \leq r \leq \sqrt{2}n$, somit ist der Wert des zweiten Integrals positiv und $\leq \int_{n/2}^{n/\sqrt{2}} dr/r = \frac{1}{2} \log 2$.

Präziser argumentiert: Wegen $1 - \cos(x) \geq 2x^2/\pi^2$ für $0 \leq x \leq \pi$ ist

$$0 \leq \sum_{z=1}^{\lfloor n/2 \rfloor} \frac{2}{1 - \cos(2\pi z/n)} \leq \sum_{z=1}^{\lfloor n/2 \rfloor} \frac{1}{4z^2/n^2} \leq \frac{n^2}{4} \sum_{z=1}^{\infty} \frac{1}{z^2} = O(n^2) \quad (2.126)$$

Zu $\varepsilon > 0$ wähle $\delta > 0$, so dass

$$\frac{1 - \varepsilon}{2\pi^2(x^2 + y^2)} \leq \frac{1}{2 - \cos(2\pi x) - \cos(2\pi y)} \leq \frac{1 + \varepsilon}{2\pi^2(x^2 + y^2)} \quad (2.127)$$

für alle $0 < x, y \leq \delta$ gilt. (Dies ist wegen $\cos(x) = 1 - x^2/2 + O(x^3)$ für $x \rightarrow 0$ möglich.)

(2.124) zusammen mit (2.126), (2.127) liefert für genügend großes n

$$\begin{aligned} \mathbf{t}_{\text{meet}}^{(n)} &\leq Cn^2 + \sum_{\substack{z_1, z_2=1 \\ a_n \leq z_1^2 + z_2^2 \leq \delta^2 n^2}}^{\lfloor \delta n \rfloor} \frac{4(1 + \varepsilon)}{2\pi^2(z_1^2 + z_2^2)/n^2} + \sum_{\substack{z_1, z_2=1 \\ \delta^2 n^2 \leq z_1^2 + z_2^2 \leq n^2}}^n \frac{1}{2(z_1^2 + z_2^2)/n^2} \\ &\leq Cn^2 + \frac{2n^2(1 + 2\varepsilon)}{\pi^2} \int_1^{n/2} \int_1^{n/2} \mathbb{1}_{\{a_n \leq z_1^2 + z_2^2 \leq \delta^2 n^2\}} \frac{1}{z_1^2 + z_2^2} dz_1 dz_2 \\ &\quad + n^2 \int_1^n \int_1^n \mathbb{1}_{\{\delta^2 n^2 \leq z_1^2 + z_2^2 \leq n^2\}} \frac{1}{z_1^2 + z_2^2} dz_1 dz_2 \\ &\leq Cn^2 + \frac{2n^2(1 + 2\varepsilon)}{\pi^2} \int_{\sqrt{a_n}}^{\delta n} \int_0^{\pi/2} \frac{1}{r^2} d\varphi r dr + n^2 \int_{\delta n}^n \int_0^{\pi/2} \frac{1}{r^2} d\varphi r dr \end{aligned}$$

$$= Cn^2 + \frac{n^2(1+2\varepsilon)}{\pi} (\log(n) + \log(\delta) - \log(\sqrt{a_n})) - \frac{n^2}{2\pi} \log(\delta) \quad (2.128)$$

und daher

$$\limsup_{n \rightarrow \infty} \frac{\mathbf{t}_{\text{meet}}^{(n)}}{n^2 \log n} \leq \frac{1+2\varepsilon}{\pi} \quad (2.129)$$

Mit analogen Argumenten ergibt sich $\liminf_{n \rightarrow \infty} \mathbf{t}_{\text{meet}}^{(n)} / (n^2 \log n) \geq (1-2\varepsilon)/\pi$, mit $\varepsilon \downarrow 0$ folgt (2.109) \square

