Westfälische Wilhelms-Universität Münster

Institut für Mathematische Statistik

Top-to-Random-Shuffles

Diplomarbeit

vorgelegt von

Christian Palmes

Thema gestellt von

Prof. Dr. Gerold Alsmeyer

15. März 2010

gewidmet
meinen Eltern

Inhaltsverzeichnis

Einleitung

Diese Arbeit beschäftigt sich mit *Kartenmischsystemen* und ist in das Gebiet der *endlichen Markovketten* einzuordnen. Genauer wird der *Top-to-Random-Shuffle* analysiert. Bei diesem wird in jedem Mischschritt die oberste Karte des zu durchmischenden Kartendecks abgehoben und an eine zufällige Position wieder in das verbleibende Kartendeck zurückgesteckt. Die zentrale Frage lautet, wie häufig obiger Mischvorgang wiederholt werden muss, bis ein anfangs vorsortiertes Kartendeck in nahezu völlige Unordnung gebracht wird. Interessant ist, dass hierbei ein *Cutoff-Effekt* besteht. Das bedeutet, dass nach einer gewissen Anzahl von Mischvorgängen das Kartendeck quasi sprunghaft durchmischt ist, d.h. dass weniger Mischvorgänge nicht ausreichend sind und mehr Mischvorgänge unnötig sind. Die Konvergenz gegen die Gleichverteilung auf der Menge aller Kartenanordnungen erfolgt also nicht kontinuierlich, sondern der Übergang zwischen „undurchmischt" und „durchmischt" ist deutlich erkennbar. Folgende Abbildung veranschaulicht das oben Gesagte, wobei die Abzisse die Anzahl der ausgeführten *Top-to-Random-Shuffles* angibt und die Ordinate den *Variationsabstand* zur Gleichverteilung beschreibt.

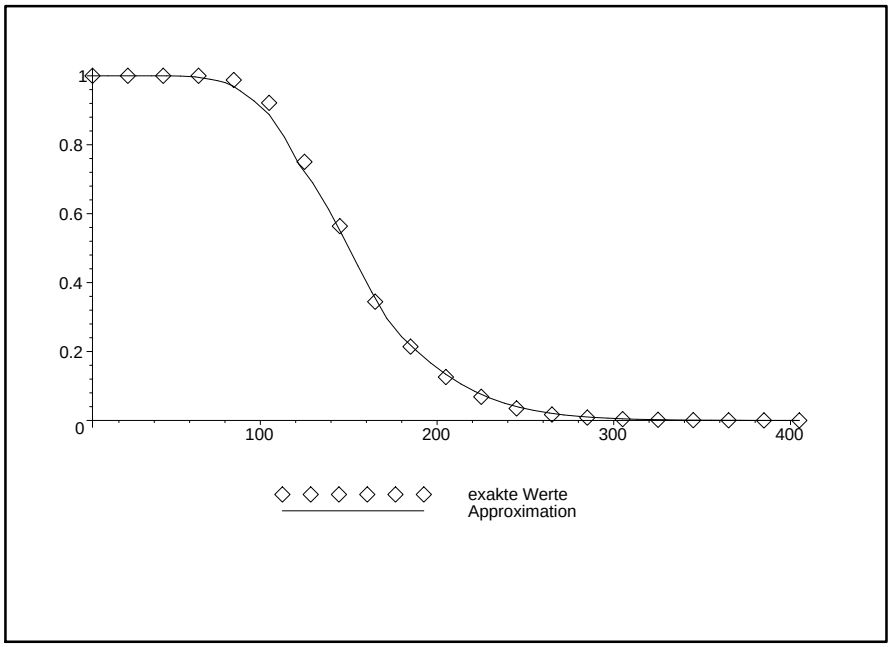

Abbildung 1: Cutoff-Effekt beim *Top-To-Random-Shuffle* mit 52 Karten

Obige Abbildung wurde mit dem Computeralgebrasystem *Maple* für ein Kartendeck von 52 Karten errechnet. Die durchgezogene Kurve ist eine Approximation der exakten Werte und entstammt der in Abschnitt 2.3 entwickelten Theorie.

In Kapitel 1 werden grundlegende Erkenntnisse der Theorie endlicher Markovketten wiederholt, wobei auch auf einen Spezialfall endlicher Markovketten eingegangen wird, der alle hier beschriebenen Kartenmischsysteme modelliert. Da dieses Kapitel eher wiederholenden Charakter hat, werden wir an mehreren Stellen auf einen Beweis verzichten und auf entsprechende Literatur verweisen. Dieses Kapitel basiert im Wesentlichen auf Alsmeyer [5] und Aldous, Diaconis [2].

