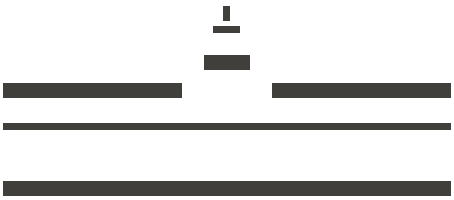

Westfälische
Wilhelms-Universität
Münster

Über die Verteilung der Länge externer Zweige in Koaleszenzbäumen: Genetische Vielfalt innerhalb einer Spezies

Diplomarbeit
von
Thomas Uckelmann

September 2012

Betreut durch Prof. Dr. Gerold Alsmeyer
Institut für Mathematische Statistik

Inhaltsverzeichnis

Einleitung

Der n-Koaleszenzprozess nach Kingman [15] ergibt sich als asymptotisches Modell für die Geneaologie von n Individuen einer haploiden Population der Größe $2N$ unter dem Wright-Fisher Modell. Dabei lässt sich der n-Koaleszenzprozess in natürlicher Weise als baumwertiger Prozess auffassen.

In [17] wurde als mögliches Maß für die genetische Vielfalt einer Population die *Einzigartigkeit eines Individuums* als Zeit bis zum zeitlich nähesten gemeinsamen Vorfahren eines Individuums und seines nähesten Verwandten innerhalb der Population eingeführt.

Für diese wurde unter Betrachtung von Laplace-Transformierten von Caliebe et al. in [7] das Ergebnis geliefert, dass sie asymptotisch einer Verteilung mit Dichte $x \to 8/(2+x)^3$, $x \geq 0$, genügt. Genauer wurde gezeigt, dass die Länge Z_n eines zufällig gewählten externen Zweiges eines Koaleszenzbaumes der Konvergenzeigenschaft $nZ_n \xrightarrow{\mathrm{d}} Z$ genügt, wobei die Dichte von Z gerade der eben gegebenen entspricht.

Eine unweigerliche Konsequenz ist damit, dass die genetische Vielfalt einer Population auf eine geringe Anzahl Individuen konzentriert ist. Insbesondere kann daher der Verlust einzelner Individuen große Auswirkungen auf die genetische Vielfalt der Population haben.

Eine andere Betrachtungsweise ergibt sich mit Blick auf [12]. Dort wurde die Länge aller externen Zweige im Verhältnis zur Länge der internen Zweige untersucht. Die dahinterliegende Frage ist, ob die Annahme von neutraler Selektion für eine gegebene Population mit Mutationseinfluss gerechtfertigt ist. Die Länge aller externen Zweige korrespondiert dabei zu der Zeit, in der Mutationen auftretenen können, welche nur ein Individuum betreffen können.

In [13] wurde gezeigt, dass für die Gesamtlänge L_n der externen Zweige eines Koaleszenzprozesses nach Kingman mit n externen Zweigen die Verteilungskonvergenz

$$\frac{1}{2}\sqrt{\frac{n}{\ln n}}\,(L_n - 2) \xrightarrow{\mathrm{d}} \mathrm{N}(0,1)$$

für $n \to \infty$ gilt, wobei wie üblich mit $\mathrm{N}(0,1)$ eine Normalverteilung mit Erwartung 0 und Varianz 1 bezeichnet sei.

Vermöge der Unabhängigkeit des Koaleszenzprozesses und des Mutationsprozesses in einem Koaleszenzprozess mit Mutation, kann die Länge der externen Zweige herangezogen werden, um die Anzahl von Mutationen auf den externen Zweigen zu bestimmen. Die Länge eines zufällig gewählten externen Zweiges korrespondiert unter dieser Betrachtung mit der Zeit, in der Mutationen das ausgewählte Individuum im Vergleich zu seinem nähesten Verwandten verändern können.

Entsprechend des asymptotischen Ergebnisses für die Länge eines externen Zweiges lässt sich die Verteilung der Anzahl der Mutationen auf einem externen Zweig angeben.

In dieser Arbeit wollen wir das Ergebnis von Caliebe et al. rekapitulieren.

Ich danke Herrn Prof. Dr. Gerold Alsmeyer für diesen Themenvorschlag und im Besonderen für seine Geduld mit mir. Desweiteren möchte ich mich bei meiner Familie für die Unterstützung während meiner Studienzeit bedanken.

1. Allgemeine Grundlagen

Dieses Kapitel dient dazu, einige grundlegende Definitionen und Ergebnisse der Theorie der Markov-Sprungprozesse anzugeben, auf die im Folgenden verwiesen wird. Darüber hinaus werden wir größtenteils auf Beweise verzichten.

Bevor wir damit allerdings beginnen, wollen wir noch einige allgemeinere Definitionen treffen und halten entsprechend fest, dass wir mit einer Exponentialverteilung mit Parameter λ, kurz $\mathrm{Exp}(\lambda)$, eine Verteilung bezeichnen, deren Verteilungsfunktion durch $(1-e^{-\lambda x})\mathbb{1}_{[0,\infty)}(x)$ gegeben ist. Insbesondere berechnet sich der Erwartungswert in diesem Fall zu $1/\lambda$. Ebenfalls wollen wir $\mathcal{O}$ in der üblichen Bedeutung des Landau-Kalküls verstehen.

1.1 Definition
Ein topologischer Raum $(S, \mathfrak{T})$ heißt *polnisch*, wenn eine die Toplogie $\mathfrak{T}$ erzeugende Metrik ϱ existiert, so dass (S, ϱ) ein separabler, vollständiger metrischer Raum ist.

1.2 Definition
Sei (M, d) ein metrischer Raum und bezeichne $\mathbb{R}_+ = [0, \infty)$. Dann heißt eine Funktion $f : \mathbb{R}_+ \to M$ *càdlàg (aus dem Französischen »continue à droite, limite à gauche«)*, falls

(i) $f(t) = \lim_{s \downarrow t} f(s)$ für alle $t \geq 0$, und

(ii) $f(t-) = \lim_{s \uparrow t} f(s)$ für alle $t > 0$ existiert.

f ist also rechtsseitig stetig und besitzt linksseitige Limiten.

Wir nennen ferner einen stochastischen Prozess $(X_t)_{t \geq 0}$ *càdlàg*, falls dies für seine Pfade der Fall ist.

1.3 Definition
Gegeben einen Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ mit einer Filtration $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$, heißt eine nichtnegative Zufallsvariable τ *Stoppzeit bezüglich* $\mathcal{F}$, falls

$$\{\tau \leq t\} \in \mathcal{F}_t, \tag{1.1}$$

für alle $t \geq 0$ gilt. Wir bezeichnen ferner mit

$$\mathcal{F}_\tau \stackrel{\text{def}}{=} \{A \in \mathfrak{A} : A \cap \{\tau \leq t\} \in \mathcal{F}_t\} \tag{1.2}$$

die σ-Algebra der τ-Vergangenheit.

1.1. Markov-Prozesse in stetiger Zeit

Für die allgemeine Betrachtung von Markov-Prozessen in stetiger Zeit in [1], auf die wir hier für die Beweise der Sätze verweisen, setzen wir voraus, dass der Zustandsraum $(S, \mathfrak{S})$ eines stochastischen Prozesses $X = (X_t)_{t\geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ lokalkompakt und mit abzählbarer Basis sei, sowie $\mathfrak{S}$ die zugehörigen Borelschen σ-Algebra bezeichne. Da damit $(S, \mathfrak{S})$ insbesondere polnisch ist, existieren die im Folgenden auftretenden (regulär) bedingten Verteilungen (vgl. Satz 53.4 in [2]).

1.4 Definition (Markov-Prozess)
Sei $X = (X_t)_{t\geq 0}$ ein stochastischer Prozess auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ mit Zustandsraum $(S, \mathfrak{S})$ und $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ eine Filtration.

X heißt *Markov-Prozess bezüglich* $\mathcal{F}$, wenn er $\mathcal{F}$-adaptiert ist und die *Markov-Eigenschaft*, gegeben durch

$$\mathrm{P}(X_t \in A|\mathcal{F}_s) = \mathrm{P}(X_t \in A|X_s) \quad \text{P-f.s.} \tag{1.3}$$

für alle $0 \leq s < t < \infty$ und $A \in \mathfrak{S}$, besitzt. Ferner bezeichne

$$\mathbb{P}_{s,t}(X_s, A) \stackrel{\text{def}}{=} \mathrm{P}(X_t \in A|X_s) \tag{1.4}$$

für alle $0 \leq s < t < \infty$ und $A \in \mathfrak{S}$ die zugehörigen Übergangskerne. Falls diese so gewählt werden können, dass sie lediglich über $s - t$ von s und t abhängen, das heißt, falls

$$\mathrm{P}(X_t \in A|\mathcal{F}_s) = \mathrm{P}(X_t \in A|X_s) = \mathbb{P}_{0,t-s}(X_s, A) \quad \text{P-f.s.}$$

gilt, so heißt X *zeitlich homogen*.

Im Fall der natürlichen Filtration $\mathcal{G} = (\mathcal{G}_t)_{t\geq 0}$, mit $\mathcal{G}_t = \sigma(X_s, s \leq t)$, verzichten wir auf den Zusatz *bezüglich* $\mathcal{G}$.

Ist der Zustandsraum $(S, \mathfrak{S})$ abzählbar, so nennen wir X einen *Markov-Sprungprozess* oder auch *zeitstetige Markovkette*.

1.5 Definition
Wir sagen, ein Markov-Prozess $X = (X_t)_{t\geq 0}$ bezüglich einer Filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ mit Zustandsraum $(S, \mathfrak{S})$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$, erfülle die *starke Markov-Eigenschaft bezüglich* τ, wenn, bedingt unter $\{\tau < \infty\}$,

$$\mathrm{P}(X_{\tau+t} \in A|\mathcal{F}_\tau) = \mathrm{P}(X_{\tau+t} \in A|X_\tau) = \mathbb{P}_{\tau,\tau+t}(X_\tau, A) \quad \mathrm{P\text{-f.s.}} \tag{1.5}$$

für alle $A \in \mathfrak{S}$ und $t \geq 0$ für eine Stoppzeit τ bezüglich $\mathcal{F}$ gilt. Im Fall eines homogenen Prozesses vereinfacht sich die Gestalt von (1.5) zu

$$\mathrm{P}(X_{\tau+t} \in A|\mathcal{F}_\tau) = \mathbb{P}_{0,t}(X_\tau, A) \quad \mathrm{P\text{-f.s.}} \tag{1.6}$$

1.6 Bemerkung
Die Definition der *starken Markov-Eigenschaft* für beliebige Markov-Prozesse in stetiger Zeit verwendet rechtsseitig stetige σ-Algebren, beziehungsweise rechtsseitig stetige Erweiterungen (vgl. Seite 57 in [8]). Für einen Markov-Prozess $X = (X_t)_{t\geq 0}$ bezüglich einer rechtsseitig stetigen Filtration $\mathcal{F} = (\mathcal{F}_t)_{t\geq 0}$ sagen wir dann, dass dieser die *starke Markov-Eigenschaft* besitze, falls für jede Stoppzeit τ bezüglich $\mathcal{F}$ bedingt unter $\{\tau < \infty\}$ (1.5) gilt.

Die im Folgenden betrachteten Markov-Prozesse seien nunmehr stets zeitlich homogen. Wir führen deshalb $\mathbb{P}_t = \mathbb{P}_{0,t}$ als abkürzende Schreibweise für die Übergangskerne ein. Darüber hinaus setzen wir $\mathbb{P}_0(x, \cdot) \overset{\mathrm{def}}{=} \delta_x$, wobei δ_x die Einpunktverteilung in x bezeichne.

Für die Übergangskerne eines homogenen Markov-Prozesses gelten die *Kolmogorov-Chapman-Gleichungen* als einfache Konsequenz der Rechenregeln für stochastische Kerne:

1.7 Lemma
Für die Familie $(\mathbb{P}_t)_{t\geq 0}$ der Übergangskerne eines zeitlich homogenen Markov-Prozesses gelten die *Kolmogorov-Chapman-Gleichungen*, das heißt

$$\mathbb{P}_{s+t} = \mathbb{P}_s \mathbb{P}_t. \tag{1.7}$$

$(\mathbb{P}_t)_{t\geq 0}$ bildet damit eine Halbgruppe (bezüglich der Hintereinanderschaltung von Kernen).

Wir weisen ebenfalls darauf hin, dass zu jeder Halbgruppe $(\mathbb{P}_t)_{t\geq 0}$ von Übergangskernen auf $(S, \mathfrak{S})$ mit $\mathbb{P}_0(x, \cdot) = \delta_x$ ein Markov-Prozess existiert. Darüber hinaus kann die Anfangsverteilung $\lambda \in \mathfrak{W}(S)$, wobei $\mathfrak{W}(S)$ die Menge der Wahrscheinlichkeitsmaße auf S

bezeichne, beliebig gewählt werden. Dies ergibt sich als Anwendung des Konsistenzsatzes von Daniell-Kolmogorov.

Insbesondere erlaubt uns dies die Betrachtung von Räumen

$$\left(\Omega, \mathfrak{A}, (X_t)_{t\geq 0}, (\mathrm{P}_\lambda)_{\lambda\in\mathfrak{W}(S)}\right), \tag{1.8}$$

in denen $X = (X_t)_{t\geq 0}$ unter jedem P_λ ein Markov-Prozess mit Übergangskernen $\mathbb{P}_t$ und Anfangsverteilung λ ist. Wir nennen dann (1.8) ein *Standardmodell* für $(\mathbb{P}_t)_{t\geq 0}$.

1.8 Bemerkung
Es bezeichne $b\mathfrak{S}$ den Raum der beschränkten Funktionen $f : \mathfrak{S} \to \mathfrak{S}$, versehen mit der Supremumsnorm $\|\cdot\|_\infty$. Wir setzen ferner

$$\mathbb{P}_t f(x) \stackrel{\text{def}}{=} \int_S f(y)\mathbb{P}_t(x, dy) \tag{1.9}$$

für alle $f \in b\mathfrak{S}$. Vermöge der $\mathfrak{A}$-Messbarkeit von $x \to \mathbb{P}_t(x, A)$ für alle $A \in \mathfrak{S}$, definiert dies einen linearen Operator von $b\mathfrak{S}$ nach $b\mathfrak{S}$. Für diesen Operator gilt ferner $\mathbb{P}_t f \geq 0$ für alle nichtnegativen $f \in b\mathfrak{S}$, $\|\mathbb{P}_t f\|_\infty \leq \|f\|_\infty$ für alle $f \in b\mathfrak{S}$, das heißt, $\mathbb{P}_t$ ist eine *positive Kontraktion*, und man nennt $(\mathbb{P}_t)_{t\geq 0}$ deshalb auch eine *positive Kontraktionshalbgruppe*.

1.9 Definition
Eine Übergangshalbgruppe $(\mathbb{P}_t)_{t\geq 0}$ heißt *stochastisch*, falls $\|\mathbb{P}_t\| = 1$ für alle $t \geq 0$ gilt und andernfalls *substochastisch*. Dabei bezeichnet $\|\cdot\|$ die Operatornorm, definiert durch

$$\|\mathbb{P}_t\| \stackrel{\text{def}}{=} \sup\{\|\mathbb{P}_t f\| : \|f\|_\infty = 1\}. \tag{1.10}$$

1.2. Markov-Sprungprozesse

Sei nunmehr der Zustandsraum $(S, \mathfrak{S})$ abzählbar. Für einen Markov-Sprungprozess gemäß 1.4 stellen wir fest:

1.10 Definition
Gegeben einen Markov-Sprungprozess $X = (X_t)_{t\geq 0}$ auf einem abzählbaren Zustandsraum $(S, \mathfrak{S})$ mit Übergangshalbgruppe $(\mathbb{P}_t)_{t\geq 0}$, wird dieser bereits eindeutig durch $p_{ij}(t) \stackrel{\text{def}}{=} \mathbb{P}_t(i, \{j\})$ für $i, j \in S$ festgelegt. Wir definieren deshalb

$$\mathbf{P}(t) \stackrel{\text{def}}{=} (p_{ij}(t))_{i,j\in S} \tag{1.11}$$

für alle $t \geq 0$ und nennen $(\mathbf{P}(t))_{t\geq 0}$ die zu X gehörende *Übergangsmatrixfunktion*.

1.11 Bemerkung
Für die Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$ eines Markov-Sprungprozesses $X = (X_t)_{t\geq 0}$ gelten

(i) $p_{ij}(t) \geq 0$ für alle $t \geq 0$ und $i, j \in S$, sowie $p_{ij}(0) = \delta_{ij}$, wobei δ_{ij} das Kronecker-Delta bezeichnet, und damit $\mathbf{P}(0) = (\delta_{ij})_{i,j\in S} = I$,

(ii) $\sum_{j\in S} p_{ij}(t) \leq 1$ für alle $t \geq 0$ und $i, j \in S$,

(iii) $p_{ij}(s+t) = \sum_{k\in S} p_{ik}(s)p_{kj}(t)$, genannt *Kolmogorov-Chapman-Gleichungen*; die Halbgruppeneigenschaft lässt sich also bezüglich der gewöhnlichen Matrix-Multiplikation als

$$\mathbf{P}(s+t) = \mathbf{P}(s)\mathbf{P}(t) \tag{1.12}$$

darstellen.

1.12 Definition
Eine Matrix $\mathbf{P} = (p_{ij})_{i,j\in S}$ heißt *stochastisch*, falls

$$p_{ij} \geq 0, \text{ und } \sum_{j\in S} p_{ij} = 1 \tag{1.13}$$

für alle $i, j \in S$ gilt.

Dementsprechend nennen wir eine Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$ eines Markov-Sprungprozesses *stochastisch*, falls $\mathbf{P}(t)$ für alle $t \geq 0$ stochastisch ist. Andernfalls nennen wir $(\mathbf{P}(t))_{t\geq 0}$ *substochastisch*.

1.13 Definition
Gegeben eine Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$, nennen wir diese eine *Standard-Übergangsmatrixfunktion*, falls ihre einzelnen Komponenten stetig in 0 sind, das heißt, falls

$$\lim_{t\downarrow 0} p_{ij}(t) = p_{ij}(0) = \delta_{ij} \tag{1.14}$$

für alle $i, j \in S$ gilt.

1.14 Satz
Gegeben eine Standard-Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$, so ist jedes $p_{ij}(t)$ für alle $t > 0$ stetig differenzierbar. Für $t = 0$ gilt zumindet noch die (rechtsseitige) Differenzierbarkeit. Bezeichnet nun

$$q_{ij} \stackrel{\text{def}}{=} \lim_{t\downarrow 0} t^{-1}(p_{ij}(t) - p_{ij}(0)), \tag{1.15}$$

so gilt ferner $|q_{ij}| < \infty$, falls $i \neq j$ und $q_{ii} < \infty$ (kann aber den Wert $-\infty$ annehmen).

1.15 Definition
Gegeben eine Standard-Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$, eines Markov-Sprungprozesses $X = (X_t)_{t\geq 0}$, setzen wir

$$\mathbf{Q} \stackrel{\text{def}}{=} (q_{ij})_{i,j\in S} \tag{1.16}$$

und bezeichnen $\mathbf{Q}$ als *Q-Matrix* von X.

Gegeben eine Q-Matrix $\mathbf{Q}$ eines Markov-Sprungprozesses $X = (X_t)_{t\geq 0}$, gilt $q_{ij} \geq 0$ für alle $i, j \in S$ mit $i \neq j$ vermöge $t^{-1}(p_{ij}(t) - p_{ij}(t)) = t^{-1}p_{ij}(t) \geq 0$ für alle $t \geq 0$. Eine ähnliche Überlegung liefert darüber hinaus auch $q_{ii} \leq 0$. Setze nunmehr

$$q_i \stackrel{\text{def}}{=} -q_{ii} \text{ für alle } i \in S. \tag{1.17}$$

Eine Anwendung von Fatous Lemma bezüglich des Zählmaßes auf den Paaren (i, j) mit $i \neq j$ liefert ferner

$$\sum_{i\neq j} q_{ij} = \sum_{i\neq j} \lim_{t\downarrow 0} \frac{p_{ij}(t)}{t} \leq \liminf_{t\downarrow 0} \sum_{i\neq j} \frac{p_{ij}(t)}{t} = \lim_{t\downarrow 0} \frac{1 - p_{ii}(t)}{t} = q_i. \tag{1.18}$$

1.16 Definition
Eine Q-Matrix $\mathbf{Q}$ heißt *konservativ*, falls

$$\sum_{j\neq i} q_{ij} = q_i < \infty \tag{1.19}$$

für alle $i \in S$ gilt.

1.17 Satz
Gegeben eine Standard-Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$ mit konservativer Q-Matrix $\mathbf{Q}$. Dann ist $\mathbf{Q}$ der Generator der Halbgruppe mit Definitionsbereich

$$\mathcal{D}(\mathbf{Q}) = \left\{ f \in b\mathfrak{S} : \exists g \in b\mathfrak{S} : \lim_{t\downarrow 0} \|t^{-1}(\mathbf{P}(t)f - f) - g\|_\infty = 0 \right\}.$$

1.18 Satz
Jeder càdlàg-Markov-Sprungprozess, das heißt jeder Markov-Sprungprozess mit rechtsseitig stetigen, stückweise konstanten Pfaden, besitzt die starke Markov-Eigenschaft.

1.19 Definition
Für einen Markov-Sprungprozess $X = (X_t)_{t\geq 0}$ auf einem Zustandsraum $(S, \mathfrak{S})$ mit Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$ und zugehöriger Q-Matrix $\mathbf{Q} = (q_{ij})_{i,j\in S}$ und $q_i = -q_{ii}$ für $i \in S$, bezeichnen wir einen Zustand $i \in S$ als

(i) *stabil*, falls $0 < q_i < \infty$,

(ii) *absorbierend*, falls $q_i = 0$,

(iii) *augenblicklich*, falls $q_i = \infty$.