Literaturverzeichnis

- [A82] David J. Aldous, Markov chains with almost exponential hitting times, *Stochastic Process. Appl.* 13 (1982), no. 3, 305–310 und Corrigenda: “Markov chains with almost exponential hitting times,” *Stochastic Process. Appl.* 14 (1983), no. 1, 107.
- [B16] Matthias Birkner, *Stochastische Modelle der Populationsbiologie*, Skript zu einer Vorlesung an der Johannes Gutenberg-Universität Mainz, Wintersemester 2015/2016. <http://www.staff.uni-mainz.de/birkner/SMPB1516/smpb1516.pdf>
- [B17] Matthias Birkner, *Stochastik II*, Skript zu einer Vorlesung an der Johannes Gutenberg-Universität Mainz, Wintersemester 2017/18. http://www.staff.uni-mainz.de/birkner/StochII_1718/Stochastik_II_WS1718.pdf
- [B18] Matthias Birkner, *Stochastik III*, Skript zu einer Vorlesung an der Johannes Gutenberg-Universität Mainz, Sommersemester 2018. http://www.staff.uni-mainz.de/birkner/StochIII_18/Stochastik3_SS2018.pdf
- [B] Leo Breiman, *Probability*, Addison-Wesley, 1968.
- [C89] J. Theodore Cox, Coalescing random walks and voter model consensus times on the torus in Z^d , *Ann. Probab.* 17 (1989), no. 4, 1333–1366.
- [CCC16] Yu-Ting Chen, Jihyeok Choi, J. Theodore Cox, On the convergence of densities of finite voter models to the Wright-Fisher diffusion, *Ann. Inst. Henri Poincaré Probab. Stat.* 52 (2016), no. 1, 286–322.
- [CC18] Yu-Ting Chen, J. Theodore Cox, Weak atomic convergence of finite voter models toward Fleming-Viot processes, *Stoch. Proc. Appl.* 128 (2018), 2463–2488.
- [D19] Iulia Dahmer, *Stochastische Populationsmodelle*, Vorlesung an der Johannes Gutenberg-Universität Mainz, Wintersemester 2018/2019.
- [DJ] Evgenij B. Dynkin, Alexander A. Juschkewitsch, *Sätze und Aufgaben über Markoffsche Prozesse*, Springer (1969).
- [EK] Stewart N. Ethier, Thomas G. Kurtz, *Markov processes: characterization and convergence*, Wiley, 1986.
- [JS] Jean Jacod, Albert N. Shiryaev, *Limit theorems for stochastic processes*, Springer, 1987.

- [KW] Götz Kersting, Anton Wakolbinger, *Stochastische Prozesse*, Birkhäuser, 2014.
- [Kl] Achim Klenke, *Wahrscheinlichkeitstheorie*, 2. Aufl., Springer, 2008.
- [LPW] David A. Levin, Yuval Peres, with contributions by Elizabeth L. Wilmer, *Markov Chains and Mixing Times*, (with a chapter on “Coupling from the Past”, by James G. Propp and David B. Wilson), 2nd Ed., AMS 2017. <https://pages.uoregon.edu/dlevin/MARKOV/mcmt2e.pdf>
- [L85] Thomas M. Liggett, *Interacting particle systems*, Springer, 1985.
- [L99] Thomas M. Liggett, *Stochastic interacting systems: contact, voter and exclusion processes*, Springer, 1999.
- [L10] Thomas M. Liggett, *Continuous time Markov processes*, American Mathematical Society, 2010.
- [M58] Patrick Alfred Pierce Moran, Random processes in genetics, *Proc. Cambridge Philos. Soc.* 54, 60–71, (1958).
- [N] James R. Norris, *Markov chains*, Cambridge University Press, 1999.
- [SSS17] Emmanuel Schertzer, Rongfeng Sun, Jan M. Swart, The Brownian web, the Brownian net, and their universality, *Advances in Disordered Systems, Random Processes and Some Applications*, 270–368, Cambridge University Press, 2017. [arXiv:1506.00724](https://arxiv.org/abs/1506.00724)
- [T84] Simon Tavaré, Line-of-descent and genealogical processes, and their applications in population genetics models, *Theoretical Population Biology* 26 (2), 119–164, (1984).

Anhang A

Hintergrundmaterial und Ergänzungen

A.1 Zeitkontinuierlichen Markovketten (mit endlichem Zustandsraum)

Wir stellen hier knapp Material über zeitkontinuierlichen Markovketten zusammen, siehe auch z.B. [Kl, Kap. 17.3] oder Norris [N].

Sei E endliche Menge, $\widehat{p} = (\widehat{p}(x, y))_{x, y \in E}$ stochastische Matrix (d.h. $\widehat{p}(x, y) \geq 0$, $\sum_{y \in E} \widehat{p}(x, y) = 1$ für alle $x \in E$), $\widehat{X} = (\widehat{X}_n)_{n \in \mathbb{N}_0}$ (zeitdiskrete, homogene) \widehat{p} -Markovkette (d.h. $\mathbb{P}(\widehat{X}_0 = x_0, \widehat{X}_1 = x_1, \dots, \widehat{X}_n = x_n) = \mathbb{P}(\widehat{X}_0 = x_0) \widehat{p}(x_0, x_1) \times \dots \times \widehat{p}(x_{n-1}, x_n)$ für $x_0, x_1, \dots, x_n \in E$).