Kapitel 2 beschäftigt sich speziell mit *Top-to-Random-Shuffles*, wobei hier verschiedene Varianten des *Top-to-Random-Shuffles* untersucht werden und nicht nur die anfänglich beschriebene, bei der in jedem Schritt immer *genau eine* Karte abgehoben wird. Für alle hier untersuchten Varianten steht der Nachweis eines *Cutoff-Effekts* im Zentrum der Betrachtung, weshalb wir an den zugehörigen Stellen auch längere Rechnungen explizit ausführen werden. Dieses Kapitel basiert hauptsächlich auf Diaconis, Fill, Pitman [8] und Aldous, Diaconis [1].

Danksagungen

Der mit Abstand größte Dank gebührt meinen Eltern für die finanzielle und vor allem ideelle Unterstützung.

Herrn Prof. Dr. G. Alsmeyer danke ich für die Vergabe und Betreuung dieses interessanten Themas. Frau Dipl.-Math. Andrea Winkler, Herr Dipl.-Math. Dominik Menning und Herr Alexander Kupsch haben mir ferner durch Verifizierung einiger Beweise und mit vielen hilfreichen Ratschlägen ebenfalls sehr geholfen, wofür ich mich an dieser Stelle bedanken möchte. Schließlich bedanke ich mich bei Polly für die Unterstützung in den letzten Monaten und wünsche ihr alles Gute für ihr anstehendes Jura Examen.

Kapitel 1

Grundlegende Betrachtungen

In diesem Kapitel soll ein Einblick in die allgemeine Theorie, die allen Kartenmischsystemen zugrunde liegt, gegeben werden. Wie bereits in der Einleitung erwähnt, wird mehrfach anstelle eines Beweises die entsprechende Literatur zitiert werden.

Abschnitt 1.1 formalisiert mit Definition 1.1.1, was mit einem Kartenmischsystem gemeint ist, und zeigt mittels Satz 1.1.2, dass wir uns in dem Gebiet der endlichen, homogenen Markov-Ketten befinden.

Abschnitt 1.2 wiederholt einige Fakten diskreter, homogener Markov-Ketten. Es werden allgemeiner homogene Markov-Ketten auf einem abzählbar unendlichen Zustandsraum behandelt, da dies für unsere Zwecke nur einen geringen Mehraufwand darstellt und wir so auf die zusätzlichen Eigenschaften im Spezialfall endlicher Markov-Ketten expliziter aufmerksam machen können. Die beiden zentralen Aussagen in diesem Abschnitt beinhalten Theorem 1.2.7 und Satz 1.2.13. Theorem 1.2.7 liefert eine wichtige Aussage über die Konvergenz und Konvergenzrate spezieller Markov-Ketten, und Satz 1.2.13 ist die Kernaussage des Prinzips der *stark stationären Zeiten*, das uns noch an einigen Stellen in Kapitel 2 als nützliches technisches Hilfsmittel dienen wird.

Abschnitt 1.3 beinhaltet einige Informationen bzgl. endlicher Markov-Ketten, deren Zustandsraum eine endliche Gruppe G ist und deren Übergangsmatrix die spezielle Form (1.12) besitzt. Dieser Spezialfall ist von Bedeutung, da wir uns für $G = \mathfrak{S}_n$ (symmetrische Gruppe) in der Situation von Kartenmischsystemen befinden. Eine interessante Aussage bildet hier Satz 1.3.7, der eine Relation zwischen Separations- und Variationsabstand enthält. Dieser wird hier vollständig bewiesen.

1.1 Kartenmischen im Allgemeinen

Es sei ein Kartendeck mit n Karten gegeben (n Kartendeck). Da alle Karten als unterscheidbar angenommen werden, resultieren hieraus $n!$ verschiedene Anordnungsmöglichkeiten. Wir gehen davon aus, dass dieses Kartendeck zu Beginn perfekt sortiert ist (z.B. ein neu gekauftes, bei dem gerade die Schutzfolie entfernt wurde) und eine Person dieses Kartendeck in endlich vielen Schritten in möglichst große Unordnung bringen möchte. Jeder Schritt entspricht hierbei einem Mischvorgang. Bei jedem vernünftigen Mischvorgang werden zwei Aussagen immer erfüllt sein.