Darüber hinaus bezeichnen wir eine Übergangsmatrixfunktion als *stabil*, wenn alle Zustände $i \in S$ stabil sind.

In konservativen Markov-Sprungprozessen gibt es keine augenblickliche Zustände.

1.20 Definition
Ein konservativer Markov-Sprungprozess $X = (X_t)_{t\geq 0}$ mit Werten in $\mathbb{N}_0$ heißt *Geburts- und Todesprozess* (in stetiger Zeit), falls er von jedem Zustand n lediglich in die Nachbarzustände $n-1, n+1$ springen kann. Die dazugehörige Q-Matrix ist von der Gestalt

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdots \\ \vdots & & & & \ddots & \end{pmatrix},$$

wobei $\lambda_n \in [0,\infty), n \geq 0$ die *Geburtsraten* und $\mu_n \in [0,\infty), n \geq 1$ die *Sterberaten* bezeichnen. Entsprechend bezeichnen wir einen Geburts- und Todesprozess als *(reinen) Geburtsprozess*, falls sämtliche Sterberaten $\mu_n = 0$ sind, beziehungsweise als *(reinen) Todesprozess*, falls sämtliche Geburtsraten $\lambda_n = 0$ sind. Im letzten Fall ist $\mathbf{Q}$ von der Gestalt

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -\mu_1 & 0 & 0 & \cdots \\ 0 & \mu_2 & -\mu_2 & 0 & \cdots \\ \vdots & & & \ddots & \end{pmatrix}.$$

Insbesondere ist der Zustand 0 ein absorbierender Zustand für einen reinen Todesprozess.

1.3. Die minimale Konstruktion

Im Hinblick auf 1.14, stellt sich die Frage, ob auch die Umkehrung gilt, sprich, ob zu gegebener konservativer Q-Matrix $\mathbf{Q}$ ein Markov-Prozess $X = (X_t)_{t\geq 0}$ existiert, dessen Q-Matrix gerade durch $\mathbf{Q}$ gegeben ist.

1.21 Definition
Für einen Markov-Sprungprozess $X = (X_t)_{t\geq 0}$, definiere die *Sprungzeiten*

$$\sigma_0 \stackrel{\text{def}}{=} 0, \quad \sigma_n \stackrel{\text{def}}{=} \inf\{t > \sigma_{n-1} : X_t \neq X_{\sigma_{n-1}}\} \tag{1.20}$$

für $n \geq 1$, sowie die *Absorptionszeit*

$$\varrho_A \stackrel{\text{def}}{=} \sup\{\sigma_k : \sigma_k < \infty\} \tag{1.21}$$

und die sukzessiven Eintrittszeiten in eine Menge $A \in S$

$$\sigma_0(A) \stackrel{\text{def}}{=} 0, \quad \sigma_n(A) \stackrel{\text{def}}{=} \inf\{\sigma_k > \sigma_{n-1}(A) : X_{\sigma_{k-1}} \in A^c, X_{\sigma_k} \in A\}. \tag{1.22}$$

1.22 Satz
Für einen Markov-Sprungprozess $X = (X_t)_{t\geq 0}$, mit rechtsseitig stetigen, stückweise konstanten Pfaden sind die in (1.20) und (1.22) definierten Zufallsvariablen, sowie die *erste Explosionszeit*, definiert als $\sigma_\infty = \sup_{n\geq 1} \sigma_n$, Stoppzeiten.

1.23 Satz
Gegeben einen Markov-Sprungprozess $X = (X_t)_{t\geq 0}$ in einem Standardmodell mit Standard-Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$, konservativer Q-Matrix $\mathbf{Q}$ und kanonischer Filtration $\mathcal{F}$, sei ferner vorausgesetzt, dass $\sigma_\infty = \sup_{n\geq 1} \sigma_n = \infty$ gelte. Definiere weiter

$$\tau_n \stackrel{\text{def}}{=} (\sigma_n - \sigma_{n-1})\mathbb{1}_{\{\sigma_{n-1}<\infty\}} + \infty \mathbb{1}_{\{\sigma_{n-1}=\infty\}} \tag{1.23}$$

$$\hat{X}_n \stackrel{\text{def}}{=} X_{\sigma_n}\mathbb{1}_{\{\sigma_n<\infty\}} + X_{\varrho_A}\mathbb{1}_{\{\sigma_n=\infty\}}. \tag{1.24}$$

Dann existiert eine Übergangsmatrix $\hat{\mathbf{P}} = (\hat{p}_{ij})_{i,j\in S}$ mit

$$\hat{p}_{ii} = \begin{cases} 0, & \text{falls } q_i \in (0,\infty) \\ 1, & \text{falls } q_i = 0, \end{cases}$$

so dass für alle $n \in \mathbb{N}_0$, $j \in S$ und $t \geq 0$

$$\mathrm{P}(\hat{X}_{n+1} = j, \tau_{n+1} > t | \mathcal{F}_{\sigma_n}) = \sum_{i\in S} \hat{p}_{ij} e^{-q_j t}\mathbb{1}_{\{\hat{X}_n=i\}} \quad \mathrm{P}_\lambda\text{-f.s} \tag{1.25}$$

für alle $\lambda \in \mathfrak{W}(S)$ gilt. Insbesondere bildet $\hat{X}$ unter jedem P_λ eine diskrete Markov-Kette mit Zustandsraum S, Übergangsmatrix $\hat{\mathbf{P}}$ und Startverteilung λ. Ferner sind $\tau_1, \tau_2, \ldots$ bedingt unter $\hat{X}$ stochastisch unabhängig und genügen jeweils einer Exponentialverteilung mit Parameter $q_{\hat{X}_{n-1}}$, das heißt $\tau_n \sim \mathrm{Exp}(q_{\hat{X}_{n-1}})$ für alle $n \in \mathbb{N}$.

1.24 Satz
Gegeben die Voraussetzungen des vorangegangen Satzes ist die Übergangsmatrix $\hat{\mathbf{P}}$ der eingebetteten Markov-Kette $(\hat{X}_n)_{n\geq 0}$ wie folgt durch $\mathbf{Q}$ bestimmt: Falls $0 < q_i < \infty$, gilt

$$\hat{p}_{ii} = 0 \text{ und } \hat{p}_{ij} = q_{ij}/q_i \text{ für } i \neq j. \tag{1.26}$$

Im Fall $q_i = 0$ gilt $\hat{p}_{ij} = \delta_{ij}$ für alle $j \in S$.

Mithilfe der Sätze 1.23 und 1.24 lässt sich nunmehr zu gegebener konservativer Q-Matrix $\mathbf{Q}$ ein Verfahren zur Konstruktion eines Markov-Sprungprozess $X = (X_t)_{t\geq 0}$, für den $\mathbf{Q}$ die Q-Matrix darstellt, angeben. Man konstruiert eine diskrete Markov-Kette $\hat{X} = (\hat{X}_n)_n$ mit Übergangsmatrix $\hat{\mathbf{P}}$ gemäß 1.24, sowie eine Folge $(\tau_n)_{n\geq 1}$ von exponentialverteilten Verweildauern, $\tau_n \sim \mathrm{Exp}(q_{\hat{X}_n})$, die bedingt unter $\hat{X}$ stochastisch unabhängig sind. Man definiert dann $\sigma_0 = 0$, $\sigma_n = \tau_1 + \tau_2 + \cdots + \tau_n$ für die Sprungzeiten und setzt

$$X_t \overset{\text{def}}{=} \hat{X}_n, \text{ für alle } t \in [\sigma_n, \sigma_{n+1}) \tag{1.27}$$

und $X_t \overset{\text{def}}{=} \hat{X}_n$, falls $\sigma_n = \infty$. Dabei kann es allerdings passieren, dass unendlich viele Sprünge in endlicher Zeit auftreten. Man spricht in diesem Fall von *Explosion*, charakterisiert durch $\mathrm{P}_{\delta_i}(\sup_{n\geq 1} \sigma_n < \infty) > 0$ für ein $i \in S$.

Bei der *minimalen Konstruktion* wird deshalb der Zustandsraum um einen absorbierenden Punkt $\Delta \notin S$ erweitert und der Prozess zum Explosionszeitpunkt in diesem »Friedhof« geparkt. Dies liefert dann ebenfalls einen Markov-Sprungprozess mit rechtsseitig stetigen, stückweise konstanten Pfaden.

Da Explosion in unserem Fall später nicht auftritt, verzichten wir an dieser Stelle auf diese Ausführung, notieren aber zumindest noch, warum der Begriff der minimalen Konstruktion gerechtfertigt ist, ehe wir Bedingungen an $\mathbf{Q}$ angeben, die sicherstellen, dass die minimale Konstruktion nicht-explodierend ist.

1.25 Satz
Für jede zu $\mathbf{Q}$ gehörende substochastische Standard-Übergangsmatrixfunktion $\tilde{\mathbf{P}}(t) = (\tilde{p}_{ij}(t))_{i,j\in S}$ gilt

$$\tilde{p}_{ij}(t) \geq p_{ij}(t) \tag{1.28}$$

für alle $i, j \in S$ und $t \geq 0$. Ist $\mathbf{P}(t) = (p_{ij})_{i,j\in S}$ stochastisch, gilt also $p_{i\Delta}(t) = 0$ für alle $i \in S$ und $t \geq 0$, so ist $\mathbf{P}(t)$ die einzige zu $\mathbf{Q}$ gehörende substochastische Standard-Übergangsmatrixfunktion.

1.26 Satz (Reuters Explosionskriterium)
Die minimale Konstruktion ist genau dann nicht-explodierend, wenn $x = 0$ die einzige nichtnegative und beschränkte Lösung der Gleichung $\mathbf{Q}x = x$ bildet.

1.27 Proposition
Hinreichende Bedingungen dafür, dass die minimale Konstruktion nicht-explodierend ist, sind

(i) Der Zustandsraum S ist endlich,

(ii) $\sup_{i \in S} q_i < \infty$,

(iii) Die eingebettete diskrete Markovkette ist rekurrent.

2. ATGC's of life

Um eine Motivation für die Betrachtung von Koaleszenzbäumen zu liefern, soll nach einem oberflächlichen Einblick in die Genetik, inklusive einiger grundlegender Begriffsbildungen, auf das Wright-Fisher-Modell eingegangen werden.

Wir folgen Durrett [9] in seinem gleichnamigen Kapitel in der Feststellung, dass das Erbgut der meisten Organismen in der Desoxyribonukleinsäure (DNS) kodiert ist. Diese besteht im Regelfall aus zwei zu einer Doppelhelix verwobenen, komplementären Strängen, bestehend aus einer Sequenz aus vier verschiedenen Nukleotiden. Ein Nukleotid besteht dabei aus einer der vier verschiedenen Nukleobasen Adenin (A), Guanin (G), Cytosin (C) und Thymin (T), einer Phosphorsäure (P) und einem Zucker, im Fall der DNS, der Desoxyribose (dR). Ferner verbinden sich stets Adenin mit Thymin (über 2 Wasserstoffbrücken), sowie Cytosin mit Guanin (über drei Wasserstoffbrücken).

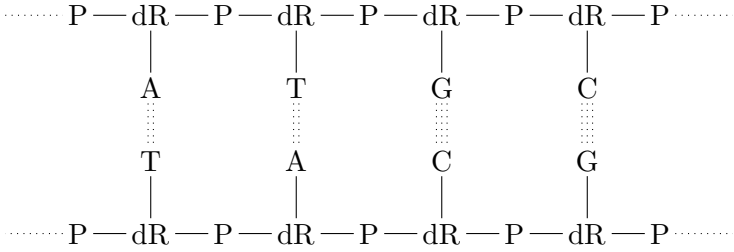

Abbildung 2.1.: Vereinfachte Darstellung eines (inneren) DNS-Abschnittes.

Das grundlegende Prinzip der Vervielfältigung von DNS besteht in der Aufteilung der Doppelhelix in zwei Stränge und der Neukonstruktion komplementärer Stränge, die durch die paarweise Bindung der Basen gegebenen sind. Auf eine genauere Darstellung des Reproduktionsverganges verzichten wir an dieser Stelle ebenfalls, fügen aber noch an, dass das Genom eines Organismus über mehrere DNS-Stränge verteilt sein kann; beim Menschen befindet sich beispielsweise ein Großteil des Erbgutes in den Zellkernen und ist dort innerhalb der Chromosomen unterteilt. Wir wollen verallgemeinert im Folgenden mit Chromosomen stets die verschiedenen DNS-Stränge eines Organismus bezeichnen.

Bevor wir nun eine einfache Variante des Wright-Fisher Modells betrachten, fassen wir zunächst noch einige Begriffe in der folgenden Definition zusammen.

2.1 Definition
Gegeben eine Population, bezeichnen wir diese als

(i) *haploid*, falls die Chromosomen eines Individuums stets nur in einfacher Form vorliegen,

(ii) *diploid*, falls die Chromosomen eines Individuums zweifach vorliegen.

Ferner bezeichnen wir eine bestimmte Ausprägung eines Gens an einem bestimmten Genort als *Allel*.

2.1. Das Wright-Fisher Modell

Als besonders einfache Variante des Wright-Fisher Modells wollen wir nun ein *Modell für zufällige Reproduktion ohne Mutationen und Selektion* hinsichtlich zweier Allele A und a einer haploiden Population konstanter Größe $2N$ mit nichtüberlappenden Generationen betrachten.

Dabei lässt sich der Evolutionsprozess wie folgt mit Hilfe eines Urnenmodelles beschreiben: Bezeichnen A_n beziehungsweise a_n die Anzahl der Individuen vom Typ A beziehungsweise a der n-ten Generation, so ergibt sich die $(n+1)$-ste Generation durch $2N$-faches Ziehen mit Zurücklegen aus einer Urne mit A_n Kugeln vom Typ A (und a_n Kugeln vom Typ a).

Damit erhalten wir für die Verteilung der Anzahl der Allele vom Typ A in der $(n+1)$-sten Generation eine Binomialverteilung mit Parametern $2N$ und $p_n^{(A)} \overset{\text{def}}{=} A_n/2N$ gemäß der Faltungseigenschaft von Bernoulli-Verteilungen mit Parameter $p_n^{(A)}$ (vgl. Satz 26.2 in [2]).

Dabei ist zu beachten, dass die Allel-Verteilung der $(n+1)$-sten Generation lediglich von der Verteilung der n-ten Generation abhängt: $(A_n)_{n\in\mathbb{N}}$ bildet folglich einen (diskrete) Markov-Kette.

Vermöge des Urnenmodells für den Generationswechsel, notieren wir ferner die offensichtliche Feststellung, dass 0 und $2N$ absorbierende Zustände für $(A_n)_{n\in\mathbb{N}}$ darstellen. Darüber hinaus stellen wir fest, dass einer dieser Zustände nach hinreichend langer Zeit angenommen wird.

In natürlicher Weise stellt sich damit die Frage, wie lange es dauert, bis ein Allel ausgestorben ist.

Wir folgen weiter [9] und bezeichnen mit

$$H_n^\circ = \frac{2A_n(2N - A_n)}{2N(2N-1)}$$

die Wahrscheinlichkeit, dass zwei zufällig gewählte Individuen (Ziehen ohne Zurücklegen) zum Zeitpunkt n von unterschiedlichem Typ sind. In Anlehnung an diploide Organismen, bei denen Heterozygotie (in Bezug auf ein bestimmtes Gen) bedeutet, dass im Chromosomensatz zwei verschieden Allele vorliegen, nennt man H_n° die *Heterozygotie* einer Population.

Wir notieren ohne Beweis:

2.2 Satz
In einem haploiden Modell für zufällige Reproduktion ohne Mutationen und Selektion hinsichtlich zweier Allele einer Population konstanter Größe $2N$ gilt für die Heterozygozität H_n° der Population zum Zeitpunkt n

$$\mathrm{E}H_n^\circ = \left(1 - \frac{1}{2N}\right)^n \mathrm{E}H_0^\circ.$$

Diese Beziehung gilt auch, wenn man in der Definition der Heterozygotie ohne Zurücklegen zieht. Das heißt, wenn man

$$H_n = \frac{2A_n(2N - A_n)}{(2N)^2} = \frac{2N-1}{2N} H_n^\circ$$

setzt, so gilt $\mathrm{E}H_n = (1 - 1/2N)^n \,\mathrm{E}H_0$.

Mit Verweis auf §8, Satz 2 in [11] erhalten wir aus der Reihendarstellung der Exponentialfunktion

$$e^{-x} = 1 - x + R(x), \text{ mit } |R(x)| \leq x^2,$$

für alle x mit $|x| \leq 3/2$.

Dies liefert für hinreichend großes $2N$ die asymptotische Beziehung

$$\mathrm{E}H_n \approx e^{-n}\,\mathrm{E}H_0$$

2.3 Bemerkung
Das beschriebene Urnenmodell lässt sich leicht um Abstammungsinformationen erweitern: Dazu macht man die Kugeln der Anfangsgeneration unterscheidbar, beispielsweise durch (zufällige) Durchnummerierung von 1 bis $2N$. Wir können dann ein Individuum der $(n+1)$-sten Generation als Kind eines Individuums der n-ten Generation identifizieren. Dementsprechend bezeichnen wir mit einer *Ahnenlinie* die aufsteigende Linie der Vorfahren eines Individuums der n-ten Generation.

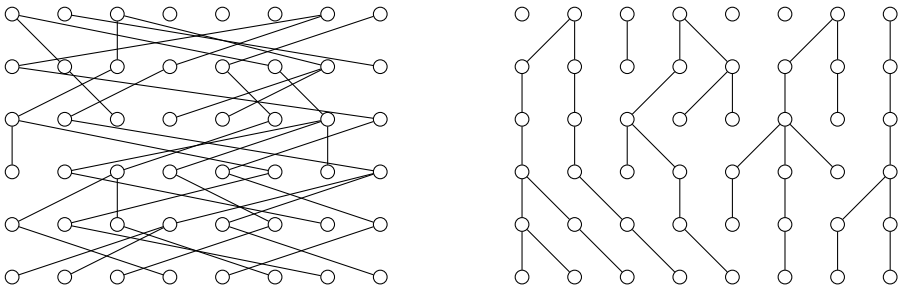

Abbildung 2.2.: Realisierung des Wright-Fisher Modells für eine Population der Größe 8 über 6 Generationen mit eingetragenen Ahnenlinien (links) und einer »entworrenen« Version (rechts).

2.4 Satz
Unter Reskalierung der Zeit bezüglich $2N$ Generationen genügt die Zeit τ_k, für die k verschiedene Ahnenlinien existieren, einer Exponentialverteilung mit Parameter $\binom{k}{2}$.

Beweis. Wir vermerken zunächst, dass bei der Betrachtung von Ahnenlinien, also der Betrachtung eines Wright-Fisher Modells in rückwärtiger Zeit, diese lediglich verschmelzen können. Dies ist der Fall, wenn wenigstens zwei Individuen einen direkten gemeinsamen Vorfahren haben.

Wir betrachten deshalb zunächst die Wahrscheinlichkeit, dass k Individuen einer Generation verschiedene Eltern haben. Dazu erinnern wir an den Generationswechsel und notieren, dass nach Ziehen von m verschiedenen Kugeln noch $2N-m$ davon verschiedene Kugeln existieren. Damit berechnet sich die gewünschte Wahrscheinlichkeit zu

$$\prod_{i=0}^{k-1} \frac{2N-i}{2N} = \prod_{i=1}^{k-1} \left(1 - \frac{i}{2N}\right) = 1 - \frac{\sum_{i=0}^{k-1} i}{2N} + \mathcal{O}\left(\frac{1}{N^2}\right),$$

wobei im letzten Schritt das Produkt ausmultipliziert und nach den Potenzen von $2N$ umsortiert wurde (vgl. auch [16]). Für die Wahrscheinlichkeit, dass k Ahnenlinien über die ersten n Generationen bestehen bleiben, erhalten wir daher

$$\left(1 - \frac{k(k-1)}{2} \cdot \frac{1}{2N} + \mathcal{O}\left(\frac{1}{N^2}\right)\right)^n$$

Reskalieren wir also die Zeit mit $2N$, setzen wir also $t \stackrel{\text{def}}{=} n/2N$, so erhalten wir

$$\lim_{N\to\infty} \left(1 - \frac{k(k-1)}{2} \cdot \frac{1}{2N} + \mathcal{O}\left(\frac{1}{N^2}\right)\right)^{2Nt} = \exp\left(-\frac{k(k-1)}{2}t\right)$$

und damit für die skalierte Zeit bis zum ersten Verschmelzungsereignis eine exponentialverteilte Zeit mit Parameter $\binom{k}{2}$.

Mit der Feststellung, dass die obige Argumentation für jeden Zeitpunkt, zu dem k Ahnenlinien existieren, greift, folgt damit die Behauptung. □

Für die Wahrscheinlichkeit, dass mehr als zwei Individuen dieselben Eltern haben, notieren wir, dass diese in $\mathcal{O}\left(\frac{1}{N^2}\right)$ liegt und mithin ebenfalls asymptotisch vernachlässigbar ist. Dazu überlegen wir uns, dass die Wahrscheinlichkeit, dass es ein Individuum der Vorgängergeneration mit mindestens 3 Kindern gibt, durch $2N\binom{k}{3}\left(\frac{1}{2N}\right)^3 \in \mathcal{O}\left(\frac{1}{N^2}\right)$ gegeben ist.

2.5 Bemerkung
Unter den Annahmen von 2.4 verschmelzen k verschiedene Ahnenlinien nach einer exponentialverteilten Zeit τ_k mit Parameter $\binom{k}{2}$ zu $(k-1)$ verschiedenen Ahnenlinien.

Vermöge der Feststellung, dass sich in einem Modell zufälliger Reproduktion ohne Mutation und Selektionsdruck hinsichtlich zweier Allele einer diploiden Population konstanter Größe N, jedes Gen zu jedem Zeitpunkt eindeutig zurückverfolgen lässt, kann man dieses auch als Modell für eine haploide Population konstanter Größe $2N$ betrachten.