Sei $(M_t)_{t \geq 0}$ Poissonprozess mit Rate $\lambda > 0$, unabhängig von \widehat{X} ,

$$X_t := \widehat{X}_{M_t}, \quad t \in [0, \infty),$$

so ist (\widehat{p}^m bezeichne die m -te Potenz von \widehat{p} , I die $E \times E$ -Identitätsmatrix)

$$\begin{aligned} p_t(x, y) &:= \mathbb{P}(X_t = y \mid X_0 = x) = \sum_{m=0}^{\infty} \mathbb{P}(M_t = m, X_t = y \mid X_0 = x) \\ &= \sum_{m=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^m}{m!} \widehat{p}^m(x, y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \frac{(\lambda t)^m}{m!} ((-I)^n \widehat{p}^m)(x, y) \\ &= \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} \lambda^\ell \sum_{m=0}^{\ell} \binom{\ell}{m} (\widehat{p}^m (-I)^{\ell-m})(x, y) = \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} (\lambda(\widehat{p} - I))^\ell(x, y) \\ &= \sum_{\ell=0}^{\infty} \frac{t^\ell}{\ell!} Q^\ell(x, y) = (e^{tQ})(x, y) \end{aligned} \tag{A.1}$$

(wobei die Matrix $Q = (q_{x,y})_{x,y \in E}$ Einträge $q_{x,y} = \lambda(\widehat{p}(x, y) - \delta_{xy})$ besitzt; obige Reihe konvergiert, denn $\max_{x,y \in E} |Q_{xy}^n| \leq (|E| \max_{x,y \in E} |Q_{x,y}|)^n$; beachte auch, dass \widehat{p}^m und I^n kommutieren) und analoge Rechnungen, die die Unabhängigkeit der Zuwächse von $(M_t)_{t \geq 0}$ ausnutzen, zeigen

$$\mathbb{P}(X_{t_1} = x_1, \dots, X_{t_n} = x_n \mid X_0 = x_0) = \prod_{i=1}^n p_{t_i - t_{i-1}}(x_{i-1}, x_i)$$

für $0 = t_0 < t_1 < \dots < t_n$, $x_0, x_1, \dots, x_n \in E$.

Die Matrix Q heißt die *Sprungratenmatrix* (auch: *Ratenmatrix* oder Q -Matrix) der zeitkontinuierlichen Markovkette X , sie hat die Eigenschaften

$$q_{x,y} \geq 0 \text{ für } x \neq y, \quad \sum_{y \in E, y \neq x} q_{x,y} = -q_{x,x}.$$

Zur Interpretation der Einträge von Q als Sprungraten: Für $x \neq y$ und $h \downarrow 0$ ist

$$\mathbb{P}(X_{t+h} = y \mid X_t = x) = (e^{hQ})(x, y) = Q^0(x, y) + hQ^1(x, y) + O(h^2) = hq_{x,y} + O(h^2).$$

Die Schar von Übergangsmatrizen $P_t = e^{tQ}$, $t \geq 0$ bildet eine Halbgruppe, d.h. $P_{t+s} = P_t P_s$ für $s, t \geq 0$ (gemäß den Rechenregeln für [Matrix-]Exponentiale), eintragsweise ausgesprochen lautet dies

$$p_{t+s}(x, y) = \sum_{z \in E} p_t(x, z) p_s(z, y), \quad x, y \in E \quad (\text{A.2})$$

(die sogenannten Chapman-Kolmogorov-Gleichungen).

A.1.1 Kolmogorovs Vorwärts- und Rückwärtsgleichungen

Sei Q die Sprungratenmatrix einer zeitkontinuierlichen Markovkette auf der endlichen Menge E . Dann gilt

$$\frac{\partial}{\partial t} \exp(tQ) = Q \exp(tQ) = \exp(tQ)Q, \quad (\text{A.3})$$