(i) Jeder Mischvorgang ist unabhängig von allen vorherigen.

(ii) Vor der Ausführung eines Mischvorgangs ist die aktuelle Anordnung des Kartendecks der Person, die das Kartendeck durchmischt, unbekannt.

Identifizieren wir jede Karte mit einer natürlichen Zahl aus $\{1, 2, \ldots, n\}$, so lässt sich die Anordnung des Kartendecks eindeutig durch ein Element π der *symmetrischen Gruppe* $\mathfrak{S}_n$ angeben. Genauer bezeichne $\pi(i)$ die Karte an i-ter Position, wobei zur besseren Vorstellung die *erste* Position der *obersten* Karte des Kartendecks entspreche und dann naheliegend durchnummeriert wird.

Sei z.B. $n = 2m$, $m \in \mathbb{N}$ und jeder Mischschritt teile das Kartendeck *zuerst* in *exakt* zwei gleich große Kartenstapel, d.h. es werden genau m Karten abgehoben, und füge diese *anschließend* quasi fächerförmig zusammen. Genauer sollen die Karten beider Kartenstapel abwechselnd aufeinandergelegt werden, so dass wir einen „perfekten" *Riffle Shuffle* (siehe Lemma 2.5.11) durchgeführt haben. Formal lässt sich dies durch eine einfache Abbildung beschreiben: $\pi \mapsto \pi \circ \sigma$, wobei π die Anordnung des Decks vor dem Mischvorgang und $\pi \circ \sigma$ die Anordnung nach dem obigen Mischvorgang bezeichnet. Offenbar gilt hier

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & \ldots & n-1 & n \\ 1 & m+1 & 2 & m+2 & \ldots & m & 2m \end{pmatrix} \in \mathfrak{S}_n,$$

und "$\circ$„ bezeichne die gewöhnliche Komposition von Abbildungen in $\mathfrak{S}_n$. Nach $k \in \mathbb{N}$ Schritten haben wir folglich $\pi \circ \sigma^k$. Insbesondere ergibt sich nach $k = \operatorname{ord} \sigma$ Mischschritten wieder das Ausgangsdeck π, wobei $\operatorname{ord} \sigma$ die Anzahl der Elemente der durch σ erzeugten zyklischen Untergruppe $\langle \sigma \rangle$ bezeichne. Für $n \geq 3$ kann ferner unmöglich jede Kartenkonstellation in $\mathfrak{S}_n$ erreicht werden, da $\langle \sigma \rangle \subsetneq \mathfrak{S}_n$, denn $\langle \sigma \rangle$ ist kommutativ, $\mathfrak{S}_n$ jedoch nicht. Mit z.B. 10 Karten erreichen wir das Ausgangskartendeck schon nach 6 Schritten, wir haben also nur 6 von den $10! = 3.628.800$ möglichen Kartenanordnungen erreicht. Natürlich gilt mit $n \geq 3$ für *jedes* $\tau \in \mathfrak{S}_n$ die echte Mengeninklusion $\langle \tau \rangle \subsetneq \mathfrak{S}_n$, so dass wir in unserem Bestreben nach einer guten Durchmischung unterschiedliche $\tau \in \mathfrak{S}_n$ in jedem Schritt zulassen werden. Welches $\tau \in \mathfrak{S}_n$ in dem jeweiligen Schritt gewählt werden soll, liefert der Ausgang eines Zufallsexperiments auf $\mathfrak{S}_n$, genauer die Realisierung einer Zufallsvariablen $Y_i : (\Omega, \mathfrak{A}, P) \to (\mathfrak{S}_n, \mathcal{P}(\mathfrak{S}_n))$, wobei $(\Omega, \mathfrak{A}, P)$ ein geeigneter Wahrscheinlichkeitsraum ist und mit $\mathcal{P}(\mathfrak{S}_n)$ die Potenzmenge von $\mathfrak{S}_n$ gemeint ist. $i = 1, 2, \ldots$ bezeichne hierbei den i-ten Mischschritt. Offensichtlich ist in diesem Modellansatz obige Bedingung (ii) erfüllt. $\left(P^{Y_i}\right)_{i=1,2,\ldots}$ charakterisiert das Mischverfahren, welches im Schritt $i = 1, 2, \ldots$ zur Anwendung kommt. Es ist durchaus denkbar, dass dieses nicht immer dasselbe ist, z.B. wenn zum Schluss das Kartendeck noch abgehoben wird. Dennoch werden wir in dieser Arbeit immer $P^{Y_i} = P^{Y_1}$, $i = 1, 2, \ldots$ annehmen und dem Mischverfahren anhand von P^{Y_1} einen Namen zuteilen. Falls P^{Y_1} z.B. von der Form (2.3) ist, wird das Mischsystem *Top-m-to-Random-Shuffle* genannt. Sei