Zu beachten ist allerdings, dass bei dieser Betrachtungsweise auch Selbstbefruchtung zu berücksichtigen ist. Mit Bezug auf (14) in [16] notieren wir allerdings, dass für einen geeignet gewählten Skalierungsparameter $2N_e = 2N/(1+F)$, wobei F abhängig von der Wahrscheinlichkeit für Selbstbefruchtung ist, die Ahnenlinien sich gemäß der obigen Bemerkung verhalten.

Ähnlich erhält man für eine zweigeschlechtliche, diploide Population der Größe N mit N_m männlichen und N_f weiblichen Individuen, also $N = N_m + N_f$, einen Skalierungsparameter $2N_e = 8N_mN_f/(N_m+N_f)$, bezüglich dem die Ahnenlinien sich asymptotisch erneut gemäß der obigen Bemerkung verhalten (vgl (16) in [16]).

3. Der Koaleszenzprozess nach Kingman

Wir haben gesehen, dass für eine haploide Population der Größe $2N$, beziehungsweise alternativ für eine diploide Population der Größe N mit zufälliger Rekombination, ohne Mutation und Selektionsdruck, die Zeit, bis k Ahnenlinien zu $k-1$ Ahnenlinien verschmelzen, asymptotisch gemäß einer Exponentialverteilung mit Parameter $\binom{k}{2}$ verhält. Kingman [14],[15] definierte 1982 den n-Koaleszenzprozess als zeitstetige Markov-Kette auf den Äquivalenzrelationen von $\{1, 2, \ldots, n\}$ als Approximation der Geneaolgie einer (haploiden) Population unter dem Wright-Fisher-Model.

Bemerkenswert ist dabei, dass die Betrachtung zeitlich rückwärts stattfindet und lediglich die Genealogie einer Population und nicht die expliziten genetischen Strukturen jeder Generation betrachtet. Insbesondere werden bereits ausgestorbene Erblinien nicht betrachtet, was im Hinblick auf Simulationsverfahren auch den Rechenaufwand reduziert.

3.1. Der n-Koaleszenzprozess

In diesem Abschnitt sei stets $n \in \mathbb{N}$ eine natürliche Zahl. Ferner sei, sofern nicht anders angegeben, stets $(\Omega, \mathfrak{A}, \mathrm{P})$ ein Wahrscheinlichkeitsraum.

3.1 Definition (i) $\mathscr{E}_n$ bezeichne die Menge der Äquivalenzrelationen auf $\{1, 2, \ldots, n\}$.

(ii) Die Äquivalenzrelationen $\Delta_n, \Theta_n \in \mathscr{E}_n$ seien durch

$$\Delta_n \stackrel{\text{def}}{=} \{(i,i) | i \in \{1, 2, \ldots, n\}\}, \tag{3.1}$$

$$\Theta_n \stackrel{\text{def}}{=} \{(i,j) | i, j \in \{1, 2, \ldots, n\}\} \tag{3.2}$$

gegeben. Dann ist Δ_n die Äquivalenzrelation, in der jedes Element aus $\{1, 2, \ldots, n\}$ nur mit sich selbst äquivalent ist und Θ_n die Äquivalenzrelation, in der alle Elemente aus $\{1, 2, \ldots, n\}$ miteinander äquivalent sind.

(iii) Für $\xi \in \mathscr{E}_n$ bezeichne $|\xi|$ die Anzahl der Äquivalenzklassen von ξ. Darüber hinaus bezeichne für $k < l \leq n$

$$k \sim_\xi l \stackrel{\text{def}}{\Leftrightarrow} (k,l) \in \xi$$

die Äquivalenz von k und l bezüglich ξ, sowie $[k]$ die Äquivalenzklasse von k.

(iv) Für $\xi, \eta \in \mathscr{E}_n$ bezeichne $\xi \prec \eta$, dass η aus ξ durch das Verschmelzen zweier Äquivalenzklassen hervorgeht, also dass

$$\xi \prec \eta \overset{\text{def}}{\Leftrightarrow} \xi \subset \eta, \quad |\xi| = |\eta| + 1 \tag{3.3}$$

gilt.

Wir folgen nun Kingman in der Defintion des n-Koaleszenzprozesses in [14].

3.2 Definition
Ein Markov-Sprungprozess $(R_t)_{t\geq 0}$ mit Zustandsraum $\mathscr{E}_n$ heißt *n-Koaleszenzprozess*, falls

$$R_0 = \Delta_n \tag{3.4}$$

gilt und die Übergangsraten

$$q_{\xi\eta} = \lim_{h\downarrow 0} h^{-1}\,\mathrm{P}(R_{t+h} = \eta | R_t = \xi)$$

mit $\xi, \eta \in \mathscr{E}_n, \xi \neq \eta$ durch

$$q_{\xi\eta} = \begin{cases} 1 & \text{falls } \xi \prec \eta, \\ 0 & \text{sonst} \end{cases} \tag{3.5}$$

gegeben sind.

Damit nun $\mathbf{Q} \overset{\text{def}}{=} (q_{\xi,\eta})_{\xi,\eta}$ konservativ wird, ergibt sich für die Diagonaleinträge unweigerlich:

3.3 Korollar
Es sei $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $\xi \in \mathscr{E}_n$. Für die Gesamtaustrittsrate

$$q_\xi = \lim_{h\downarrow 0} h^{-1} P(R_{t+h} \neq \xi | R_t = \xi) = \sum_{\eta \neq \xi} q_{\xi\eta}$$

aus ξ gilt

$$q_\xi = \frac{1}{2}|\xi|(|\xi| - 1). \tag{3.6}$$

Beweis. Die einzigen Zustände $\eta \neq \xi$, für die der Summand ungleich 0 ist, sind definitionsgemäß alle η mit $\xi \prec \eta$, also diejenigen η, welche aus ξ durch die Verschmelzung zweier

Äquivalenzklassen von ξ hervorgehen. Dies entspricht der Auswahl aller 2-elementigen Teilmengen aus einer $|\xi|$-elementigen Menge. Damit gilt

$$q_\xi = \sum_{\eta \neq \xi} q_{\xi\eta} = \sum_{\xi \prec \eta} q_{\xi\eta} = \sum_{\xi \prec \eta} 1 = \binom{|\xi|}{2} = \frac{1}{2}|\xi|(|\xi| - 1)$$

für alle $\xi \in \mathscr{E}_n$. □

3.4 Bemerkung
Solche Markov-Sprungprozesse existieren und besitzen alle dieselben endlich-dimensionalen Verteilungen. Darüber hinaus können diese so konstruiert werden, dass ihre Pfade rechtsseitig stetig und stückweise konstant sind.

Vermöge der Verteilungs-Eindeutigkeit, ist es mitunter auch üblich, von *dem* n-Koaleszenzprozess zu reden.

Beweis. Wir beginnen mit der Feststellung, dass $\mathscr{E}_n$ versehen mit der diskreten Topologie insbesondere polnisch ist. Ferner haben wir $\mathbf{Q}$ gerade so gewählt, dass $\mathbf{Q}$ konservativ ist. Mit Hinweis auf 1.27 folgt wegen der Endlichkeit des Zustandsraumes damit die Eindeutigkeit der minimalen Konstruktion. Speziell gilt für die Übergangsmatrixfunktion $(\mathbf{P}(t))_{t\geq 0}$

$$\mathbf{P}_t = e^{\mathbf{Q}t} \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{t^k}{k!}\mathbf{Q}^k,$$

für $t \geq 0$ (vgl. Korollar 3.5 in [3]). □

Insbesondere gilt vermöge 1.18, dass n-Koaleszenzprozesse (mit rechtsseitig stetigen und stückweise konstanten Pfaden) die starke Markov-Eigenschaft besitzen.

3.5 Bemerkung
Sei $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $\xi \in \mathscr{E}_n$ mit $|\xi| = k$ für ein $k \in \mathbb{N}$, dann ist die Verweildauer in ξ exponentialverteilt mit Parameter $d_k = \frac{1}{2}k(k-1)$ und hängt somit einzig von $|\xi|$ ab.

Beweis. Aufgrund der Markov-Eigenschaft ist klar, dass die Verweildauer eine gedächtnislose Verteilung haben muss und die einzige zeitstetige gedächtnislose Verteilung ist die Exponentialverteilung. Mit Gleichung (3.6) folgt schließlich die Behauptung. □

3.6 Definition und Satz
Gegeben einen n-Koaleszenzprozess $(R_t)_{t\geq 0}$, definieren wir

$$D_t \stackrel{\text{def}}{=} |R_t|. \tag{3.7}$$

Dann ist $(D_t)_{t\geq 0}$ eine zeitstetige Markov-Kette mit Zustandsraum $\{1, 2, \ldots, n\}$ und Übergangsraten

$$\lim_{h\downarrow 0} h^{-1} \operatorname{P}(D_{t+h} = l | D_t = k) = \begin{cases} d_k & \text{falls } l = k-1, \\ 0 & \text{falls } l \neq k, k-1, \end{cases} \tag{3.8}$$

mit $d_k = \frac{1}{2}k(k-1)$.

Beweis. Sei $\xi \in \mathscr{E}_n$. Für jeden Zustandswechsel von ξ in einen Zustand $\eta \in \mathscr{E}_n$ muss gemäß Definition $|\eta| = |\xi| - 1$ gelten. Jeder Zustandswechsel in $(R_t)_{t\geq 0}$ entspricht also genau einem Zustandswechsel in $(D_t)_{t\geq 0}$ und umgekehrt. Der Prozess $(D_t)_{t\geq 0}$ verhält sich also in seiner Sprungdynamik wie der Koaleszenzprozess und ist mithin selbst ein Markov-Sprungprozess. Die Übergangsrate von einem Zustand k in einen Zustand $k-1$ in $(D_t)_{t\geq 0}$ korrespondiert demnach mit den Austrittsraten q_ξ aus einem Zustand ξ mit $|\xi| = k$ im Koaleszenzprozess.
Sei $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $(D_t)_{t\geq 0}$ gemäß 3.6 definiert. Sei ferner $\omega \in \Omega$ und gelte $D_t(\omega) = k$. Dann gibt es ein $\xi(\omega) \in \mathscr{E}_n$ mit $|\xi(\omega)| = k$ und $R_t(\omega) = \xi(\omega)$. Seien ferner $k, l \in \{1, 2, \ldots, n\}$, dann gilt:

$$\begin{aligned} &\lim_{h\downarrow 0} h^{-1} \operatorname{P}(D_{t+h}(\omega) = l | D_t(\omega) = k) \\ =&\lim_{h\downarrow 0} h^{-1} \operatorname{P}(D_{t+h}(\omega) = l | R_t(\omega) = \xi(\omega)) \\ =&\lim_{h\downarrow 0} h^{-1} \sum_{\substack{\eta(\omega)\in\mathscr{E}_n \\ |\eta(\omega)|=l}} \operatorname{P}(R_{t+h}(\omega) = \eta(\omega)) | R_t(\omega) = \xi(\omega)) \end{aligned}$$

Für $l = k-1$ bedeutet dies gerade $q_\xi(\omega)$ und für $l \neq k, k-1$ sind die einzelnen Summanden $= 0$. □

In der üblichen Terminologie (vgl. 1.20) handelt es sich bei $(D_t)_{t\geq 0}$ um einen reinen Todesprozess mit Anfangszustand n und Sterberaten d_k.

Darüber hinaus ist die in 3.1 definierte Äquivalenzrelation Θ ein absorbierender Zustand für $(R_t)_{t\geq 0}$ und dazu korrespondierend der Zustand 1 für $(D_t)_{t\geq 0}$ absorbierend.

3.7 Proposition
Sei $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $(D_t)_{t\geq 0}$ der assoziierte Todesprozess gemäß 3.6. Es bezeichne τ_k die Verweildauer von $(D_t)_{t\geq 0}$ im Zustand k für alle $k \in \{2, \ldots, n\}$.

(i) Dann sind die τ_k unabhängig und jeweils gemäß einer Exponentialverteilung mit Parameter $d_k = \frac{1}{2}k(k-1)$ verteilt, das heißt $\tau_k \sim \operatorname{Exp}(d_k)$.

(ii) Die Durchgangszeit

$$T \stackrel{\text{def}}{=} \inf\{t \geq 0 | R_t = \Theta\} = \inf\{t \geq 0 | D_t = 1\}$$

kann mittels der τ_k als

$$T = \sum_{k=2}^{n} \tau_k$$

dargestellt werden.

Beweis. zu (i) Analog zu 3.5 liefert auch hier die Markov-Eigenschaft zusammen mit den Übergangsraten von $(D_t)_{t\geq 0}$, dass die τ_k jeweils exponentialverteilt mit Parameter d_k sind. Ebenfalls liefert die Markov-Eigenschaft die Unabhängigkeit vermöge der Feststellung, dass die Aufenthaltsdauer in einem Zustand unabhängig vom Eintrittszeitpunkt ist. □

Ein typischer Pfad von $(R_t)_{t\geq 0}$ ist nun eine Folge von Äquivalenzrelationen

$$\Delta = \mathscr{R}_n \prec \mathscr{R}_{n-1} \prec \mathscr{R}_{n-2} \prec \cdots \prec \mathscr{R}_2 \prec \mathscr{R}_1 = \Theta, \tag{3.9}$$

wobei der Prozess jeweils eine $\text{Exp}(d_k)$-verteilte Zeit τ_k in $\mathscr{R}_k$ verbringt. Offensichtlich gilt

$$|\mathscr{R}_k| = k. \tag{3.10}$$

Die Folge (3.9) entspricht einem typischen Pfad der eingebetteten diskreten Markovkette des n-Koaleszenzprozesses, auch Sprung-Kette genannt.

3.2. Die Sprung-Kette

3.8 Satz
Es sei $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $(D_t)_{t\geq 0}$ der assoziierte Todesprozess, sowie $(\mathscr{R}_k)_{k=n,n-1,\ldots,1}$ die eingebettete diskrete Markovkette. Dann sind $(\mathscr{R}_k)_{k=n,n-1\ldots,1}$ und $(D_t)_{t\geq 0}$ unabhängig und es gilt

$$R_t = \mathscr{R}_{D_t} \tag{3.11}$$

für alle $t \geq 0$. Die Übergangswahrscheinlichkeiten der Markov-Kette $(\mathscr{R}_k)$ sind für $\xi \in \mathscr{E}_n$, $|\xi| = k$ und $2 \leq k \leq n$ durch

$$\mathrm{P}(\mathscr{R}_{k-1} = \eta | \mathscr{R}_k = \xi) = \begin{cases} \frac{2}{k(k-1)} & \text{falls } \xi \prec \eta, \\ 0 & \text{sonst,} \end{cases} \tag{3.12}$$

gegeben. Die absoluten Wahrscheinlichkeiten sind durch

$$\mathrm{P}(\mathscr{R}_k = \xi) = \frac{(n-k)!k!(k-1)!}{n!(n-1)!}\lambda_1!\lambda_2!\cdots\lambda_k!, \tag{3.13}$$

gegeben, wobei $\lambda_1, \lambda_2, \ldots, \lambda_k$ die Größen der Äquivalenzklassen von ξ bezeichne.

Beweis. Gemäß der Theorie über Sprung-Ketten (vgl. 1.24), sind die Übergangswahrscheinlichkeiten von der Form

$$\frac{q_{\xi\eta}}{q_\xi} \qquad (\xi \neq \eta),$$

sofern $q_\xi > 0$ gilt. Zudem sind die Verweildauern bedingt unter der Sprung-Kette unabhängig und exponentialverteilt mit Parameter q_ξ.

Es gilt also für $\xi, \eta \in \mathscr{E}_n$, $|\xi| = k$ und $\xi \prec \eta$

$$P(\mathscr{R}_{k-1} = \eta | \mathscr{R}_k = \xi) = \frac{q_{\xi\eta}}{q_\xi} = q_\xi^{-1} = \frac{2}{|\xi|(|\xi|-1)}$$

und damit (3.12) unter Berücksichtigung von (3.10).

Für den Beweis von (3.13) führen wir eine Rückwärtsinduktion über k durch. Führe dafür zunächst die abkürzende Schreibweise

$$p_k(\xi) \overset{\text{def}}{=} \mathrm{P}(\mathscr{R}_k = \xi), \quad \xi \in \mathscr{E}_n, |\xi| = k$$

ein. Für den Fall $k = n$ ist die Aussage offensichtlich richtig, denn in diesem Fall folgt schon $\xi = \Delta_n$ und damit

$$1 = p_n(\Delta_n) = \frac{(n-n)!n!(n-1)!}{n!(n-1)!}\underbrace{1!\cdots 1!}_{=\lambda_1!\cdots\lambda_n!} = 1$$

Gemäß (3.12) gilt ferner für $\eta \in \mathscr{E}_n$

$$p_{k-1}(\eta) = \sum_{\xi \prec \eta} \frac{2}{k(k-1)} p_k(\xi),$$

denn:

$$\begin{aligned}\mathrm{P}(\mathscr{R}_{k-1} = \eta) &= \sum_{\xi \prec \eta} \mathrm{P}(\mathscr{R}_k = \xi)\,\mathrm{P}(\mathscr{R}_{k-1} = \eta | \mathscr{R}_k = \xi) \\ &= \sum_{\xi \prec \eta} \frac{2}{|\xi|(|\xi|-1)}\,\mathrm{P}(\mathscr{R}_k = \xi)\end{aligned}$$

Wenn nun $\lambda_1, \lambda_2, \dots \lambda_{k-1}$ die Größen der Äquivalenzklassen von η bezeichnen, sind die von ξ gerade $\lambda_1, \lambda_2, \dots, \lambda_{l-1}, \nu, \lambda_l - \nu, \lambda_{l+1}, \dots, \lambda_{k-1}$ für ein $l, 1 \leq l \leq k-1$ und ein $\nu, 1 \leq \nu \leq \lambda_l - 1$. Nehmen wir nun also an, die Aussage sei bereits für k bewiesen, so müssen wir, um über alle möglichen ξ aus unserer Rekursionsgleichung aufzusummieren, über alle möglichen Indizes l und alle Möglichkeiten, die l-te Äquivalenzklasse von η auf zwei verschiedene Klassen aufzuteilen, aufsummieren. Für jedes $1 \leq \nu \leq \lambda_l - 1$ gibt es dann $\binom{\lambda_l}{\nu} = \frac{\lambda_l!}{\nu!(\lambda_l-\nu)!}$ Möglichkeiten, ν Elemente aus der l-ten Äquivalenzklasse von η für eine neue Klasse zu entnehmen. Nun vertauschen aber die Rollen der ν-ten und der l-ten Klasse miteinander, weshalb dies mit einem Faktor von $\frac{1}{2}$ korrigiert werden muss. In Formeln geschrieben gilt dann

$$
\begin{aligned}
&p_{k-1}(\eta) \\
=&\sum_{l=1}^{k-1}\sum_{\nu=1}^{\lambda_l-1} \frac{2}{k(k-1)} \frac{(n-k)!k!(k-1)!}{n!(n-1)!} \lambda_1! \cdots \lambda_{l-1}! \nu! (\lambda_l - \nu)! \lambda_{l+1} \cdots \lambda_{k-1}! \frac{1}{2} \binom{\lambda_l}{\nu} \\
=&\frac{(n-k)!(k-1)!(k-2)!}{n!(n-1)!} \sum_{l=1}^{k-1} \lambda_1! \cdots \lambda_{l-1}! \lambda_{l+1}! \cdots \lambda_{k-1}! \sum_{\nu=1}^{\lambda_l-1} \binom{\lambda_l}{\nu} (\lambda_l - \nu)! \nu! \\
=&\frac{(n-k)!(k-1)!(k-2)!}{n!(n-1)!} \sum_{l=1}^{k-1} \lambda_1! \cdots \lambda_{l-1}! \lambda_{l+1}! \cdots \lambda_{k-1}! \sum_{\nu=1}^{\lambda_l-1} \frac{\lambda_l! (\lambda_l - \nu)! \nu!}{(\lambda_l - \nu)! \nu!} \\
=&\frac{(n-k)!(k-1)!(k-2)!}{n!(n-1)!} \sum_{l=1}^{k-1} \lambda_1! \cdots \lambda_{l-1}! \lambda_l! \lambda_{l+1}! \cdots \lambda_{k-1}! \sum_{\nu=1}^{\lambda_l-1} 1 \\
=&\frac{(n-k)!(k-1)!(k-2)!}{n!(n-1)!} \lambda_1! \lambda_2! \cdots \lambda_{k-1}! \sum_{l=1}^{k-1}\sum_{\nu=1}^{\lambda_l-1} 1. \qquad (3.14)
\end{aligned}
$$

Die Doppelsumme in (3.14) berechnet sich ferner zu

$$
\sum_{l=1}^{k-1}\sum_{\nu=1}^{\lambda_l-1} 1 = \sum_{l=1}^{k-1} (\lambda_l - 1) = \sum_{l=1}^{k-1} \lambda_l - \sum_{l=1}^{k-1} 1 = n - (k-1),
$$

womit der Induktionsschritt bewiesen ist. □

3.9 Satz
Gegeben ein n-Koaleszenzprozess $(R_t)_{t \geq 0}$ mit assoziierter Sprungkette $(\mathscr{R}_k)_k$, sowie $l < k, \xi, \eta \in \mathscr{E}_n, |\xi| = k, |\eta| = l, \xi \subset \eta$, gilt für die bedingten Verteilungen von $\mathscr{R}_k$

$$
\mathrm{P}(\mathscr{R}_l = \eta | \mathscr{R}_k = \xi) = \frac{(k-l)!l!(l-1)!}{k!(k-1)!} \lambda_1! \lambda_2! \dots \lambda_l!, \qquad (3.15)
$$

wobei $\lambda_1, \lambda_2, \ldots, \lambda_l$ wie folgt definiert sind: Seien η_i, $1 \leq i \leq l$ die Äquivalenzklassen von η und ξ_j, $1 \leq j \leq k$ die Äquivalenzklassen von ξ, so ist λ_i die Anzahl der Äquivalenzklassen von ξ, welche zu η_i verschmelzen, das heißt

$$\lambda_i = |\{\xi_j | 1 \leq j \leq k, \xi_k \subset \eta_i\}|.$$

Beweis. Wir verzichten an dieser Stelle auf einen Beweis, merken aber an, dass man einen Beweis führen kann, wenn man die Äquivalenzklassen von ξ als Individuen eines nunmehr $|\xi|$-Koaleszenzprozesses auffasst. Dieser Idee werden wir im folgenden Abschnitt weiter nachgehen und verweisen deshalb an dieser Stelle auf das noch ausstehende Ergebnis 3.11. □

3.3. Der Koaleszenzprozess

Nachdem wir nun den n-Koaleszenzprozess für beliebige natürliche Zahlen n kennengelernt haben, stellt sich hinsichtlich tieferer Analysen, wie zum Beispiel Grenzwertverhalten von n-Koaleszenzprozessen, die Frage, inwiefern Koaleszenzprozesse für verschiedene n in Verbindung stehen. In den Kapiteln 6 und 7 in [15], stellt Kingman zwei verschiedene Konzepte der Einbettung eines m-Koaleszenzprozesses in einen n-Koaleszenzprozess vor, *Temporal Coupling* beziehungsweise *Natural Coupling*. Als Konsequenz dieser Konsistenzeigenschaft ergibt sich darüber hinaus sogar die Existenz eines Prozesses $(R_t)_{t\geq 0}$ auf der Menge $\mathscr{E}$ der Äquivalenzklassen auf den natürlichen Zahlen, so dass dieser n-Koaleszenzprozesse für alle $n \in \mathbb{N}$ enthält (vgl. Kapitel 8 in [15], oder [14]).