(wie man sich durch gliedweise Differentiation der Reihe in (A.1) überzeugt), demnach für $t \geq 0$ und $P_t = e^{tQ} = (p_t(x, y))_{x, y \in E}$

$$\frac{\partial}{\partial t} P_t = Q P_t, \text{ d.h. } \forall x, y \in E : \quad \frac{\partial}{\partial t} p_t(x, y) = \sum_z q_{x,z} p_t(z, y) = \sum_z q_{x,z} (p_t(z, y) - p_t(x, y)), \quad (\text{A.4})$$

$$\frac{\partial}{\partial t} P_t = P_t Q, \text{ d.h. } \forall x, y \in E : \quad \frac{\partial}{\partial t} p_t(x, y) = \sum_z p_t(x, z) q_{z,y} = p_t(x, y) q_{y,y} + \sum_{z \neq y} p_t(x, z) q_{z,y}. \quad (\text{A.5})$$

Die Gleichungen (A.4) und (A.5) haben eine stochastische Interpretation: (A.4) heißt Kolmogorovs *Rückwärtsgleichung*, denn es gilt (wir schreiben $\mathbb{P}_x(\cdot)$ für $\mathbb{P}(\cdot \mid X_0 = x)$)

$$\begin{aligned} \frac{p_{t+h}(x, y) - p_t(x, y)}{h} &= \frac{1}{h} (\mathbb{P}_x(X_{t+h} = y) - \mathbb{P}_x(X_t = y)) \\ &= \frac{1}{h} \left(\sum_z \mathbb{P}_x(X_{t+h} = y \mid X_h = z) \mathbb{P}_x(X_h = z) - \mathbb{P}_x(X_t = y) \right) \\ &= \frac{1}{h} \left(\sum_z p_t(z, y) (1_{\{x=z\}} + hq_{x,z} + o(h)) - p_t(x, y) \right) = \sum_z q_{x,z} p_t(z, y) + o(1), \end{aligned}$$

man leitet sie also aus her durch „Rückwärtszerlegung“ des Prozesses X im Intervall $[0, t+h]$ gemäß dem Verhalten am Anfang des Intervalls.

Analog heißt (A.5) Kolmogorovs *Vorwärtsgleichung*, sie entsteht durch Zerlegung gemäß dem Wert bei t :

$$\begin{aligned} \frac{p_{t+h}(x, y) - p_t(x, y)}{h} &= \frac{1}{h} \left(\sum_z \mathbb{P}_x(X_{t+h} = y | X_t = z) \mathbb{P}_x(X_t = z) - \mathbb{P}_x(X_t = y) \right) \\ &= \frac{1}{h} \left(\sum_z p_t(x, z) (1_{\{z=y\}} + h q_{y,z} + o(h)) - p_t(x, y) \right) \\ &= \sum_z q_{x,z} p_t(z, y) + o(1). \end{aligned}$$

Sowohl (A.4) als auch (A.5) sind (im Fall $|E| < \infty$) eindeutig lösbar und beide bestimmen die Halbgruppe von Übergangsmatrizen $(P_t)_{t \geq 0}$ – es sind beides endliche Systeme linearer Differentialgleichungen mit konstanten Koeffizienten.

A.2 Zur Skorokhod-Topologie auf dem Raum der càdlàg-Pfade

Sei E ein polnischer Raum (mit einer Metrik d_E , sagen wir), $D([0, \infty), E)$ die Menge aller Funktionen $f : [0, \infty) \rightarrow E$, die rechtsstetig sind und an jeder Stelle einen linken Limes besitzen (oft mit dem französischen Akronym càdlàg bezeichnet).

Sei $\Lambda := \{\lambda : [0, \infty) \rightarrow [0, \infty) : \lambda \text{ bijektiv und stetig}\}$, die Lipschitz-Konstante von $\lambda(\cdot)$ bezeichnen wir mit $\gamma(\lambda) := \sup_{0 \leq s < t} \frac{\lambda(t) - \lambda(s)}{t - s}$ ($\leq \infty$). Für $f, g \in D([0, \infty), E)$ definiert man die Skorokhod¹-Metrik als

$$d(f, g) := \inf_{\lambda \in \Lambda} \left\{ \gamma(\lambda) \vee \int_0^\infty e^{-t} \sup_{s \geq 0} d_E(f(s \wedge t), g(\lambda(s) \wedge t)) dt \right\} \quad (\text{A.6})$$

Dies ist tatsächlich eine Metrik (d.h.: symmetrisch, $d(f, g) = 0 \iff f = g$, die Dreiecksungleichung gilt), damit ausgestattet ist $D([0, \infty), E)$ ein vollständiger und separabler metrischer Raum, siehe z.B. [EK, Ch. 3.5].

¹nach Anatolii Volodymyrovych Skorokhod, 1930–2011