$$X_k \stackrel{\text{def}}{=} Y_1 \circ \ldots \circ Y_k, \quad k = 1, 2, \ldots \text{ und } X_0 \stackrel{\text{def}}{=} \mathrm{id}.$$

X_k, $k \geq 0$ beschreibt folglich die Anordnung eines Kartendecks nach dem k-ten Mischschritt, falls ein vorsortiertes ($\mathrm{id} \in \mathfrak{S}_n$) Kartendeck vor dem ersten Schritt zugrunde liegt. Zusammen mit (i) gelangen wir nun endgültig zur mathematischen Struktur oben beschriebener Kartenmischverfahren:

Definition 1.1.1. Mit einem *Mischverfahren* oder *Mischsystem* ist eine Verteilung auf $\mathfrak{S}_n$ gemeint.

Satz 1.1.2. *$(X_k)_{k=0,1,\dots}$ ist eine diskrete, homogene Markov-Kette auf dem Zustandsraum $\mathfrak{S}_\mathrm{n}$ mit der Startverteilung δ_{id} und der Übergangsmatrix*

$$M(\sigma,\tau) = P^{Y_1}(\sigma^{-1}\circ\tau), \quad \sigma,\tau\in\mathfrak{S}_\mathrm{n},\ M\in\mathbb{R}^{\mathfrak{S}_\mathrm{n}\times\mathfrak{S}_\mathrm{n}}.$$

Beweis. Es gilt für alle $k\geq 0$ und $\pi_0,\dots,\pi_{k+1}\in\mathfrak{S}_\mathrm{n}$

$$\begin{aligned} & P(X_{k+1}=\pi_{k+1}|X_k=\pi_k,\dots,X_0=\pi_0)\\ = \ & P(X_k\circ Y_{k+1}=\pi_{k+1}|X_k=\pi_k,\dots,X_0=\pi_0)\\ = \ & P(Y_{k+1}=\pi_k^{-1}\circ\pi_{k+1}|X_k=\pi_k,\dots,X_0=\pi_0)\\ = \ & P^{Y_1}(\pi_k^{-1}\circ\pi_{k+1}),\end{aligned}$$

wobei die letzte Gleichung wegen (i) und $\sigma(X_0,\dots,X_k)\subseteq\sigma(Y_1,\dots,Y_k)$ gilt. Sei Z eine Zufallsvariable. $\sigma(Z)$ bezeichnet dann wie gewohnt die kleinste σ-Algebra bzgl. derer Z messbar ist. □

Diese Arbeit ist daher in die Theorie der diskreten (genauer endlichen), homogenen Markov-Ketten einzuordnen. Die Existenz solcher Prozesse folgt direkt aus dem Satz von *Ionescu Tulcea* (siehe Alsmeyer [3, S.311]). Im nächsten Abschnitt werden einige allgemeingültige Fakten über diskrete, homogene Markov-Ketten wiederholt, wovon allerdings die meisten nicht bewiesen werden. Der Leser sei hierzu auf [5] verwiesen. In dem übernächsten Abschnitt 1.3 werden wir dann wieder etwas konkreter und betrachten Random Walks auf endlichen Gruppen, wie z.B. $(X_k)_{k=0,1,\dots}$ in Satz 1.1.2.

1.2 Diskrete, homogene Markov-Ketten

Sei $\mathcal{S}$ eine nichtleere, abzählbare Menge und $X\overset{\text{def}}{=}(X_k)_{k=0,1,\dots}$ eine Folge von Zufallsvariablen $X_k:(\Omega,\mathfrak{A},P)\to(S,\mathcal{P}(\mathcal{S}))$.

Definition 1.2.1. X heißt *diskrete Markov-Kette* (DMK), falls sie die *Markoveigenschaft*

$$P^{X_{k+1}|X_0,\dots,X_k} = P^{X_{k+1}|X_k} \quad P\text{-f.s.}, \quad k\in\mathbb{N}_0 \tag{1.1}$$

erfüllt.