3.10 Bemerkung
Es sei $\xi \in \mathscr{E}_n$ eine Äquivalenzrelation mit Äquivalenzklassen $\xi_1, \xi_2, \ldots, \xi_k$. Wir vermerken zunächst, dass wir die Äquivalenzklassen nach der natürlichen Ordnung ihrer kleinsten Elemente sortieren können, das heißt, wir schreiben

$$\xi_i \leq \xi_j \overset{\text{def}}{\Leftrightarrow} \min\{k : k \in \xi_i\} \leq \min\{k : k \in \xi_j\},$$

wobei wir $\min \emptyset \overset{\text{def}}{=} \inf$ setzen. Dann können wir jede Äquivalenzrelation in $\mathscr{E}_n$ als n-Tupel $(\nu_1, \nu_2 \ldots, \nu_n)$ darstellen, mit $\nu_i < \nu_j$ für alle $0 < i < j \leq |\xi|$ und $\nu_k = \emptyset$ für $k > |\xi|$.

3.11 Proposition (Temporal Coupling)
Es sei $(R_t^{(n)})_{t\geq 0}$ ein n-Koaleszenzprozess und $(D_t)_{t\geq 0}$ der assoziierte Todesprozess. Es bezeichne τ_k die Verweildauer von $(D_t)_{t\geq 0}$ im Zustand k. Setze $T_k \overset{\text{def}}{=} \sum_{i=0}^{k-1} \tau_{n-i}$. Definiere

ferner eine Abbildung $\kappa_{m,n} : \mathscr{E}_n \to \mathscr{E}_m$ gemäß

$$\kappa_{m,n}(\xi_1, \xi_2, \ldots, \xi_n) = ([1], [2], \ldots, [|\xi|], \emptyset, \ldots, \emptyset),$$

wobei $(\xi_1, \xi_2, \ldots, \xi_n)$ die in 3.10 vorgestellte Darstellung bezeichnet. Dann gilt

$$(\kappa_{m,n} R^{(n)}_{T_m+t})_{t\geq 0} \stackrel{\mathrm{d}}{=} (R^{(m)}_t)_{t\geq 0}$$

für einen m-Koaleszenzprozess $(R^{(m)}_t)_{t\geq 0}$, das heißt $(\kappa_{m,n} R_{T_m+t})_{t\geq 0}$ ist ein m-Koaleszenzprozess.

Beweis. Wir beginnen mit der Feststellung, dass es ausreichend ist, die Behauptung für den Übergang $n \to (n-1)$ zu zeigen. Bezeichne also $\tau_n = \inf\{t > 0 : R_t \neq R_0\}$ die erste Sprungzeit und $\kappa \stackrel{\mathrm{def}}{=} \kappa_{n-1,n}$. Vermöge der starken Markov-Eigenschaft ist $(R'_t)_{t\geq 0} \stackrel{\mathrm{def}}{=} (R_{\tau_n+t})_{t\geq 0}$ ein Markov-Sprungprozess mit derselben Sprungdynamik wie $(R_t)_{t\geq 0}$, gestartet in R_{τ_n}. Offensichtlich gilt $R'_0 = \Delta_m$ und jeder Sprung von $(R'_t)_{t\geq 0}$ induziert einen Sprung von $(\kappa R'_t)_{t\geq 0} = (\kappa R_{\tau_n+t})_{t\geq 0}$. □

Rein anschaulich kann man also für $m < n$ einen m-Koaleszenzprozess innerhalb eines n-Koaleszenzprozesses $(R_t)_{t\geq 0}$ finden, indem man bis zu dem Zeitpunkt wartet, an dem $(R_t)_{t\geq 0}$ aus m Äquivalenzklassen besteht, und dann diese als Individuen auffasst.

Im Hinblick auf den n-Koaleszenzprozess als Modell für die Genealogie einer Auswahl von n Individuen einer Population der Größe $2N$, ist es allerdings natürlicher, statt n Individuen eine Subpopulation von $m < n$ Individuen zu betrachten. In Bezug zu einem n-Koaleszenzprozess bedeutet dies die Einschränkung auf die Äquivalenzrelationen auf m Elementen. Die positive Antwort, dass auch auf diese Art ein m-Koaleszenzprozess innerhalb eins n-Koaleszenzprozesses gefunden werden kann, liefert nunmehr

3.12 Proposition (Natural Coupling)
Sei $(R^{(n)}_t)_{t\geq 0}$ ein n-Koaleszenzprozess, sowie $m < n$, $m, n \in \mathbb{N}$. Sei ferner $\rho_{m,n} : \mathscr{E}_n \to \mathscr{E}_m$ definiert durch

$$\rho_{m,n}(\xi) \stackrel{\mathrm{def}}{=} \{(i,j) | 1 \leq i, j \leq m, (i,j) \in \xi\}$$

für alle $\xi \in \mathscr{E}_n$. Dann ist $(\rho_{m,n} R^{(n)}_t)_{t\geq 0}$ ein m-Koaleszenzprozess, das heißt

$$(\rho_{m,n} R^{(n)}_t)_{t\geq 0} \stackrel{\mathrm{d}}{=} (R^{(m)}_t)_{t\geq 0} \tag{3.16}$$

für einen m-Koaleszenzprozess $(R^{(m)}_t)_{t\geq 0}$.

Beweis. Wie im Beweis zu 3.11 vermerken wir auch hier, dass es ausreicht, den Übergang $n \to (n-1)$ zu betrachten. Dies ergibt sich leicht aus der Feststellung

$$\rho_{m,n} = \rho_{m,n-1} \circ \rho_{n-1,n}$$

für $m < n-1$.

Sei also $(R_t)_{t\geq 0}$ ein n-Koaleszenzprozess und $\rho \overset{\text{def}}{=} \rho_{n-1,n} : \mathscr{E}_n \to \mathscr{E}_{n-1}$ die oben definierte Einschränkungsabbildung. Seien ferner $\xi \in \mathscr{E}_{n-1}$ und wähle $\nu \in \mathscr{E}_n$ derart, dass $\rho\nu = \xi$ gilt.

Darüber hinaus sei an dieser Stelle davon ausgegangen, dass $(R_t)_{t\geq 0}$ in ν startet (vgl. [5]): Bezeichne $\sigma_\nu \overset{\text{def}}{=} \inf\{t \geq 0 : R_t = \nu\}$ die (Erst-)Eintrittszeit in den Zustand ν. Dann ist ν eine Stoppzeit (vgl. (1.22); da der Wert 0 nur angenommen wird, wenn $\nu = \Delta_n$ gilt) und vermöge der starken Markov-Eigenschaft, sowie der zeitlichen Homogenität ist $(R_{\sigma_\nu+t})_{t\geq 0}$ ein Markov-Prozess mit denselben Übergangsraten wie ein n-Koaleszenzprozess, der jedoch in ν startet.

Falls nun ein $m \in \{1, 2, \ldots, n-1\}$ mit $m \sim_\nu n$ existiert, das heißt, falls die Äquivalenzklasse von n in ν nicht einelementig ist, so gilt bereits $|\nu| = |\xi|$ und für jedes $\mu \in \mathscr{E}_n$ mit $\nu \prec \mu$ gilt $\rho\nu \prec \rho\mu$. Darüber hinaus gilt natürlich auch $\rho\mu \neq \rho\mu'$ für alle $\mu, \mu' \in \mathscr{E}_n$ mit $\nu \prec \mu$, $\nu \prec \mu'$, $\mu \neq \mu'$. Damit stimmt das Übergangsverhalten von $(R_t)_{t\geq 0}$ und $(\rho R_t)_{t\geq 0}$ überein.

Sei also nun die Äquivalenzklasse von n in ν einelementig und setze $k \overset{\text{def}}{=} |\nu|$. Bezeichne σ die Zeit bis zum ersten Sprung aus ν heraus. Dann genügt σ einer Exponentialverteilung mit Parameter $\binom{k}{2}$. Bezeichne mit $\mu = R_\sigma$ den Zustand von R_t nach dem ersten Sprung. Dann entspricht die Verteilung von μ der Gleichverteilung auf der Menge der Äquivalenzrelationen, welche durch Verschmelzung zweier Äquivalenzklassen von ν entstehen. Davon sind $k-1$ dieser Relationen durch Verschmelzungen der Äquivalenzklassen von n mit einer der übrigen $k-1$ Äquivalenzklassen entstanden. Bezeichne also mit N das Ereignis, dass die Äquivalenzklasse von n in μ weiterhin einelementig ist. Offensichtlich gilt dann $\mathrm{P}(N) = \frac{k-2}{k}$ und, bedingt unter N, entspricht die Verteilung von $\rho\mu$ der Gleichverteilung auf den $\binom{k-1}{2}$ möglichen Verschmelzungen zweier Äquivalenzrelationen von $\rho\nu$.

Falls also N nicht eintritt, so bezeichne mit σ' die Zeit bis zum zweiten Sprung aus ν heraus. Dies entspricht der Zeit bis zum ersten Sprung aus μ heraus und genügt damit einer Exponentialverteilung mit Parameter $\binom{k-1}{2}$. Bezeichne mit $\mu' \overset{\text{def}}{=} R_{\sigma+\sigma'}$ den Zustand von $(R_t)_{t\geq 0}$ nach dem zweiten Sprung. Dann entspricht die Verteilung von $\rho\mu'$ der Gleichverteilung auf den $\binom{k-1}{2}$ möglichen Äquivalenzrelationen, die durch Verschmelzung zweier Äquivalenzklassen von $\rho\nu$ hervorgehen.

ρR_t verweilt nun also bis zum Zeitpunkt $\sigma + \mathbb{1}_{\{N^c\}}\sigma'$ im Zustand ξ und springt dann beim Eintreten von N in den Zustand μ und beim Nichteintreten von N in den Zustand μ'. Darüber hinaus hängt das Sprungziel nicht von den Zeitpunkten der Sprünge ab.

Es ist leicht einzusehen, dass die Voraussetzungen von A.6 gegeben sind. Damit genügt also $\sigma + \mathbb{1}_{\{N^c\}}\sigma'$ einer Exponentialverteilung mit Parameter $\binom{k-1}{2}$.

Insgesamt verweilt also $(\rho R_t)_{t\geq 0}$ eine $\binom{|\xi|}{2}$-verteilte Zeit im Zustand ξ und springt dann mit gleicher Wahrscheinlichkeit in eine der $\binom{|\xi|-1}{2}$ Äquivalenzrelationen, welche durch Verschmelzung zweier Äquivalenzklassen von ξ hervorgehen. □

Im Kontext des n-Koaleszenzprozesses als asymptotisches Modell für die Genealogie einer Population, erinnern wir daran, dass die Äquivalenzklassen mit Individuen assoziiert sind. Die spezielle Wahl der Einschränkungsabbildung $\rho_{m,n}$ bevorzugt bei der Betrachtung von Subpopulationen der Größe k einer Population der Größe n gerade die ersten k Individuen.

Man überlegt sich allerdings, dass die Assoziation von Individuen und Äquivalenzklassen auch in anderer Reihenfolge getroffen werden kann, und dass dies keine Auswirkung auf die Gestalt der Genealogie (unter neutraler Evolution) haben sollte.

Formalisiert wird dies durch den Begriff der Austauschbarkeit (vgl. unter anderem [14] oder [15]), auf den im Folgenden eingegangen wird.

3.13 Definition
Für $n \in \mathbb{N}$ bezeichne

(i)
$$S_n \overset{\text{def}}{=} \{\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\} | \pi \text{ bijektiv}\},$$
die *symmetrische Gruppe* von $\{1, 2, \ldots, n\}$,

(ii)
$$S_{\mathbb{N}} \overset{\text{def}}{=} \{\pi : \mathbb{N} \to \mathbb{N} | \pi \text{ bijektiv}\},$$
die *(unendliche) symmetrische Gruppe* von $\mathbb{N}$ und

(iii)
$$S_{\mathbb{N}}^{\text{fin}} \overset{\text{def}}{=} \{\pi \in S_{\mathbb{N}} | |\text{supp}(\pi)| < \infty\},$$
die *(unendliche) symmetrische Gruppe* von $\mathbb{N}$ *mit endlichem Träger*, wobei wir mit $\text{supp}(\pi) \overset{\text{def}}{=} \{n \in \mathbb{N} | \pi(n) \neq n\}$ den Träger von π bezeichnen.

3.14 Definition
Eine (zufällige) Äquivalenzrelation $\xi \in \mathscr{E}$ heißt *austauschbar*, falls ihre Verteilung invariant unter der Wirkung jeder Permutation $\pi \in S_{\mathbb{N}}^{\text{fin}}$ mit endlichem Träger ist, wobei die Wirkung von π auf ξ gemäß

$$\pi\xi = \xi_\pi \stackrel{\text{def}}{=} \{(i,j) : i,j \in \mathbb{N}, (\pi(i),\pi(j)) \in \xi\} \tag{3.17}$$

definiert ist.

Für (zufällige) Äquivalenzrelationen $\xi \in \mathscr{E}_n$ reduziert sich 3.14 offensichtlich auf die Invarianz der Verteilung von ξ unter der Wirkung der Permutationen $\pi \in S_n$.

3.15 Bemerkung
Man überlegt sich leicht, dass die in (3.17) definierte Abbildung $S_{\mathbb{N}}^{\text{fin}} \times \mathscr{E} \to \mathscr{E}$, $(\pi, \xi) \to \xi_\pi$ tatsächlich eine Wirkung im Sinne der Gruppentheorie ist.

3.16 Proposition
Der n-Koaleszenzprozess nach Kingman ist austauschbar. Dabei bezeichnen wir einen stochastischen Prozess $(R_t)_{t\geq 0}$ mit Werten in $\mathscr{E}_n$ als austauschbar, wenn

$$(R_t)_{t\geq 0} \stackrel{\text{d}}{=} (\pi R_t)_{t\geq 0} \tag{3.18}$$

für alle $\pi \in S_n$ gilt.

Beweis. Der Beweis folgt unmittelbar aus der Tatsache, dass das Verhalten eines n-Koaleszenzprozesses lediglich von der Anzahl der Äquivalenzklassen abhängt. Diese ist invariant unter $\pi \in S_n$. Darüber hinaus ist π nicht nur mit der Anzahl der Äquivalenzklassen verträglich, sondern auch mit der Größe der Äquivalenzklassen. □

Bevor wir uns nun mit der Frage nach der Existenz eines stochastischen Prozesses befassen, der n-Koaleszenzprozesse für alle $n \in \mathbb{N}$ enthält, wollen wir zunächst festhalten, dass der Raum $\mathscr{E}$ der Äquivalenzrelationen auf $\mathbb{N}$, versehen mit einer geeigneten Topologie, ein polnischer Raum ist.

Wir folgen dazu Lemma 2.6 in [5], definieren aber zunächst noch Einschränkungabbildungen $\rho_n : \mathscr{E} \to \mathscr{E}_n$, $n \in \mathbb{N}$, analog zu 3.12, gemäß

$$\rho_n(\xi) \stackrel{\text{def}}{=} \{(i,j) | 1 \leq i,j \leq n : (i,j) \in \xi\} \tag{3.19}$$

für alle $\xi \in \mathscr{E}$.

3.17 Lemma
Die Abbildung $\mathrm{d} : \mathscr{E} \times \mathscr{E} \to \mathbb{R}$, definiert durch

$$\mathrm{d}(\xi, \xi') \stackrel{\text{def}}{=} \frac{1}{\sup\{n \in \mathbb{N} : \rho_n \xi = \rho_n \xi'\}}, \tag{3.20}$$

für $\xi, \xi' \in \mathscr{E}$, mit der Konvention $1/\sup \mathbb{N} = 0$, ist eine Ultrametrik.

Beweis. Wir beginnen mit der Feststellung, dass lediglich der Nachweis der starken Dreiecksungleichung nicht offensichtlich ist. Seien also $\xi, \eta, \nu \in \mathscr{E}$ verschiedene Äquivalenzrelationen. Sei ferner $\mathrm{d}(\xi, \nu) = 1/k$ und $\mathrm{d}(\xi, \eta) = 1/n$. Für $n \leq k$ folgt $\mathrm{d}(\eta, \nu) = 1/n \geq 1/k$. Im Fall $n > k$ gilt aber bereits $\mathrm{d}(\eta, \nu) = 1/k$. Denn angenommen η und ν würden auch unter ρ_{k+1} übereinstimmen, so wäre dies auch für ξ und ν der Fall, was einen Widerspruch zur Annahme liefern würde. Insgesamt ergibt sich nun

$$\mathrm{d}(\xi, \nu) \leq \max\{\mathrm{d}(\xi, \eta), \mathrm{d}(\eta, \nu)\}, \tag{3.21}$$

d ist also eine Ultrametrik. □

3.18 Proposition
Der Raum $(\mathscr{E}, \mathrm{d})$ der Äquivalenzrelationen auf $\mathbb{N}$, versehen mit der in (3.20) definierten Metrik d, ist kompakt.

Beweis. Sei $(\xi_n)_{n\in\mathbb{N}} \subset \mathscr{E}$ eine Folge von Äquivalenzrelationen auf den natürlichen Zahlen.

Wir stellen zunächst fest, dass $(\rho_1 \xi_n)_n$ eine identische Folge ist. Mithin finden wir eine Teilfolge $(\xi_{1,n})_n$ derart, dass $\mathrm{d}(\xi_{1,n}, \xi_{1,n'}) \leq 1$ für alle n, n' gilt. Ferner überlegen wir uns, dass für eine Folge $(\nu_{k,n})_n \subset \mathscr{E}$ mit $\mathrm{d}(\nu_{k,n}, \nu_{k,n'}) \leq 1/k$ für alle n, n' die Menge $\{\rho_{k+1}\nu_{k,n} : n \in \mathbb{N}\}$ höchstens $(k+1)$-elementig ist. Dies liefert die Existenz einer Teilfolge $(\nu_{k+1,n})_n$ derart, dass $\mathrm{d}(\nu_{k+1,n}, \nu_{k+1,n'}) \leq 1/(k+1)$ für alle n, n'. Zusammen ergeben sich so Teilfolgen $\xi^{(k)} \stackrel{\text{def}}{=} (\xi_{k,n})_n$ derart, dass $\xi^{(k+1)}$ eine Teilfolge von $\xi^{(k)}$ ist, und $\mathrm{d}(\xi_{k,n}, \xi_{k,n'}) \leq 1/k$ für alle n, n'.

Für die diagonale Folge $(\nu_n)_n \stackrel{\text{def}}{=} (\xi_{n,n})_n$ ergibt sich damit für alle $k \in \mathbb{N}$

$$\rho_k \nu_n = \rho_k \nu_{n'} \tag{3.22}$$

für alle $n, n' \geq k$, und damit $\mathrm{d}(\nu_n, \nu_{n'}) \leq 1/k$ für alle $n, n' \geq k$.

Insbesondere können wir $(\nu_n)_n$ als Äquivalenzrelation auf den natürlichen Zahlen auffassen, indem wir eine Äquivalenzrelation $\xi \in \mathscr{E}$ als Element $(\rho_n \xi)_n \in \bigtimes_{i=1}^{\infty} \mathscr{E}_n$ auffassen.

Das Bild von $\mathscr{E}$ unter der so beschriebenen Einbettung sind dann gerade die Elemente von $\bigtimes_{i=1}^{\infty} \mathscr{E}_n$, die *regulär* im Sinne von (3.22) sind.

Damit besitzt jede Folge in $(\mathscr{E}, \mathrm{d})$ eine konvergente Teilfolge und $(\mathscr{E}, \mathrm{d})$ ist demnach ein kompakter metrischer Raum. □

3.19 Bemerkung
Gemäß 3.3 in [10] erzeugt die Metrik $\mathrm{d}'(\xi, \eta) \stackrel{\text{def}}{=} \sup_{n\in\mathbb{N}} 2^{-n} \mathbb{1}_{\{\rho_n\xi \neq \rho_n\eta\}}$ die von den Einschränkungsabbildungen ρ_n erzeugte Topologie, das heißt die schwächste Topologie, bezüglich der die Einschränkungsabbildungen ρ_n stetig sind. Ferner ist gemäß Lemma 9 in [10] der Raum $(\mathscr{E}, \mathrm{d}')$ kompakt und total unzusammenhängend und insbesondere polnisch (vgl. auch [4]).

Man überlegt sich leicht, dass d und d′ vermöge der Beziehung

$$\mathrm{d}(\xi, \xi') = 1/k \Leftrightarrow \mathrm{d}'(\xi, \xi') = 1/2^k$$

äquivalente Metriken bilden.

3.20 Satz
Es existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ und ein eindeutig verteilter stochastischer Prozess $(R_t)_{t\geq 0}$ mit Werten in $\mathscr{E}$, so dass für die Einschränkungsabbildungen $\rho_n : \mathscr{E} \to \mathscr{E}_n$, $n \in \mathbb{N}$, gemäß (3.19), die eingeschränkten Prozesse $(\rho_n R_t)_{t\geq 0}$ jeweils n-Koaleszenzprozesse sind. $(R_t)_{t\geq 0}$ heißt dann *Kingmans Koaleszenzprozess* oder *Koaleszenzprozess nach Kingman*.

Beweis. Wir folgen [4] in der Feststellung, dass eine Äquivalenzrelation ξ auf $\mathbb{N}$ als eine Funktion $\delta : \mathbb{N} \to \mathbb{N}$ verstanden werden kann, indem wir jede natürliche Zahl auf die kleinste zu ihr bezüglich ξ äquivalente natürliche Zahl abbilden, das heißt

$$\delta(n) \stackrel{\text{def}}{=} \min\{k \in \mathbb{N} : k \sim_\xi n\}.$$

Man überlegt sich leicht, dass für jede Äquivalenzrelation genau eine solche Funktion existiert und auch jede solche Funktion eine eindeutige Äquivalenzrelation beschreibt.

Damit lässt sich ein Prozess $(R_t)_{t\geq 0}$ auf $\mathscr{E}$ formal als Prozess mit Indexmenge $\mathbb{N}$ und Zustandsraum $\mathbb{N}^{[0,\infty)}$ auffassen.

Im Hinblick auf die endlich-dimensionalen Verteilungen dieses Prozesses notieren wir, dass für jede endliche Teilmenge $I \subset \mathbb{N}$ offensichtlich ein $k \in \mathbb{N}$ existiert, derart dass $I \subset \{1, 2, \ldots, k\}$. Sei ferner $J \subset I$. Vermöge der Austauschbarkeit von n-Koaleszenzprozesen dürfen wir ferner annehmen, dass I, J von der Gestalt $\{1, 2, \ldots, l\}$ mit $l \in \{|I|, |J|\}$ sind.

Mit dem Satz von Daniell-Kolmogorov (vgl. Satz 54.7 in [2]) folgt damit die Existenz eines eindeutigen Maßes Q auf den Äquivalenzrelationen von $\mathbb{N}$ mit den vorgegebenen Randverteilungen und mithin die Existenz eines eindeutig verteilten stochastischen Prozesses $(R_t)_{t\geq 0}$ auf $\mathscr{E}$, derart dass $(\rho_n R_t)_{t\geq 0}$ der Verteilung eines n-Koaleszenzprozesses genügt. □

Für konstruktive Beweise sei an dieser Stelle unter anderem auf Theorem 3 in [14] oder Proposition 2.1 in [4] verwiesen.

3.21 Bemerkung
Wie schon im Fall der n-Koaleszenzprozesse führt auch hier die Eindeutigkeit in Verteilung dazu, dass auch von *dem* Koaleszenzprozess gesprochen wird.

3.22 Proposition
Der Koaleszenzprozess nach Kingman ist austauschbar.

Beweis. Es sei $(R_t)_{t\geq 0}$ ein Koaleszenzprozess und $\pi \in S_{\mathbb{N}}^{\text{fin}}$ eine Permutation der natürlichen Zahlen mit endlichem Träger, dann gibt es ein $n \in \mathbb{N}$ derart, dass $\text{supp}(\pi) \subset \{1, 2, \ldots, n\}$ gilt. Wir wir in 3.16 bereits gesehen haben, ist $(\rho_k R_t)_{t\geq 0}$ austauschbar für alle $k \geq n$. Damit ist aber auch $(R_t)_{t\geq 0}$ bereits austauschbar. □

Der Startzustand des Koaleszenzprozesses nach Kingman ist, wie im endlichen Fall, erneut die Äquivalenzrelation, in der jede Äquivalenzklasse einelementig ist. Damit gilt insbesondere $|R_0| = \infty$. Man kann jedoch zeigen, dass der Koaleszenzprozess nach positiver Zeit fast sicher nur endlich viele Äquivalenzklassen hat. Wir notieren (vgl. Theorem 2.1 in [4]):

3.23 Proposition
Für einen Koaleszenzprozess $(R_t)_{t\geq 0}$, gilt

$$\mathrm{P}(|R_t| < \infty, t > 0) = 1,$$

das heißt, ein Koaleszenzprozess besitzt zu jedem Zeitpunkt $t > 0$ fast sicher nur endlich viele Äquivalenzklassen.

Beweis. Es ist ausreichend zu zeigen, dass für jedes $\epsilon > 0$ ein $M > 0$ derart existiert, dass $\mathrm{P}(|R_t| > M) \leq \epsilon$. Dazu betrachten wir die Einschränkungen auf die eingebetteten n-Koaleszenzprozesse $(R_t^{(n)})_{t\geq 0} \overset{\text{def}}{=} (\rho_n R_t)_{t\geq 0}$ und wählen exponentialverteilte Zufallsvariablen τ_n mit Parameter $\binom{n}{2}$. Die Wahrscheinlichkeit, dass $(R_t^{(n)})_{t\geq 0}$ zum Zeitpunkt t noch mehr als M Äquivalenzklassen hat, ergibt sich als die Wahrscheinlichkeit, dass die

Summe der ersten M Verweildauern des assoziierten Todesprozesse $(D_t^{(n)})_{t\geq 0}$ den Wert t nicht überschreitet. Vermöge der Markov-Ungleichung (vgl. Satz 17.4 in [2]) für die monotone identische Abbildung, erhalten wir

$$\begin{aligned}
\mathrm{P}\left(|R_t^{(n)}| > M\right) &= \mathrm{P}\left(\sum_{k=M}^{n} \tau_k > t\right) \\
&\leq \frac{1}{t}\,\mathrm{E}\left(\sum_{k=M}^{n} \tau_k\right) \\
&\leq \frac{1}{t}\sum_{k=M}^{\infty} \frac{2}{k(k-1)} \\
&= \frac{2}{t}\sum_{k=M}^{\infty}\left(\frac{1}{k-1} - \frac{1}{k}\right) = \frac{2}{t(M-1)}.
\end{aligned}$$

Daraus ergibt sich

$$\limsup_{n\to\infty} \mathrm{P}(|R_t^{(n)}| > M) \leq \frac{2}{t(M-1)}$$

und folglich $\limsup_{n\to\infty} \mathrm{P}(|R_t^{(n)}| > M) \leq \epsilon$ für alle $M > \lceil \frac{2}{\epsilon t} \rceil + 1$ und mithin

$$\mathrm{P}(|R_t| > M) \leq \epsilon.$$

Sei nun ϵ_n eine monotone Nullfolge. Dies liefert eine monoton wachsende Folge $(M_n)_n$ mit $\lim_{n\to\infty} M_n = \infty$. Ferner definiert $A_n \stackrel{\text{def}}{=} \{|R_t| > M_n\}$ eine antitone Mengenfolge mit $A \stackrel{\text{def}}{=} \lim_{n\to\infty} A_n = \{|R_t| = \infty\}$. Vermöge der Stetigkeit von oben (vgl. Satz 2.3 in [2]), gilt

$$\mathrm{P}(|R_t| = \infty) = \mathrm{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathrm{P}(A_n) \leq \lim_{n\to\infty} \epsilon_n = 0,$$

und damit die Behauptung. □

3.4. Genealogische Bäume

Als Prozess auf den Äquivalenzklassen von $\{1, \ldots, n\}$ ist ein n-Koaleszenzprozess eine relativ abstrakte Realisierung der vergleichsweise einfachen Sprungdynamik. Mit der Zerlegung des Prozesses in seine eingebettete diskrete Markov-Kette und zugehörigen Todesprozess (vgl. 3.6), lässt sich auf natürliche Weise ein n-Koaleszenzprozess $(R_t)_{t\geq 0}$ mit einem zufälligen, binären Wurzelbaum assoziieren, in welchem die Blätter mit den verschiedenen Äquivalenzklassen korrespondieren, und die Zweige proportional zu den sukzessiven Sprungzeiten von $(D_t)_{t\geq 0}$ dargestellt werden (siehe auch [19]), das heißt, wir können Koaleszenzprozesse als baumwertige Prozesse auffassen.

3.24 Definition
Ein gerichteter Graph G ist ein Paar (V, E), bestehend aus einer *Knotenmenge* V und einer *Kantenmenge* $E \subset V \times V$. Seien $v, w \in V$ zwei Knoten. Wir sagen dann, dass eine Kante von v nach w geht, falls $(v, w) \in V$. Da dies für unsere Fälle ausreichend ist, sei ferner G *schleifenfrei*, das heißt, $(v, v) \notin E$ für alle $v \in V$.

Wir sagen, dass für zwei Knoten $v, w \in V$ ein *Pfad* oder *Weg* von v nach w existiert, wenn eine Folge $(x_n)_{1 \leq n \leq k}$ mit $x_1 = v$ und $x_k = w$ existiert, so dass

$$(x_i, x_{i+1}) \in V \text{ für alle } i \in \{1, 2, \ldots, k-1\}$$

Ferner bezeichnen wir mit

$$\deg_{\text{out}}(v) \stackrel{\text{def}}{=} |\{w \in V : (v, w) \in E\}| \tag{3.23}$$

$$\deg_{\text{in}}(v) \stackrel{\text{def}}{=} |\{w \in V : (w, v) \in E\}| \tag{3.24}$$

den *Ausgangsgrad*, beziehungsweise den *Eingangsgrad* von v.

3.25 Definition
Ein *Wurzelbaum* ist ein gerichteter Graph $T = (V, E)$ mit einer ausgezeichneten Wurzel $r \in V$ derart, dass

(i) für alle $v \in V \setminus \{r\}$ ein Pfad von v nach r existiert, oder

(ii) für alle $v \in V \setminus \{r\}$ ein Pfad von r nach v existiert.

T heißt ferner *binär*, falls

(i') $\deg_{\text{in}}(v) \in \{0, 2\}$ und $\deg_{\text{out}}(v) \in \{0, 1\}$, im Fall (i),

(ii') $\deg_{\text{out}}(v) \in \{0, 2\}$ und $\deg_{\text{in}}(v) \in \{0, 1\}$, im Fall (ii),

für alle Knoten $v \in V$ gilt.

3.26 Definition
Wir betrachten an dieser Stelle einen schleifenfreien, ungerichteten Graph G als Paar (V, E), bestehend aus der Knotenmenge V und der Kantenmenge $E \subset \{\{v, w\} : v \neq w \in V\}$. Eine Kante ist demnach eine zweielementige Teilmenge der Knotenmenge. Wir bezeichnen mit

$$\deg(v) \stackrel{\text{def}}{=} |\{e \in E : v \in e\}|$$

den Knotengrad von v.

3.27 Bemerkung
Vermöge der Feststellung, dass für einen binären Wurzelbaum mit Wurzel $r \in V$

$$\deg_{\text{in}}(v) + \deg_{\text{out}}(v) \begin{cases} = 2, & \text{falls } v = r \\ \in \{1, 3\}, & \text{sonst.} \end{cases}$$

gilt, können wir einen binären Wurzelbaum auch als ungerichteten Graphen auffassen.

Ein binärer Wurzelbaum lässt sich demnach auf sehr natürliche Form als Darstellung einer Realisierung eines n-Koaleszenzprozesses $(R_t)_{t\geq 0}$ auffassen und konstruieren.

3.28 Bemerkung
Beginnend mit n Knoten, korrespondierend zu den Äquivalenzklassen von $R_0 = \Delta_n$ und einer leeren Kantenmenge, fügt man für jedes Koaleszenzereignis einen neuen Knoten, sowie neue Kanten zwischen diesem und den Knoten, welche zu den am Verschmelzungsereignis beteiligten Äquivalenzklassen assoziiert sind, ein. Offensichtlich liefert diese Konstruktion dann einen binären Wurzelbaum.

3.29 Definition
Sei $G = (V, E)$ ein schleifenfreier, ungerichteter Graph, sowie $|V| > 1$. Ein Knoten $v \in V$ heißt *äußerer* oder *externer* Knoten, falls $\deg(v) = 1$ gilt und *innerer* Knoten andernfalls.

Entsprechend heißt eine Kante *äußere* oder *externe Kante*, falls einer ihrer Endpunkte ein äußerer Knoten ist, und *innere Kante* andernfalls.

Ist G ein Baum, so bezeichnen wir Kanten auch als *Zweige* und externe Knoten als *Blätter*.

Die externen Knoten eines zur Realisierung eines n-Koaleszenzprozesses assoziierten binären Wurzelbaumes sind damit gerade die den einelementigen Äquivalenzklassen zugeordneten Knoten.

3.30 Definition
Ein gewichteter Graph W ist ein Tripel (V, E, w), bestehend aus einem Graphen $G = (V, E)$ und einer Gewichtsfunktion $w : E \to \mathbb{R}$.

Für eine positive Gewichtsfunktion können wir die Gewichte über die Kantenlängen darstellen. Darüber hinaus lässt sich die Konstruktion gemäß 3.28 einfach um die Konstruktion einer Gewichtsfunktion mithilfe der Sprungzeiten des zugrundeliegenden Markov-Sprungprozesses erweitern. Dabei entsteht ein Baum, in welchem alle externen Knoten denselben Abstand zur Wurzel haben.

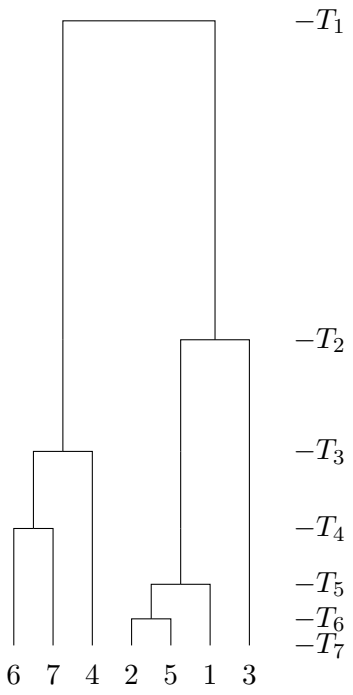

Abbildung 3.1.: Zu einem Koaleszenzprozess assoziierter zufälliger, binärer Wurzelbaum mit externen Knoten $1, 2, \ldots, 7$. Hierbei bezeichnen T_i die sukzessiven Sprungzeitpunkte, d.h. $T_{i+1} - T_i = \tau_i$.

3.31 Definition
Für einen n-Koaleszenzprozess $(R_t)_{t\geq 0}$ bezeichnen wir den zufälligen, binären Wurzelbaum, der sich gemäß des oben angegebenen, pfadweisen Konstruktionsverfahrens ergibt, als *Koaleszenzbaum* (zu $(R_t)_{t\geq 0}$).

3.32 Bemerkung
In Koaleszenzbäumen nimmt 3.11 die einfache Gestalt an, den unteren Teil des Baumes abzuschneiden.

3.33 Proposition
Für die Anzahl der Knoten eines binären Wurzelbaumes $T = (V, E)$ gilt:

(i) Es existiert ein $k \in \mathbb{N}_0$ mit $|V| = 2k - 1$.

(ii) Es existieren in diesem Fall genau k externe Knoten

Darüber hinaus gibt es dann auch genau $2k - 2$ Kanten.

Beweis. Die Aussagen ergeben sich leicht per Induktion aus der Überlegung, dass jeder Wurzelbaum $T = (V, E)$ mit Wurzelknoten r genau zwei Unterwurzelbäume T', T'' enthält. Die Wurzelknoten von T' und T'' sind gerade die zwei Kindknoten von r. Für die Knotenzahl von T' beziehungsweise T'' existieren dann per Induktionsvoraussetzung $k', k'' \in \mathbb{N}_0$, so dass wir

$$|V| = 1 + (2k' - 1) + (2k'' - 1) = 2(k' + k'') - 1$$

erhalten. Analog ergibt sich, dass die Anzahl externer Knoten gerade die Summe über die Anzahl externer Knoten der beiden Teilbäume ist. Als unweigerliche Konsequenz aus der Tatsache, dass an einem Verschmelzungsvorgang stets genau 2 Kanten beteiligt sind, wird die Wurzel des Baumes nach $k - 1$ Verschmelzungsvorgängen erreicht, also gibt es $2(k - 1)$ Kanten. □

3.34 Definition
Für einen gewichteten, binären Wurzelbaum $T = (V, E, w : E \to \mathbb{R}_{>0})$ mit Wurzel r, bezeichnen wir den Abstand zweier Knoten $v, w \in V$ als

$$\mathrm{d}(v, w) = \inf \left\{ \sum_k w(e_k) : (e_k)_k \subset E \text{ ist ein Pfad von v nach w} \right\},$$

wobei $\inf \emptyset = -\infty$ vereinbart sei und ferner die *Höhe* des Baumes mit

$$h(T) \stackrel{\text{def}}{=} \sup \{\mathrm{d}(v, r) : \deg(v) = 1\},$$

falls $|V| > 1$.

3.35 Definition
Wir bezeichnen einen gewichteten, binären Wurzelbaum $T = (V, E, w : E \to \mathbb{R}_{>0})$ mit Wurzel r und $|V| > 1$ als *regulär* beziehungsweise *zulässig*, falls

$$\mathrm{d}(v, r) = h(T)$$

für alle externen Knoten v gilt.

Demnach können wir jeden zulässigen, gewichteten, binären Wurzelbaum als Darstellung eines Pfades eines Koaleszenzprozesses auffassen. Entsprechend können wir ein Konstruktionsverfahren für zufällige, gewichtete binäre Wurzelbäume angeben, so dass ihre Verteilungen gerade denen von n-Koaleszenzprozessessen entsprechen (vgl. [19] oder [9])

Wir haben im Beweis von 3.33 gesehen, dass sich ein binärer Wurzelbaum als Baum, bestehend aus einem Wurzelknoten und zwei Subwurzelbäumen darstellen lässt. Entsprechend können wir zwei Wurzelbäume zu einem einzelnen Wurzelbaum mit neuer Wurzel wie folgt verschmelzen:

3.36 Proposition
Seien $S = (V_S, E_S, w_S : E_S \to \mathbb{R}_{>0})$ und $S' = (V_{S'}, E_{S'}, w_{S'} : E_{S'} \to \mathbb{R}_{>0})$ gewichtete, binäre Wurzelbäume mit $V_S \cap V_{S'} = \emptyset$ und Wurzeln s, s'. Sei ferner $r \notin V_S \cup V_{S'}$, sowie $g \in \mathbb{R}_{>0}$. Ohne Einschränkung sei darüber hinaus $h(S) \geq h(S')$ angenommen. Vermöge der Definitionen

$$V \stackrel{\text{def}}{=} \{r\} \cup V_S \cup V_{S'}$$
$$E \stackrel{\text{def}}{=} \{s, r\} \cup \{s', r\} \cup E_S \cup E_{S'}$$

$$w(e) \stackrel{\text{def}}{=} \begin{cases} w_S(e), & \text{für alle } e \in E_S, \\ w_{S'}(e), & \text{für alle } e \in E_{S'}, \\ g, & \text{für } e = \{s, r\}, \\ g + h(S) - h(S'), & \text{für } e = \{s', r\} \end{cases}$$

liefert dies einen gewichteten, binären Wurzelbaum $T = (V, E, w)$ mit Höhe $h(T) = h(S) + g$. Dieser ist ferner regulär, falls S, S' regulär sind.

Damit lässt sich nun folgendes Konstruktionsverfahren angeben:

(i) Beginne mit einem Wald $T_0 = (T_1, T_2, \ldots, T_n)$ von einelementigen Wurzelbäumen, etwa $T_i = (i, \emptyset, w_0^{(i)})$, wobei $w_0^{(i)}$ die leere Abbildung bezeichne.

(ii) Gegeben einen Wald $T_k = (T_1, T_2, \ldots, T_{n-k})$ regulärer, gewichteter, binärer Wurzelbäume, generiere (unabhängig) eine exponentialverteilte Zeit t_{n-k} mit Parameter $\binom{n-k}{2}$, und führe dann das obige Vorgehen 3.36 mit $g = t_{n-k}$ solange sukzessive durch, bis der Wald sich auf einen einzelnen Baum reduziert hat. Dies ist offenkundig nach $n - 1$ Schritten der Fall.

4. Längenverteilung externer Zweige in Koaleszenzbäumen

Wie wir in 3.20 gesehen haben, existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ und ein stochastischer Prozess $(R_t)_{t\geq 0}$ mit Werten in $\mathscr{E}$ derart, dass die kanonischen Einschränkungen $(\rho_n R_t)_{t\geq 0}$ auf den Äquivalenzrelationen der natürlichen Zahlen in die Äquivalenzrealtionen auf $\{1, 2, \ldots, n\}$ jeweils n-Koaleszenzprozesse sind. Ferner haben wir in 3.19 gesehen, dass $\mathscr{E}$ ein polnischer Raum ist. Dies ermöglicht es uns insbesondere Konvergenzaussagen über Eigenschaften von n-Koaleszenzprozessen zu treffen. Ebenfalls haben wir gesehen, dass sich jeder n-Koaleszenzprozess als zufälliger binärer Wurzelbaum auffassen lässt.

Wir folgen Caliebe et al. [7] in der Feststellung, dass sich die Länge eines zufällig gewählten externen Zweiges als zufällige Summe exponentialverteilter Zuwächse auffassen lässt und notieren dazu zunächst eine rekursive Darstellung.

Im Folgenden sei stets ein Koaleszenzprozess $(R_t)_{t\geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$ gegeben. Aussagen zu einem Koaleszenzbaum beziehen sich dann auf den zu $(\rho_n R_t)_{t\geq 0}$ assoziierten zufälligen Baum.

4.1 Bemerkung
Für einen n-Koaleszenzbaum bezeichne Z_n die Länge eines externen Zweiges, welcher zufällig aus den n externen Zweigen des Baumes ausgewählt wurde. Aus der Struktur des Koaleszenzbaumes wird klar, dass der so gewählte Zweig entweder am ersten Koaleszenzereignis beteiligt ist, und in dem Fall gilt $Z_n = T_n$, oder, dass er an einem späteren Koaleszenzereignis beteiligt ist, in welchem Fall $Z_n = T_n + R_n$ gilt, wobei R_n vermöge 3.11 gemäß Z_{n-1} verteilt ist, das heißt $R_n \stackrel{\sim}{=} Z_{n-1}$.

Daraus ergibt sich folgende Rekursion

$$Z_n \stackrel{\sim}{=} B_n Z_{n-1} + T_n, \qquad n \geq 3, \tag{4.1}$$

sowie $Z_2 \stackrel{\sim}{=} T_2$. Dabei sind für $n \in \{3, 4, \ldots, n\}$ die B_n bernoulliverteilt mit Parameter $1 - 2/n$) und die die T_n exponentialverteilt mit Parameter $\lambda_n = \binom{n}{2}$. Desweiteren sind $Z_2, B_3, \ldots, B_n, T_3, \ldots, T_n$ unabhängig.

4.2 Proposition
Für einen zufällig gewählten externen Zweig Z_n, $n \geq 3$ gilt für die Erwartung und die Varianz

$$\mathrm{E}Z_n = \frac{2}{n}, \tag{4.2}$$

$$\mathrm{Var}(Z_n) = \frac{8H_n - 12 + 4/n}{n(n-1)}, \tag{4.3}$$

wobei $H_n \stackrel{\text{def}}{=} \sum_{j=1}^{n} 1/j$ die n-te harmonische Zahl bezeichne.

Beweis. Wir setzen zunächst $X_n \stackrel{\text{def}}{=} n(n-1)Z_n$. Unter Erinnerung an die rekursive Struktur von Z_n und die Rechenregeln für Erwartungswerte gemäß (4.1) ergibt sich dann vermöge der Unabhängigkeit der $Z_2, B_3, \ldots, B_n, T_3, \ldots, T_n$

$$\begin{aligned}
\mathrm{E}X_n &= \mathrm{E}(n(n-1)Z_n) \\
&= n(n-1)\,\mathrm{E}(B_n Z_{n-1} + T_n) \\
&= n(n-1)\,\mathrm{P}(B_n = 0)\,\mathrm{E}(B_n Z_{n-1} + T_n | B_n = 0) \\
&\quad + n(n-1)\,\mathrm{P}(B_n = 1)\,\mathrm{E}(B_n Z_{n-1} + T_n | B_n = 1) \\
&= n(n-1)\left(\frac{2}{n}\,\mathrm{E}T_n + \frac{n-2}{n}\,\mathrm{E}(Z_{n-1} + T_n)\right) \\
&= n(n-1)\,\mathrm{E}T_n + \mathrm{E}X_{n-1} \\
&= 2 + \mathrm{E}X_{n-1}
\end{aligned} \tag{4.4}$$

Iteratives Einsetzen liefert $\mathrm{E}X_n = 2(n-1)$ und damit folgt

$$\mathrm{E}Z_n = \frac{\mathrm{E}X_n}{n(n-1)} = \frac{2}{n}.$$

Zur Berechnung der Varianz, sei auf den Hinweis verwiesen, dass es sich häufig als nützlich erweist, eine Funktion der eigentlich zu betrachtenden Zufallsvariablen zu untersuchen. Wie in [12] setzen wir nunmehr $Y_n \stackrel{\text{def}}{=} n(n-1)Z_n^2$ und erhalten

$$\begin{aligned}
\mathrm{E}Y_n &= \mathrm{E}[n(n-1)Z_n^2] \\
&= \mathrm{E}[n(n-1)(B_n Z_{n-1} + T_n)^2] \\
&= n(n-1)\,\mathrm{E}(B_n Z_{n-1} + T_n)^2
\end{aligned}$$

Bedingen unter B_n liefert dann

$$\begin{aligned}
\mathrm{E}Y_n &= n(n-1)\,\mathrm{P}(B_n=1)\,\mathrm{E}\left((B_nZ_{n-1}+T_n)^2|B_n=1\right)\\
&\quad + n(n-1)\,\mathrm{P}(B_n=0)\,\mathrm{E}\left((B_nZ_{n-1}+T_n)^2|B_n=0\right)\\
&= n(n-1)\left(\mathrm{P}(B_n=1)\,\mathrm{E}(T_n+Z_{n-1})^2+\mathrm{P}(B_n=0)\,\mathrm{E}T_n^2\right)\\
&= n(n-1)\left(\frac{n-2}{n}\,\mathrm{E}(T_n+Z_{n-1})^2+\frac{2}{n}\,\mathrm{E}T_n^2\right)\\
&= (n-1)(n-2)\,\mathrm{E}(T_n^2+2T_nZ_{n-1}+Z_{n-1}^2)+2(n-1)\,\mathrm{E}T_n^2\\
&= n(n-1)\,\mathrm{E}T_n^2+2(n-1)(n-2)\,\mathrm{E}T_nZ_{n-1}+\mathrm{E}Y_{n-1}. \qquad (4.5)
\end{aligned}$$

Nun gilt aber für die $\mathrm{Exp}(\lambda_n)$-verteilte Zufallsvariable T_n

$$\mathrm{E}T_n=(\lambda_n)^{-1}=\frac{2}{n(n-1)}, \qquad (4.6)$$

$$\mathrm{E}T_n^2=2(\lambda_n^2)^{-1}=\frac{8}{n^2(n-1)^2}. \qquad (4.7)$$

Ebenso ist Z_{n-1} als Verknüpfung von $Z_2, B_3, \ldots, B_{n-1}, T_3, \ldots, T_{n-1}$ unabhängig von T_n und es gilt mit Verweis auf (4.2)

$$\mathrm{E}T_nZ_{n-1}=\mathrm{E}T_n\,\mathrm{E}Z_{n-1}=\frac{2}{n(n-1)}\cdot\frac{2}{n-1}=\frac{4}{n(n-1)^2}.$$

Damit ergibt sich für die Erwartung von Y_n vermöge (4.5)

$$\begin{aligned}
\mathrm{E}Y_n &= \frac{8}{n(n-1)}+\frac{8(n-2)}{n(n-1)}+\mathrm{E}Y_{n-1}\\
&= \frac{8(n-1)}{n(n-1)}+\mathrm{E}Y_{n-1}\\
&= \frac{8}{n}+\mathrm{E}Y_{n-1}. \qquad (4.8)
\end{aligned}$$

Iteratives Einsetzen liefert dann

$$\mathrm{E}Y_n=\frac{8}{n}+\frac{8}{n-1}+\cdots+\frac{8}{2}=8(H_n-1), \qquad (4.9)$$

wobei für den letzten Summanden die Beziehung $\mathrm{E}Y_2=2\,\mathrm{E}Z_2^2=4$ verwendet wurde.

Unter Berücksichtigung der Linearität des Erwartungswertes erhalten wir daher

$$\mathrm{E}Z_n^2=\frac{\mathrm{E}Y_n}{n(n-1)}=\frac{8(H_n-1)}{n(n-1)},$$

und damit weiter für die Varianz von Z_n

$$\mathrm{Var}(Z_n) = \mathrm{E}Z_n^2 - (\mathrm{E}Z_n)^2 = \frac{8(H_n - 1)}{n(n-1)} - \frac{4}{n^2} = \frac{8H_n - 12 + 4/n}{n(n-1)}.$$

□

Wir folgen nunmehr Caliebe et al. [7] in der Feststellung, dass die standardisierte Länge eines zufällig gewählten externen Zweiges Z_n eines Koalesezenzbaumes keiner nichttrivialen Verteilung genügt, genauer gilt:

4.3 Proposition
Bezeichne Z_n die Länge eines zufällig gewählten externen Zweiges eines Koaleszenzbaumes, dann gilt

$$\frac{Z_n - EZ_n}{\sqrt{Var(Z_n)}} \xrightarrow{L_1} 0. \tag{4.10}$$

Genauer gilt sogar

$$\left\| \frac{Z_n - EZ_n}{\sqrt{Var(Z_n)}} \right\|_1 \in \mathcal{O}\left(\frac{1}{\sqrt{\ln n}} \right). \tag{4.11}$$

Beweis. Mit der trivialen Feststellung $Z_n \geq 0$, liefert eine Anwendung der Dreiecksungleichung für die L_1-Norm unter Berücksichtigung von $\|Z_n\|_1 = \mathrm{E}Z_n$

$$\left\| \frac{Z_n - \mathrm{E}Z_n}{\sqrt{\mathrm{Var}(Z_n)}} \right\|_1 \leq \frac{\|Z_n\|_1}{\sqrt{\mathrm{Var}(Z_n)}} + \frac{\mathrm{E}Z_n}{\sqrt{\mathrm{Var}(Z_n)}} = \frac{2\,\mathrm{E}Z_n}{\sqrt{\mathrm{Var}(Z_n)}}.$$

Einsetzen von Erwartungswert und Varianz gemäß 4.2 liefert ferner

$$\begin{aligned} \limsup_{n\to\infty} \frac{2EZ_n}{\sqrt{Var(Z_n)}} \sqrt{\ln n} &= \limsup_{n\to\infty} 4 \left(\frac{n^2(8H_n - 12 + 4/n)}{n(n-1)\ln n} \right)^{-\frac{1}{2}} \\ &\leq \limsup_{n\to\infty} 4 \left(\frac{n^2(8H_n - 12 + 4/n)}{n^2 \ln n} \right)^{-\frac{1}{2}}. \end{aligned}$$

Aus der bekannten Tatsache $\lim_{n\to\infty}(H_n - \ln n) = \gamma$, wobei γ die Euler-Mascheroni-Konstante bezeichne, erhält man $\lim_{n\to\infty} H_n / \ln n = 1$ und mithin

$$\limsup_{n\to\infty} \frac{2\,\mathrm{E}Z_n}{\sqrt{\mathrm{Var}(Z_n)}} \sqrt{\ln n} \leq \sqrt{2} < \infty,$$

und somit die zweite Behauptung. □

Insbesondere konvergiert damit die standardisierte Länge eines externen Zweiges in Verteilung gegen 0. Wir folgen weiter [7] in dem Bemühen, eine alternative Normalisierung zu finden.

4.1. Eine alternative Darstellung von Z_n

Sei $n \geq 3$, und B_n wie in (4.1). Setze weiterhin $B_2 \equiv 0$. Definiere nun

$$\tau_n \stackrel{\text{def}}{=} \min\{i \in \{0, \ldots, n-2\} : B_{n-1} = 0\}. \tag{4.12}$$

Dann können wir die Verteilung von τ_n unmittelbar berechnen und erhalten:

4.4 Proposition
Für die in (4.12) definierte Zufallsvariable τ_n gilt

$$\mathrm{P}(\tau_n = k) = 2\frac{n-k-1}{n(n-1)} = \frac{2}{n} - \frac{2k}{n(n-1)} \tag{4.13}$$

für $0 \leq k \leq n-2$.

Beweis. Aus der Unabhängigkeit der B_i folgt unmittelbar

$$\begin{aligned}
\mathrm{P}(\tau_n = k) &= \mathrm{P}(B_n = B_{n-1} = \cdots = B_{n-k+1} = 1, B_{n-k} = 0) \\
&= \left(\prod_{i=0}^{k-1} \mathrm{P}(B_{n-i} = 1)\right) \cdot \mathrm{P}(B_{n-k} = 0) \\
&= \left(\prod_{i=0}^{k-1} \frac{n-i-2}{n-i}\right) \frac{2}{n-k} = 2\frac{n-k-1}{n(n-1)}.
\end{aligned}$$

□

Im Hinblick auf (4.1) ermöglicht uns dies nun die Darstellung von Z_n als zufällige Summe der T_k. Genauer erhalten wir

$$Z_n \stackrel{\sim}{=} \sum_{k=0}^{\tau_n} T_{n-k}. \tag{4.14}$$

4.5 Proposition
Für das gemäß (4.12) definierte τ_n gilt

$$\frac{\tau_n}{n} \stackrel{\mathrm{d}}{\to} W, \tag{4.15}$$

wobei W eine Zufallsvariable auf $[0,1]$ mit Verteilungsfunktion

$$\mathrm{P}(W \leq x) = x(2-x), \quad x \in [0,1] \tag{4.16}$$

ist. Für die zu W gehörende Dichte gilt

$$f_W(x) = 2(1-x), \quad x \in [0,1]. \tag{4.17}$$

Beweis. Es bezeichne $\lfloor x \rfloor$ die untere Gaußklammer von x, das heißt $\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \leq x\}$. Für $x \in [0,1)$ folgt mittels (4.12) und der einfachen Überlegung $\lfloor nx \rfloor \leq n-1$ für $x \in [0,1)$, sowie der Tatsache, dass die rechte Seite in (4.13) auch mit $k = n-1$ verträglich ist,

$$\begin{aligned} \mathrm{P}(\tau_n \leq nx) &= \sum_{k=0}^{\lfloor nx \rfloor} \mathrm{P}(\tau_n = k) \\ &= \sum_{k=0}^{\lfloor nx \rfloor} \left(\frac{2}{n} - \frac{2k}{n(n-1)} \right) \\ &= \frac{2(\lfloor nx \rfloor + 1)}{n} - \frac{\lfloor nx \rfloor (\lfloor nx \rfloor + 1)}{n(n-1)} \to 2x - x^2 = x(2-x), \end{aligned}$$

für $n \to \infty$. Für $x = 1$ ergibt sich $\mathrm{P}(\tau_n \leq n) = 1 = 2 - 1^2$. Damit folgt die gewünschte Verteilungskonvergenz gegen eine Zufallsvariable W auf $[0,1]$ mit der in (4.16) gewünschten Verteilungsfunktion.

Die Behauptung für die zu W gehörende Dichte erschließt sich leicht durch Ableitung der Verteilungsfunktion von W

$$\frac{\mathrm{d}}{\mathrm{d}x} x(2-x) = x(1-x).$$

□

4.6 Proposition
Für den Erwartungswert und die Varianz von τ_n beziehungsweise W gelten

$$\mathrm{E}\tau_n = \frac{n-2}{3}, \qquad \mathrm{E}\left(\frac{\tau_n}{n}\right) \to \mathrm{E}W = \frac{1}{3}, \tag{4.18}$$

$$\mathrm{Var}(\tau_n) = \frac{n^2 - n - 2}{18}, \qquad \mathrm{Var}\left(\frac{\tau_n}{n}\right) \to \mathrm{Var}(W) = \frac{1}{18}. \tag{4.19}$$

Beweis. Gemäß der Definition des Erwartungswertes und unter Verwendung von (4.13)

und A.4, gilt

$$\begin{aligned}
\mathrm{E}(\tau_n) &= \sum_{k=0}^{n-2} k\,\mathrm{P}(\tau_n = k) \\
&= \sum_{k=0}^{n-2} 2k\frac{n-k-1}{n(n-1)} \\
&= \frac{2(n-1)}{n(n-1)}\sum_{k=0}^{n-2} k - \frac{2}{n(n-1)}\sum_{k=0}^{n-2} k^2 \\
&= \frac{2}{n}\frac{(n-2)(n-1)}{2} - \frac{2}{n(n-1)}\frac{(n-2)(n-1)(2(n-2)+1}{6} \\
&= \frac{n-2}{3}
\end{aligned}$$

Der Erwartungswert $\mathrm{E}W$ von W berechnet sich gemäß (4.17) zu

$$\begin{aligned}
\mathrm{E}W &= \int_0^1 2x(1-x)\,\mathrm{d}x \\
&= 2\left(\frac{1}{2}x^2\right|_0^1 - 2\left(\frac{1}{3}x^3\right|_0^1 = \frac{1}{3}.
\end{aligned}$$

Es reicht nun vermöge der Linearität des Erwartungswertes der Vermerk

$$\mathrm{E}\left(\frac{\tau_n}{n}\right) = \frac{\mathrm{E}\tau_n}{n} = \frac{1}{3} - \frac{2}{3n} \to \frac{1}{3}$$

für $n \to \infty$.

Für die Varainz von τ_n betrachten wir zunächst $\mathrm{E}(\tau_n^2)$. Mit Verweis auf (A.5) für die Summe von dritten Potenzen folgt

$$\begin{aligned}
\mathrm{E}(\tau_n^2) &= \sum_{k=0}^{n-2} k^2\,\mathrm{P}(\tau_n = k) \\
&= \sum_{k=0}^{n-2} 2k^2\frac{n-k-1}{n(n-1)} \\
&= \frac{2}{n}\sum_{k=0}^{n-2} k^2 - \frac{2}{n(n-1)}\sum_{k=0}^{n-2} k^3 \\
&= \frac{2}{n}\frac{(n-2)(n-1)(2(n-2)+1)}{6} - \frac{2}{n(n-1)}\frac{(n-2)^2(n-1)^2}{4} \\
&= \frac{(n-1)(n-2)}{6} = \frac{n^2-3n+2}{6}
\end{aligned} \tag{4.20}$$

Für $\operatorname{Var}(\tau_n)$ gilt somit

$$\begin{aligned}\operatorname{Var}(\tau_n) &= \operatorname{E}(\tau_n^2) - (\operatorname{E}\tau_n)^2 \\ &= \frac{(n-1)(n-2)}{6} - \frac{(n-2)^2}{9} \\ &= \frac{(n-2)(n+1)}{18} = \frac{n^2-n-2}{18}.\end{aligned}$$

Die Varianz von W berechnet sich ähnlich der obigen Berechnung des Erwartungswertes von W zu

$$\begin{aligned}\operatorname{Var}(W) &= \operatorname{E}W^2 - (\operatorname{E}W)^2 \\ &= \int_0^1 2x^2(1-x)\,\mathrm{d}x \\ &= 2\left(\frac{1}{3}x^3 - \frac{1}{4}x^4\Big|_0^1 - \frac{1}{9} = \frac{1}{18},\right.\end{aligned}$$

und auch in diesem Fall reicht der Vermerk von

$$\operatorname{Var}\left(\frac{\tau_n}{n}\right) = \frac{\operatorname{Var}(\tau_n)}{n^2} = \frac{1}{18}\left(1 - \frac{1}{n} - \frac{2}{n^2}\right) \to \frac{1}{18}$$

für $n \to \infty$. □

4.7 Bemerkung
In [7] finden wir anstelle von (4.19) mit Verweis auf [6] die Behauptung, dass für die Varianz von τ_n

$$\operatorname{Var}(\tau_n) = \frac{n^2+31n-2}{18}$$

gelte. Dabei betrachteten Blum und François eine Zufallsvariable K, das *Koaleszenzlevel* von Individuum 1, mit $K = n - \tau_n$. Insbesondere bedeutet $K = k$ also, dass das Individuum 1 beim Übergang von $k \to (k-1)$ Äquivalenzklassen zum ersten Mal an einer Verschmelzung beteiligt ist. Bezugnehmend auf (4.13), (4.18) und (4.19) notieren wir

$$\operatorname{P}(K=k) = \operatorname{P}(\tau_n = n-k) = \frac{2}{n}\frac{k-1}{n-1}$$

für $k \in \{2,3,\ldots,n\}$ und $\operatorname{E}K = n - \operatorname{E}\tau_n = 2/3(n+1)$ in Übereinstimmung mit [6], aber

$$\operatorname{Var}(K) = \operatorname{Var}(\tau_n) = \frac{1}{18}(n^2+n-2)$$

als mögliche Korrektur.

4.8 Proposition
Für die Verteilungsfunktionen $F_n(x) \stackrel{\text{def}}{=} \mathrm{P}(\tau_n/n \leq x)$ und $F(x) \stackrel{\text{def}}{=} \mathrm{P}(W \leq x)$, $x \in [0,1]$, von τ_n/n beziehungseise W gilt

$$\|F_n - F\|_\infty = |F_n(0) - F(0)| = F_n(0) = \frac{2}{n}. \tag{4.21}$$

Beweis. Wir beginnen mit der einfachen Überlegung, dass die einzigen Punkte, für die die Supremumsnorm angenommen werden kann, vermöge der strengen Monotonie von F gerade die Unstetigkeitsstellen von F_n sind. Daher betrachten wir

$$\begin{aligned} \mathrm{P}\left(\frac{\tau_n}{n} \leq \frac{i}{n}\right) &= \sum_{k=0}^{i} \left(\frac{2}{n} - \frac{2k}{n(n-1)}\right) \\ &= \frac{2(i+1)}{n} - \frac{i(i+1)}{n(n+1)} \end{aligned} \tag{4.22}$$

für $i \in \{0, 1, \ldots, n-2\}$ und erhalten weiterhin

$$\mathrm{P}\left(\frac{\tau_n}{n} \leq \frac{i}{n}\right) - \mathrm{P}\left(W \leq \frac{i}{n}\right) = \frac{2}{n} - \frac{i(n-i)}{n^2(n+1)}, \tag{4.23}$$

womit die Behauptung bewiesen ist. □

Damit erhalten wir die gleichmäßige Konvergenz von $F_n(x) \to F(x)$, welche schon durch die Stetigkeit von $F(x)$ gegeben ist, und darüber hinaus auch einen Hinweis auf die Geschwindigkeit der Konvergenz der Verteilungsfunktionen. Insbesondere gilt $\|F_n - F\|_\infty \in \mathcal{O}(1/n)$.

Wir folgen weiter Caliebe et al. [7] in ihrem Beweis zur asymptotischen Verhalten von Z_n und notieren dazu analog zunächst das folgende Lemma:

4.9 Lemma
Für festes $t > 0$ und $n/\ln n \leq j \leq n - n/\ln n$ gilt für $n \to \infty$

$$\sum_{l=n-1}^{n} \ln\left(1 + \frac{2tn}{l(l-1)}\right) = \frac{2tj}{n-j} + \mathcal{O}\left(\frac{\ln^3 n}{n}\right), \tag{4.24}$$

wobei der Fehler unabhängig von j gewählt werden kann, aber von t abhängt.

Beweis. Definiere Restglieder R_n gemäß

$$\sum_{l=n-j}^{n} \ln\left(1 + \frac{2tn}{l(l-1)}\right) = \sum_{l=n-j}^{n} \frac{2tn}{l(l-1)} + R_n,$$

das heißt R_n ist von der Gestalt

$$R_n = \sum_{l=n-j}^{n} \ln\left(1 + \frac{2tn}{l(l-1)}\right) - \frac{2tn}{l(l-1)}.$$

Unter Verwendung von A.2 (jeder Summand ist von der Gestalt $\ln(1+x)$ für ein $x > 0$) ergibt sich

$$\begin{aligned}
|R_n| &\leq \frac{1}{2} \sum_{l=n-j}^{n} \left(\frac{2tn}{l(l-1)}\right)^2 = 2t^2n^2 \sum_{l=n-j}^{n} \frac{1}{l(l-1)} \\
&\leq 2t^2n^2 \sum_{l=n-j}^{n} \frac{1}{(l-1)^2} \leq 2t^2n^2 \sum_{l=n-j}^{n} \frac{1}{(l-1)^4} \\
&\leq 2t^2n^2 \sum_{l=n-j}^{\infty} \frac{1}{(l-1)^4} \leq 2t^2n^2 \int_{n-j-1}^{\infty} \frac{1}{(x-1)^4} dx \qquad (4.25) \\
&= \frac{2t^2n^2}{3(n-j-2)^3} \leq \frac{2}{3} \frac{t^2n^2}{(n/\ln n - 2)^3} \qquad (4.26) \\
&\leq \frac{t^2n^2}{n^3/\ln^3 n} = \frac{t^2 \ln^3 n}{n},
\end{aligned}$$

wobei in (4.25) verwendet wurde, dass $x \to 1/(x-1)^4$ monoton fallend für $x > 1$, sowie $n/\ln n \geq 2$ für $n \geq 2$ ist. Die Ungleichung in (4.26) beruht auf einer Abschätzung von j nach oben, und der letzte Schritt verwendet erneut die Beziehung $n/\ln n \geq 2$ für $n \geq 2$. Mittels dieser Abschätzung ergibt sich dann

$$\limsup_{n\to\infty} \left| R_n \frac{n}{\ln^3 n} \right| \leq \limsup_{n\to\infty} \left| \frac{t^2 \ln^3 n}{n} \frac{n}{\ln^3 n} \right| = t^2 < \infty$$

Damit reicht nun zu zeigen, dass

$$\left| \frac{j}{n-j} - \sum_{l=n-j}^{n} \frac{n}{l(k-1)} \right| \in \mathcal{O}\left(\frac{\ln^3 n}{n}\right)$$

gilt, denn dann folgt

$$\begin{aligned}\sum_{l=n-j}^{n} \ln\left(1+\frac{2tn}{l(l-1)}\right) &= \sum_{l=n-j}^{n} \frac{2tn}{l(l-1)} + R_n \\ &= 2t\Bigg(\frac{j}{n-j}\underbrace{-\frac{j}{n-j}+\sum_{l=n-j}^{n}\frac{n}{l(l-1)}}_{\in\mathcal{O}\left(\frac{\ln^3 n}{n}\right)}\Bigg) + \underbrace{R_n}_{\in\mathcal{O}\left(\frac{\ln^3 n}{n}\right)} \\ &= \frac{2tj}{n-j} + \mathcal{O}\left(\frac{\ln^3 n}{n}\right)\end{aligned}$$

und damit die Behauptung. Deshalb betrachten wir nunmehr

$$\begin{aligned}\sum_{l=n-j}^{n}\frac{n}{l(l-1)} &= n\sum_{l=n-j}^{n}\left(\frac{1}{l-1}-\frac{1}{l}\right) \\ &= n\left(\frac{1}{n-j-1}-\frac{1}{n}\right) \\ &= \frac{j+1}{n-j-1}.\end{aligned}$$

Insbesondere ist $j/(n-j) \leq (j+1)/(n-j)$, weshalb nun

$$\begin{aligned}\left|\frac{j}{n-j}-\sum_{l=n-j}^{n}\frac{n}{l(l-1)}\right| &= \frac{j+1}{n-j-1}-\frac{j}{n-j} \\ &= \frac{n}{(n-j)(n-j-1)} \\ &\leq \frac{n}{(n-n+\frac{n}{\ln n})(n-n+\frac{n}{\ln n}-1)} \\ &= n\left(\frac{n^2}{\ln^2 n}-\frac{n}{\ln n}\right)^{-1}\end{aligned} \tag{4.27}$$

folgt. Damit gilt für

$$\begin{aligned}\limsup_{n\to\infty}\left|\frac{j}{n-j}-\sum_{l=n-j}^{n}\frac{n}{l(l-1)}\right|\frac{n}{\ln^3 n} &\leq \limsup_{n\to\infty}\left|\frac{n^2}{\ln^3 n}\left(\frac{n^2}{\ln^2 n}-\frac{n}{\ln n}\right)^{-1}\right| \\ &= \lim_{n\to\infty}\left|\left(\ln n-\ln^2 n\right)^{-1}\right| = 0 < \infty.\end{aligned}$$

Insbesondere kann der Fehler, bestehend aus R_n und der linken Seite in (4.27) unabhängig von j abgeschätzt werden, hängt aber von t ab. □

Wir können nun Caliebe et al. in ihrem Resultat über die Verteilungskonvergenz externer Zweige von Koaleszenzbäumen mithilfe ihrer Laplace-Transformierten folgen und notieren:

4.10 Satz
Gegeben externe Zweige in Koaleszenzbäumen mit Längen Z_n, konvergiert nZ_n für $n \to \infty$ in Verteilung gegen $Z \stackrel{\text{def}}{=} 2W/(1-W)$, wobei W durch (4.15) gegeben ist. Damit ist die zu Z gehörende Dichtefunktion von der Form $x \to 8/(2+x)^3, x \geq 0$.

Beweis. Wir bezeichnen mit φ_n die Laplace-Transformierte von nZ_n, das heißt, $\varphi_n(t) = \mathrm{E}\exp(-tnZ_n), t > 0$. Mit der im folgenden gezeigten punktweisen Konvergenz von φ_n gegen die Laplace-Transformierte φ von Z, folgt die Verteilungskonvergenz von nZ_n gegen Z (vgl. Satz 42.4 in [2]).

Seien dazu $I_n \stackrel{\text{def}}{=} [n/\ln n, n - n/\ln n]$ und $A_n \stackrel{\text{def}}{=} \{\tau_n \in I_n\}$ mit τ_n gemäß (4.12)

In einem ersten Schritt vermerken wir $\mathrm{P}(A_n^c) \in \mathcal{O}(1/\ln n)$, denn

$$\begin{aligned}\mathrm{P}(A_n^c) = \mathrm{P}(\tau_n \notin I_n) &= \mathrm{P}\left(\frac{\tau_n}{n} \notin \left[\frac{1}{\ln n}, 1 - \frac{1}{\ln n}\right]\right) \\ &= 1 - \mathrm{P}\left(\frac{\tau_n}{n} \leq 1 - \frac{1}{\ln n}\right) + \mathrm{P}\left(\frac{\tau_n}{n} \leq \frac{1}{\ln n}\right).\end{aligned}$$

Vermöge 4.8 erhalten wir für alle $x \in [0,1]$

$$\mathrm{P}\left(W \leq x\right) - \frac{2}{n} \leq \mathrm{P}\left(\frac{\tau_n}{n} \leq x\right) \leq \mathrm{P}\left(W \leq x\right) + \frac{2}{n}, \tag{4.28}$$

und damit bereits

$$\begin{aligned}\mathrm{P}(A_n^c) &\leq 1 - \mathrm{P}\left(W \leq 1 - \frac{1}{\ln n}\right) + \mathrm{P}\left(W \leq \frac{1}{\ln n}\right) + \frac{4}{n} \\ &= \mathrm{P}\left(W \notin \left[\frac{1}{\ln n}, 1 - \frac{1}{\ln n}\right]\right) + \frac{4}{n} \\ &= \frac{2}{\ln n} + \frac{4}{n}. \end{aligned} \tag{4.29}$$

Also gilt $\limsup_{n\to\infty} |\mathrm{P}(A_n^c) \cdot \ln n| \leq \limsup_{n\to\infty} |2 + 4\ln n/n| = 2 < \infty$, und mithin $\mathrm{P}(A_n^c) \in \mathcal{O}(1/\ln n)$.

Für die Laplace-Transformierte $\varphi_n(t)$ von Z_n bedeutet dies

$$\begin{aligned}\varphi_n(t) &= \mathrm{E}\exp(-tnZ_n) = \mathrm{E}\exp\left(-tn\sum_{k=0}^{\tau_n} T_{n-k}\right)\\ &= \int_{A_n\cup A_n^c} \exp\left(-tn\sum_{k=0}^{\tau_n} T_{n-k}\right).\end{aligned}$$

Es reicht ferner, das Integral über A_n zu betrachten, da wir vermöge der Beschränkung des Integranden nach oben durch 1 nunmehr

$$\int_{A_n^c} \exp\left(-tnZ_n\right) \leq \int_{A_n^c} 1 = \mathrm{P}(A_n^c) \in \mathcal{O}\left(\frac{1}{\ln n}\right)$$

erhalten. Für die Laplace-Transformierte einer $\mathrm{Exp}(\lambda)$-verteilten Zufallsvariablen X, $\lambda > 0$ verweisen wir auf Seite 282 in [2] und notieren,

$$\varphi_X(t) = \mathrm{E}\exp(-tX) = \frac{\lambda}{t+\lambda}, \quad t > 0.$$

Damit gilt vermöge der Unabhängigkeit der T_i

$$\begin{aligned}\int_{A_n} \exp\left(-tn\sum_{k=0}^{\tau_n} T_{n-k}\right) &= \sum_{j\in I_n\cap\mathbb{N}} \mathrm{P}(\tau_n = j)\prod_{k=0}^{j} \mathrm{E}\exp\left(-tnT_{n-k}\right)\\ &= \sum_{j\in I_n\cap\mathbb{N}} \mathrm{P}(\tau_n = j)\prod_{k=0}^{j} \frac{\lambda_{n-k}}{\lambda_{n-k}+tn}\\ &= \sum_{j\in I_n\cap\mathbb{N}} \mathrm{P}(\tau_n = j)\exp\left(-\sum_{k=0}^{j}\ln\left(1+\frac{tn}{\lambda_{n-k}}\right)\right)\\ &= \sum_{j\in I_n\cap\mathbb{N}} \mathrm{P}(\tau_n = j)\exp\left(-\sum_{l=n-j}^{n}\ln\left(1+\frac{tn}{\lambda_l}\right)\right),\end{aligned}$$

wobei im vorletzten Schritt die Beziehungen $xy = e^{\ln x}e^{\ln y} = \exp(\ln x + \ln y)$, sowie $\ln(x) = -\ln(1/x)$ verwendet wurden, sowie im letzten Schritt die Summe umsortiert wurde. In Erinnerung an $\lambda_l = \binom{l}{2} = \frac{l(l-1)}{2}$ erfüllt der (negative) Exponent gerade die Voraussetzungen von 4.9. Dies liefert

$$\begin{aligned}\int_{A_n} \exp\left(-tn\sum_{k=0}^{\tau_n} T_{n-k}\right) &= \sum_{j\in I_n\cap\mathbb{N}} \mathrm{P}(\tau_n = j)\exp\left(-\frac{2tj}{n-j}+\mathcal{O}\left(\frac{\ln^3 n}{n}\right)\right)\\ &= \exp\left(\mathcal{O}\left(\frac{\ln^3 n}{n}\right)\right)\int_{A_n}\exp\left(-t\frac{2\tau_n/n}{1-\tau_n/n}\right),\end{aligned}$$

wobei im letzten Schritt verwendet wurde, dass der Fehler unabhängig von j abgeschätzt werden kann.

Es gilt nun vermöge A.1 die Konvergenz von $\exp\left(\mathcal{O}\left(\frac{\ln^3 n}{n}\right)\right) \to 1$, für $n \to \infty$. Unter Verwendung von (4.15) und der asymptotischen Vernachlässigbarkeit von $\mathrm{P}(A_n^c)$ ergibt sich insgesamt die Konvergenz der Laplace-Transformierten von Z_n gegen die Laplace-Transformierte von Z.

Zur Bestimmung der Dichte von Z berechnen wir im Weiteren zunächst die Verteilungsfunktion von Z mithilfe der Verteilung von W gemäß

$$\begin{aligned}\mathrm{P}(Z \leq x) &= \mathrm{P}\left(\frac{2W}{1-W} \leq x\right) = \mathrm{P}\left(W \leq \frac{x}{2+x}\right) \\ &= \frac{2x}{2+x} - \left(\frac{x}{2+x}\right)^2. \end{aligned} \tag{4.30}$$

Anschließendes Ableiten liefert dann die Dichte von Z

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{2x}{2+x} - \left(\frac{x}{2+x}\right)^2\right] = \frac{4}{(2+x)^2} - \frac{2x}{2+x}\frac{2}{(2+x)^2} = \frac{8}{(2+x)^3},$$

und damit die Behauptung. □

Im Hinblick auf Anwendungsbeispiele, notieren wir gemäß [7], dass die Konvergenz von nZ_n auch unter geringeren Anforderungen an die Sprungzeiten des zugrundeliegenden Koaleszenzprozesses gewährleistet ist. Im Besonderen reicht es aus, wenn die zugehörigen Erwartungen mit denen der entsprechenden Exponentialverteilungen übereinstimmen, und die Varianz nicht zu groß wird.

4.11 Satz
Die Aussage von 4.10 gilt auch, wenn man die Verteilungen der T_i lediglich vorausgesetzt wird, dass

$$\mathrm{E}T_i = \frac{2}{i(i-1)}, \tag{4.31}$$

$$\lim_{n\to\infty} n^2 \sum_{k=0}^{\lfloor nx \rfloor} \mathrm{Var}(T_{n-k}) = 0 \quad \forall x \in (0,1). \tag{4.32}$$

gilt.

Beweis. Wir beginnen mit der Feststellung, dass für festes $x_0 \in (0,1)$ und $m \in \mathbb{N}$ stets ein $n_0 \in \mathbb{N}$ zu finden ist, so dass $n - \lfloor nx_0 \rfloor \geq m$ für alle $n \geq n_0$ gilt. Vermöge der

Definition der unteren Gaußklammer erhalten wir $\lfloor y \rfloor \leq y$ für alle $y \in \mathbb{R}$ und damit $n - \lfloor nx_0 \rfloor \geq n - nx_0 = n(1 - x_0)$. Es reicht nun die Bemerkung, dass $y \to (1 - x_0)y$ eine lineare Funktion mit positiver Steigung ist.

Wir betrachten erneut (4.14) und erhalten

$$\begin{aligned} nZ_n &\stackrel{\sim}{=} \sum_{k=0}^{\tau_n} nT_{n-k} \\ &= \sum_{k=0}^{\tau_n} (nT_{n-k} - \mathrm{E}nT_{n-k}) + \sum_{k=0}^{\tau_n} \mathrm{E}nT_{n-k} \\ &= \mathrm{I} + \mathrm{II}. \end{aligned} \tag{4.33}$$

Im Folgenden wird zunächst die Konvergenz von I gegen 0 in Wahrscheinlichkeit gezeigt. Dazu sei $\epsilon > 0$ und $x_0 \in (0, 1)$, dann gilt für die Wahrscheinlichkeit, dass I einen Wert ungleich 0 annimmt

$$\mathrm{P}\left(|\mathrm{I}| > \epsilon\right) = \mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n > x_0\}}\right| > \epsilon\right) + \mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n \leq x_0\}}\right| > \epsilon\right).$$

Seien ferner $\delta > 0$ und $x_0 \in (0, 1)$ so gewählt, dass $\mathrm{P}(W > x_0) \leq \frac{\delta}{4}$ gilt. Dies ist möglich, da W eine Verteilung auf $[0, 1]$ ist. Aus der Verteilungskonvergenz der τ_n/n gegen W in Verteilung, folgt weiter die Existenz einer natürlichen Zahl $m_1 \in \mathbb{N}$ derart, dass für alle $n > m_1$

$$\left| \mathrm{P}\left(\frac{\tau_n}{n} > x_0\right) - \mathrm{P}\left(W > x_0\right) \right| \leq \frac{\delta}{4}$$

gilt und damit insbesondere

$$\begin{aligned} \mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n > x_0\}}\right| > \epsilon\right) &\leq \mathrm{P}\left(\frac{\tau_n}{n} > x_0\right) \\ &= \left| \mathrm{P}\left(\frac{\tau_n}{n} > x_0\right) - \mathrm{P}\left(W > x_0\right) + \mathrm{P}\left(W > x_0\right) \right| \\ &\leq \frac{\delta}{4} + \frac{\delta}{4} = \frac{\delta}{2} \end{aligned}$$

vermöge der Dreiecksungleichung für Beträge.

Für hinreichend großes n ergibt sich nun

$$\begin{aligned}
\mathrm{E}\left(\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\right)^2 &= \mathrm{E}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\left(\sum_{n=0}^{\tau_n}(nT_{n-k}-\mathrm{E}nT_{n-k})\right)^2 \\
&= \mathrm{E}\sum_{l=0}^{\lfloor nx_0\rfloor}\mathbb{1}_{\{\tau_n=l\}}\left(\sum_{k=0}^{l}(nT_{n-k}-\mathrm{E}nT_{n-k})\right)^2 \\
&= n^2\,\mathrm{E}\sum_{l=0}^{\lfloor nx_0\rfloor}\mathbb{1}_{\{\tau_n=l\}}\sum_{k=0}^{l}(T_{n-k}-\mathrm{E}T_{n-k})^2 \\
&\quad + 2\sum_{j>k}^{l}(T_{n-k}-\mathrm{E}T_{n-k})\,(T_{n-j}-\mathrm{E}T_{n-j}) \\
&= n^2\sum_{l=0}^{\lfloor nx_0\rfloor}\mathrm{P}(\tau_n=l)\left(\sum_{k=0}^{l}\mathrm{Var}(T_{n-k})+2\sum_{j>k}^{l}\mathrm{Cov}(T_{n-k},T_{n-j})\right) \\
&\leq n^2\sum_{k=0}^{\lfloor nx_0\rfloor}\mathrm{Var}(T_{n-k}),
\end{aligned}$$

wobei benutzt wurde, dass die T_i stochastisch unabhängig und ferner unabhängig von τ_n sind. Unter Verwendung der Voraussetzung an die Varianzen folgt damit die Konvergenz von $\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}} \overset{L_2}{\to} 0$ für $n\to\infty$. Eine Anwendung der Markov-Ungleichung liefert desweiteren damit auch die stochastische Konvergenz von $\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}$ gegen 0 vermöge

$$\mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\right|\geq\nu\right)\leq\frac{1}{\nu^2}\,\mathrm{E}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\right|\right)^2\to 0$$

für $n\to\infty$ und $\nu>0$ und damit die Existenz einer natürlichen Zahl $m_2\in\mathbb{N}$ derart, dass für alle $n>m_2$

$$\mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\right|>\epsilon\right)\leq\frac{\delta}{2}$$

gilt. Damit gilt

$$\mathrm{P}\left(|\mathrm{I}|>\epsilon\right)\leq\mathrm{P}\left(\left|\mathrm{I}\mathbb{1}_{\{\tau_n/n\leq x_0\}}\right|>\epsilon\right)+\mathrm{P}\left(\frac{\tau_n}{n}>x_0\right)\leq\delta,$$

für alle $n>\max\{m_1,m_2\}$. Dabei sind m_1,m_2 abhängig von δ, aber da $\delta>0$ beliebig war, folgt die stochastische Konvergenz von I gegen 0.

Für II gilt ferner nach Voraussetzung

$$\begin{aligned}
\text{II} &= n\sum_{k=0}^{\tau_n} \mathrm{E}T_{n-k} \\
&= n\sum_{k=0}^{\tau_n} \frac{2}{(n-k)(n-k-1)} \\
&= 2n\sum_{k=0}^{\tau_n} \left(\frac{1}{n-k-1} - \frac{1}{n-k}\right) \\
&= 2n\left(\frac{1}{n-\tau_n-1} - \frac{1}{n}\right) \\
&= \frac{2}{1-\tau_n/n-1/n} - 2 \xrightarrow{\mathrm{d}} \frac{2}{1-W} - 2 = Z,
\end{aligned}$$

also die Verteilungskonvergenz von II gegen Z und damit zusammen mit der stochastischen Konvergenz von I gegen 0 ebenso die Verteilungskonvergenz von $nZ_n \xrightarrow{\mathrm{d}} Z$ für $n \to \infty$ vermöge des Satzes von Slutsky (vgl. Satz 36.12 in [2]). □

Der Beweis basiert also auf der Aufteilung der alternativen Darstellung gemäß (4.14) in zwei Teile, von denen einer stochastisch gegen 0 konvergiert, und der andere gerade in Verteilung gegen die Grenzverteilung Z. Für die Varianz von nZ_n notieren wir numehr:

4.12 Bemerkung
Gegeben die Voraussetzungen und Bezeichnungen aus 4.11, so gilt

$$\mathrm{Var}(nZ_n) = \mathrm{Var}(\mathrm{I}) + \mathrm{Var}(\mathrm{II}). \tag{4.34}$$

Wir beweisen stattdessen allgemeiner:

4.13 Proposition
Sei $(\Omega, \mathfrak{A}, \mathrm{P})$ ein Wahrscheinlichkeitsraum und $X = (X_k)_{k\in\mathbb{N}_0} : (\Omega, \mathfrak{A}, \mathrm{P}) \to (\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ eine Familie unabhängiger, quadratisch integrierbarer Zufallsvariablen, und desweiteren $N : (\Omega, \mathfrak{A}, \mathrm{P}) \to (\mathbb{N}_0, \mathcal{P}(\mathbb{N}_0))$ eine von $(X_k)_k$ unabhängige, integrierbare Zufallsvariable mit Werten in $\mathbb{N}_0$. Ist N beschränkt, so gilt

$$\mathrm{Var}\left(\sum_{k=0}^{N} X_k\right) = \mathrm{Var}\left(\sum_{k=0}^{N}(X_k - \mathrm{E}X_k)\right) + \mathrm{Var}\left(\sum_{k=0}^{N} \mathrm{E}X_k\right). \tag{4.35}$$

Darüber hinaus erhalten wir ferner

$$\operatorname{Var}\left(\sum_{k=0}^{N}(X_k - \operatorname{E}X_k)\right) = \operatorname{E}_N\left(\sum_{k=0}^{N}\operatorname{Var}(X_k)\right) \tag{4.36}$$

$$\operatorname{Var}\left(\sum_{k=0}^{N}\operatorname{E}X_k\right) = \operatorname{Var}_N\left(\sum_{k=0}^{N}\operatorname{E}X_k\right), \tag{4.37}$$

wobei mit E_N beziehungsweise Var_N die Erwartung beziehungsweise die Varianz bezüglich der Verteilung von N bezeichnet sei.

Beweis. Gemäß der Formel für die bedingte Varianz (vgl. Seite 151 in [18]) gilt

$$\operatorname{Var}\left(\sum_{k=0}^{N}X_k\right) = \operatorname{E}_N\left(\operatorname{Var}\left(\sum_{k=0}^{N}X_k|N\right)\right) + \operatorname{Var}_N\left(\operatorname{E}\left(\sum_{k=0}^{N}X_k|N\right)\right). \tag{4.38}$$

Bezeichne mit $S_N((X_k)_k) = \sum_{k=0}^{N}X_k$. Dann gilt für die bedingte Varianz von $S_N((X_k)_k)$ gegeben N vermöge der Unabhängigkeit der beteiligten Zufallsvariablen

$$\begin{aligned}
&\operatorname{Var}\left(S_N((X_k)_k)|N=n\right)\\
&= \operatorname{Var}\left(S_n((X_k)_k)\right)\\
&= S_n((\operatorname{Var}(X_k))_k)\\
&= S_n((\operatorname{Var}(X_k - \operatorname{E}X_k))_k) + S_n((\operatorname{Var}(\operatorname{E}X_k))_k).
\end{aligned} \tag{4.39}$$

Für die bedingte Erwartung von $S_N((X_k)_k)$ gegeben N ergibt sich ferner ebenfalls aufgrund der Unabhängigkeit der beteiligten Zufallsvariablen

$$\begin{aligned}
&\operatorname{E}\left(S_N((X_k)_k)|N=n\right)\\
&= \operatorname{E}\left(S_n((X_k)_k)\right)\\
&= S_n((\operatorname{E}X_k)_k)\\
&= S_n((\operatorname{E}(X_k - \operatorname{E}X_k))_k) + S_n((\operatorname{E}X_k)_k).
\end{aligned} \tag{4.40}$$

Zusammengefasst erhalten wir damit

$$\operatorname{Var}\left(S_N((X_k)_k)\right) = \operatorname{E}_N\left(S_N((\operatorname{Var}X_k)_k)\right) + \operatorname{Var}_N\left(S_N((\operatorname{E}X_k)_k)\right).$$

Die Behauptung ergibt sich unter nochmaliger Anwendung der Formel für die bedingte Varianz für die Varianz von $S_N((X_k - \operatorname{E}X_k)_k)$ beziehungsweise $S_N((\operatorname{E}X_k)_k)$ vermöge der Feststellungen $\operatorname{E}((S_N(X_k - \operatorname{E}X_k)_k)|N=n) = 0$ und $\operatorname{Var}(S_N((\operatorname{E}X_k)_k)|N=n) = 0$. □

4.14 Korollar
Für die Varianz von nZ_n ergibt sich damit

$$\mathrm{Var}(nZ_n) = n^2\,\mathrm{E}_{\tau_n}\left(\sum_{k=0}^{\tau_n}\mathrm{Var}(T_{n-k})\right) + n^2\,\mathrm{Var}_{\tau_n}\left(\sum_{k=0}^{\tau_n}\mathrm{E}T_{n-k}\right). \tag{4.41}$$

Mit Rückblick auf 4.6 können wir ferner feststellen, dass das asymptotische Verhalten von $\mathrm{Var}(nZ_n)$ durch das Verhalten von $\mathrm{Var}(\mathrm{II})$ dominiert wird, wenn sich die Summe der Varianzen der T_i geeignet verhält.

Ferner lässt sich nunmehr eine hinreichende Bedingung für die Erfüllung von (4.32) in 4.11 angeben (vgl. ebenfalls [7]):

4.15 Korollar
Gegeben eine Familie unabhängiger Zufallsvariablen $(T_i)_{i>1}$, mit $\mathrm{E}T_i = i(i-1)/2$ und $\mathrm{Var}(T_i) = (\mathrm{E}T_i)^2$, so gilt die Aussage von 4.10; insbesondere konvergiert dann nZ_n gemäß der Definition in (4.14) in Verteilung gegen eine Verteilung, deren Dichte durch $x \to 8/(2+x)^3$ gegeben ist.

Beweis. Seien T_i, $i \geq 2$ mit $\mathrm{E}T_i = \binom{i}{2}$ und $\mathrm{Var}(T_i) = (\mathrm{E}T_i)^2$. Sei ferner $x_0 \in (0,1)$. Dann können wir ein $m \in \mathbb{N}$ derart finden, dass $n - \lfloor nx_0 \rfloor \geq 2$ für alle $n \geq m$ gilt. Dann erhalten wir

$$\begin{aligned}
&\frac{n^2}{4}\sum_{k=0}^{\lfloor nx_0\rfloor}\mathrm{Var}(T_{n-k}) = \frac{n^2}{4}\sum_{k=n-\lfloor nx_0\rfloor}^{n}\mathrm{Var}(T_k)\\
&= \frac{n^2}{(n-1)^2n^2} + \frac{n^2}{(n-2)^2(n-1)^2} + \cdots + \frac{n^2}{(n-\lfloor nx_0\rfloor-1)^2(n-\lfloor nx_0\rfloor)^2}\\
&\leq \frac{n^2(\lfloor nx_0\rfloor+1)}{(n-\lfloor nx_0\rfloor-1)^4} \leq \frac{n^3+n^2}{(n(1-x_0)-1)^4} \to 0,
\end{aligned}$$

für $n \to \infty$. Damit sind alle Voraussetzungen aus 4.11 erfüllt; inbesondere gilt damit die Konvergenz von nZ_n gegen eine Verteilung, deren Dichte durch $x \to 8/(2+x)^3$ gegeben ist. □

In Bezug auf die Interpretation der Länge eines zufällig gewählten Zweiges im Hinblick auf die Frage nach einem Test auf neutrale Selektion, analog zu [12] notieren wir ebenfalls zunächst, dass man den Koaleszenzprozess einfach um neutrale Mutation unter Annahme des »Infinite Allele«, oder des »Infinite Sites« Modells erweitern kann, indem man

einen unabhängigen Poisson-Prozess mit Rate $\theta/2$ wählt, um damit nachträglich Mutationszeitpunkte zu generieren (vgl. [9] oder [16]). Der Parameter θ bezeichnet dabei die skalierte Mutationsrate.

Damit genügt die Anzahl der Mutationen auf einem Zweig der Länge l einer Poisson-Verteilung mit Rate $l\theta/2$.

Für einen externen Zweig eines Koaleszenzbaumes mit Länge Z_n bezeichnet demgemäß $M(2Z_n)$ die Anzahl der Mutationen, die auf dem Teilbaum eines zufällig gewählten Individuums und seines nähesten Verwandten stattfinden. Als Anwendung von 4.10 ergibt sich:

4.16 Korollar
Für die Anzahl der Mutationen auf einem zufällig gewählten externen Zweig eines Koaleszenzbaumes gilt $M(2nZ_n) \xrightarrow{\mathrm{d}} M(2Z)$, wobei die Verteilung von $M(2Z)$ durch

$$\begin{aligned} \mathrm{P}(M(2Z/n) = k) &= \int \mathrm{P}(M(2x/n) = k)\, \mathrm{dP}^Z(\,\mathrm{d}x) \\ &= \int \frac{x^k \theta^k}{n^k k!} e^{-x\theta/n}\, \mathrm{dP}^Z(\,\mathrm{d}x) \\ &= \frac{(\theta/n)^k}{k!} \int_0^\infty e^{-\theta x/n} \frac{8x^k}{(2+x)^3}\,\mathrm{d}x \end{aligned} \tag{4.42}$$

für $k \in \mathbb{N}_0$ gegeben ist.

A. Appendix

An dieser Stelle seien einige Aussagen zusammengefasst, auf die im bisherigen Verlauf verwiesen wurde, die aber im Textverlauf nicht zweckmäßig untergebracht werden konnten.

A.1 Lemma
Sei $k \in \mathbb{N}$ und seien $a, c \in \mathbb{R}$. Dann gilt

$$\lim_{x\to\infty} a\frac{\ln^k(x)}{x} + c = c. \tag{A.1}$$

Beweis. Da sowohl x als auch $\ln^k(x)$ für $x \to \infty$ bestimmt gegen ∞ divergieren, lässt sich die Regel von de l'Hospital anwenden (siehe zum Beispiel §16, Satz 9 in [11]). Es gilt dann für $k > 1$

$$\lim_{x\to\infty} \frac{\ln^k(x)}{x} = \lim_{x\to\infty} \frac{\frac{\mathrm{d}}{\mathrm{d}x}\ln^k(x)}{\frac{\mathrm{d}}{\mathrm{d}x}x} = \lim_{x\to\infty} \frac{\frac{k\ln^{k-1}(x)}{x}}{1} = \lim_{x\to\infty} \frac{k\ln^{k-1}(x)}{x},$$

und somit durch iteratives Anwenden der Regel von L'Hospital und unter Verwendung von $\frac{\mathrm{d}}{\mathrm{d}x}\ln(x) = 1/x$

$$\lim_{x\to\infty} a\frac{\ln^k(x)}{x} + c = \lim_{x\to\infty} a\frac{k!}{x} + c = c.$$

□

A.2 Lemma
Für $x > 0$ gilt

$$x - \frac{x^2}{2} \leq \ln(1+x) \leq x.$$

Beweis. Sei $0 < x \leq 1$, dann gilt

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{x^n}{n},$$

als Reihenentwicklung des Logarithmus mit Konvergenzradius 1. Weiterhin gilt für jedes $k \in \mathbb{N}$

$$\left|\frac{x^k}{k}\right| > \left|\frac{x^{k+1}}{k+1}\right|$$

und damit für $k \in \mathbb{N}$

$$(-1)^{2k+1}\frac{x^{2k}}{2k} + (-1)^{2k+2}\frac{x^{2k+1}}{2k+1} < 0,$$

beziehungsweise

$$(-1)^{2k+2}\frac{x^{2k+1}}{2k+1} + (-1)^{2k+3}\frac{x^{2k+2}}{2k+2} > 0.$$

Daraus folgt

$$\sum_{n=2}^{\infty}(-1)^{n+1}\frac{x^n}{n} \leq 0 \text{ bzw. } \sum_{n=3}^{\infty}(-1)^{n+1}\frac{x^n}{n} \geq 0,$$

und damit wiederum

$$\begin{aligned} x - \frac{x^2}{2} &\leq x - \frac{x^2}{2} + \sum_{n=3}^{\infty}(-1)^n + 1\frac{x^n}{n} \\ &= \ln(1+x) \\ &= x + \sum_{n=2}^{\infty}(-1)^{n+1}\frac{x^n}{n} \leq x. \end{aligned}$$

Sei nun $x > 1$. Beobachte, dass $f(x) = -\frac{x^2}{2} + x$ ein globales Maximum bei $x_0 = 1$ besitzt und ferner $f(1) = \frac{1}{2} \leq \ln(2)$ gilt. Die zweite Ungleichung gilt vermöge der Beobachtung $\ln(1+x) \leq x \Leftrightarrow 1 + x \leq e^x$ und $2 \leq e$.

□

A.3 Lemma
Sei $x \in \mathbb{R}$, dann gilt

$$\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = e^x \tag{A.2}$$

Beweis. Sei $x \in \mathbb{R}$ gegeben, ferner sei $x \neq 0$. Setze $f_x(n) \overset{\text{def}}{=} (1 + \frac{x}{n})$ und betrachte

$$\lim_{n\to\infty}\ln(f_x(n)^n) = \lim_{n\to\infty} n\ln(f_x(n)) = \lim_{n\to\infty}\frac{\ln(f_x(n))}{n^{-1}}.$$

Damit sind die Voraussetzungen für den Satz von de l'Hospital (§16, Satz 9 in [11]) gegeben und es gilt $\lim_{n\to\infty}\ln((1+x/n)^n) = x$. □

A.4 Lemma
Für die Summe der natürlichen Zahlen von 1 bis n, ihrer Quadrate, beziehungsweise ihrer dritten Potenzen gelten

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \tag{A.3}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}, \tag{A.4}$$

$$\sum_{k=1}^{n} k^3 = \frac{n^4 + 2n^3 + n^2}{4}. \tag{A.5}$$

A.5 Lemma
Es seien $(X_k)_{1\leq k\leq n}$ unabhängige, jeweils $\mathrm{Exp}(\lambda_k)$-verteilte Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, \mathrm{P})$, $\lambda_k > 0$, $1 \leq k \leq n$. Dann ist $X \stackrel{\mathrm{def}}{=} \min\{X_k : 1 \leq k \leq n\}$ ebenfalls exponentialverteilt mit Parameter $\lambda = \sum_{k=1}^{n} \lambda_k$. Die Wahrscheinlichkeit, dass X_j, $1 \leq j \leq n$ das Minimum annimmt, ergibt sich als

$$\mathrm{P}(X_j = \min\{X_k : 1 \leq k \leq n\}) = \frac{\lambda_j}{\lambda_1 + \cdots + \lambda_n}.$$

Beweis. Wähle hierzu $x > 0$ beliebig und betrachte

$$\begin{aligned}
\mathrm{P}(X > x) &= \mathrm{P}(X_1 > x, X_2 > x, \ldots, X_n > x) \\
&= \prod_{k=1}^{n} \mathrm{P}(X_k > x) \\
&= \prod_{k=1}^{n} \exp(-x\lambda_k) = \exp\left(-x \sum_{k=1}^{n} \lambda_k\right).
\end{aligned}$$

Damit entspricht die Verteilungsfunktion von X der einer Exponentialverteilung mit Parameter λ. Für den zweiten Teil notieren wir

$$\begin{aligned}
\mathrm{P}(X_j = \min\{X_k : 1 \leq k \leq n\}) &= \mathrm{P}(X_j < X_k, j \neq k, 1 \leq k \leq n) \\
&= \mathrm{E}_{X_j}(\mathbb{1}_{\{X_j < X_k : j \neq k, 1 \leq k \leq n\}} | X_j) \\
&= \int_0^\infty \lambda_j e^{-\lambda_j t} \prod_{\substack{j \neq k \\ 1 \leq k \leq n}} \mathrm{P}(X_j > t)\, \mathrm{d}t \\
&= \int_0^\infty \lambda_j e^{-(\lambda_1 + \cdots + \lambda_n)t}\, \mathrm{d}t \\
&= \frac{\lambda_j}{\lambda_1 + \cdots + \lambda_n}.
\end{aligned}$$

□

A.6 Proposition
Sei $(\Omega, \mathfrak{A}, \mathrm{P})$ ein Wahrscheinlichkeitsraum, und seien X_1, X_2: $(\Omega, \mathfrak{A}, \mathrm{P}) \to (\mathbb{R}, \mathfrak{B})$ unabhängige, exponentialverteilte Zufallsvariablen mit Parameter $\lambda_i > 0$, das heißt $X_i \sim \mathrm{Exp}(\lambda_i)$, sowie $B : (\Omega, \mathfrak{A}, \mathrm{P}) \to (\{0,1\}, \mathscr{P}(\{0,1\})$ eine von X_1, X_2 unabhängige, bernoulliverteilte Zufallsvariable mit Parameter $p \in (0,1)$, das heißt $B \sim \mathrm{B}(1,p)$. Unter der Voraussetzung $\lambda_1 > \lambda_2$ und $p = \frac{\lambda_1 - \lambda_2}{\lambda_1}$, gilt

$$X_1 + BX_2 \stackrel{\sim}{=} \mathrm{Exp}(\lambda_2),$$

das heißt, $X_1 + BX_2$ genügt einer Exponentialverteilung mit Parameter λ_2.

Beweis. Unter der Berücksichtigung von

$$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,\infty)}(x) \qquad F_X(x) = 1 - e^{-\lambda x} \mathbb{1}_{[0,\infty)}(x),$$

wobei f_X die Dichte- und F_X die Verteilungsfunktion für eine Zufallsvariable $X \sim \mathrm{Exp}(\lambda)$ bezeichne, gilt für die Verteilungsfunktion von $X_1 + BX_2$

$$\begin{aligned}
&\mathrm{P}(X_1 + BX_2 \leq t) \\
&= \mathrm{P}(N=0)\,\mathrm{P}(X_1 + BX_2 \leq t | B=0) + \mathrm{P}(N=1)\,\mathrm{P}(X_1 + BX_2 \leq t | B=1) \\
&= (1-p)\left(1 - e^{-\lambda_1 t}\right) + p\,(\,\mathrm{E}_{X_1}(\,\mathrm{P}(X_1 + BX_2 \leq t | B=1, X_1)) \\
&= (1-p)\left(1 - e^{-\lambda_1 t}\right) + p \int_0^\infty \left(1 - e^{-\lambda_2 (t-s)}\right) \mathbb{1}_{\{t-s \geq 0\}} \lambda_1 e^{-\lambda_1 s}\,\mathrm{d}s \\
&= (1-p)\left(1 - e^{-\lambda_1 t}\right) + p \left[1 - e^{-\lambda t} - \frac{\lambda_1}{\lambda_2 - \lambda_1}\left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right)\right] \\
&= (1-p)\left(1 - e^{-\lambda_1 t}\right) + p \frac{\lambda_1}{\lambda_1 - \lambda_2}\left(e^{-\lambda_1 t} - e^{-\lambda_2 t}\right).
\end{aligned}$$

Mit Erinnerung an die Voraussetzung $p = \frac{\lambda_1 - \lambda_2}{\lambda_1}$ folgt somit die Behauptung. □

Literaturverzeichnis

[1] Alsmeyer, G.: *Stochastische Prozesse, Teil 2*. Unveröffentlichtes Skript.

[2] Alsmeyer, G.: *Wahrscheinlichkeitstheorie*. Skripten zur Mathematischen Statistik, Nr. 30, 2005. Institut für Mathematische Statistik, Fachbereich Mathematik der Westfälischen Wilhelms-Universität Münster.

[3] Asmussen, S.: *Applied Probability and Queues*. Springer, 2. Aufl., 2003.

[4] Berestycki, N.: *Recent Progress in Coalescent Theory*. Ensaios Matemáticos, 16:1–193, 2009.

[5] Bertoin, J.: *Random Fragmentation and Coagulation Processes*, Bd. 102 d. Reihe *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 13. Aufl., 2006.

[6] Blum, M. G. B. und O. François: *External branch length and minimal clade size under the neutral coalescent*. Advances in Applied Probability, 37(3):647–662, 2005.

[7] Caliebe, A., R. Neininger, M. Krawczack und U. Rösler: *On the length distribution of external branches in coalescence trees: Genetic diversity within species*. Theoretical Population Biology, 72:245–252, 2007.

[8] Chung, K. L. und J. B. Walsh: *Markov Processes, Brownian Motion, and Time Symmetry*. Springer, 2. Aufl., 2005.

[9] Durrett, R.: *Probability Models for DNA Sequence Evolution*. Springer, 2. Aufl., 2008.

[10] Evans, S. N. und J. Pitman: *Construction of Markovian coalescents*. Ann. Inst. Henri Poincaré, 34(3):339–383, 1998.

[11] Forster, O.: *Analysis 1*. Vieweg & Sohn Verlag, 7. Aufl., Juni 2004.

[12] Fu, Y.-X. und W.-H. Li: *Statistical Tests of Neutrality of Mutations*. Genetics, 133(3):693–709, 1993.

[13] Janson, S. und G. Kersting: *The external lengths in Kingman's coalescent*. arXiv:1004.5011v2 [math.PR], 2011.

[14] Kingman, J.: *The Coalescent*. Stochastic Processes and their Applications, 13:235–248, 1982.

[15] Kingman, J.: *On the Genealogy of Large Populations*. Journal of Applied Probability, 19:27–43, 1982.

[16] Nordborg, M.: *Coalescent Theory*. Techn. Ber., Department of Genetics, Lund University, 2000.

[17] Rauch, E. M. und Y. Bar-Yam: *Theory predicts the uneven distribution of genetic diversity within species*. Nature, 431:449–452, 2004.

[18] Rice, J. A.: *Mathematical Statistics And Data Analysis*. Duxbury Press, 3. Aufl., 2006.

[19] Tavaré, S.: *Ancestral Inference in Population Genetics*. Springer, 2004.

Ich versichere, dass ich die vorliegende Arbeit selbständig angefertigt und keine anderen als die im Literaturverzeichnis aufgeführten Quellen und Hilfsmittel verwendet habe.

Ahaus, den 21. September 2012

Thomas Uckelmann