Bemerkung 1.2.2.

$$P^{X_{k+1}|X_k=\cdot}:\mathcal{S}\times\mathcal{P}(\mathcal{S})\to[0,1], \quad k=0,1,2,\dots$$

ist ein stochastischer Kern von $(\mathcal{S},\mathcal{P}(\mathcal{S}))$ nach $(\mathcal{S},\mathcal{P}(\mathcal{S}))$. In anderer Notation besagt (1.1)

$$\begin{aligned} & P(X_{k+1}=x_{k+1}|X_0=x_0\dots,X_k=x_k), \quad x_0,\dots,x_{k+1}\in\mathcal{S}\\ = \ & P(X_{k+1}=x_{k+1}|X_k=x_k).\end{aligned}$$

Falls es einen stochastischen Kern $\mathbb{P}$ von $(\mathcal{S},\mathcal{P}(\mathcal{S}))$ nach $(\mathcal{S},\mathcal{P}(\mathcal{S}))$ mit

$$P^{X_{k+1}|X_k} = \mathbb{P}(X_k,\cdot) \quad P\text{-f.s.}, \quad k=0,1,\dots$$

gibt, so wird die hierzu zugehörige Markov-Kette *homogen* genannt. Durch

$$A(x_0,x_1)\overset{\text{def}}{=}P(X_1=x_1|X_0=x_0), \quad A\in\mathbb{R}^{\mathcal{S}\times\mathcal{S}}, \quad \lambda\overset{\text{def}}{=}P^{X_0}$$

ist die Markov-Kette in diesem Fall offenbar eindeutig charakterisiert. A wird weiter zu einer gewöhnlichen Matrix, der *Übergangsmatrix*, falls $\mathcal{S}$ endlich ist. Wir werden im Folgenden mit einer Markov-Kette immer eine *diskrete, homogene Markov-Kette* meinen und hierfür wieder abkürzend DMK schreiben. Falls P ein Wahrscheinlichkeitsmaß beschreibt, so dass X eine DMK auf dem zugrunde liegenden Wahrscheinlichkeitsraum mit Startverteilung λ ist, d.h. $P^{X_0} = \lambda$ gilt, so werden wir auch P_λ anstelle von P schreiben. Insbesondere verwenden wir die gebräuchliche Abkürzung $P_i \stackrel{\text{def}}{=} P_{\delta_i}$, $i \in \mathcal{S}$. Genauer existiert als Folgerung aus dem Satz von *Ionescu Tulcea* zu jedem stochastischen Kern $\mathbb{P}$ ein sogenanntes Standardmodell. Dieses besteht aus einem Messraum $(\Omega, \mathfrak{A})$, einer Familie $(P_\lambda)_{\lambda \in W(\mathcal{S})}$ von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathfrak{A})$ und Zufallsvariablen $(X_k)_{k \geq 0}$, wobei $W(\mathcal{S})$ die Menge aller Verteilungen auf $(\mathcal{S}, \mathcal{P}(\mathcal{S}))$ bezeichne. In diesem Modell ist $(X_k)_{k \geq 0}$ unter jedem P_λ eine DMK mit der Startverteilung P_λ und dem Übergangskern $\mathbb{P}$. Wir haben es somit mit nur einem Prozess zu tun, wobei sich verschiedene Anfangsverteilungen durch Zugrundelegung verschiedener P_λ ergeben. Es muss daher bei einem Wechsel der Anfangsverteilung nicht der Prozess gewechselt werden, was einen technischen Vorteil bildet. Da $\mathbb{P}$ fixiert ist, entspricht die Untersuchung einer DMK innerhalb eines Standardmodells letztendlich einer Analyse von $\mathbb{P}$.

Wir fahren mit weiteren Definitionen fort.

Definition 1.2.3. Gegeben sei eine DMK X in einem Standardmodell.

(i) X heißt *irreduzibel*, falls für alle $i, j \in \mathcal{S}$ ein $k \in \mathbb{N}_0$ existiert, so dass $P_i(X_k = j) > 0$ gilt.

(ii) X heißt *rekurrent*, falls sie irreduzibel ist und für ein $i \in \mathcal{S}$ (und damit für alle $i \in \mathcal{S}$, siehe [5]) $P_i(X_k = i, \text{ unendlich oft}) = 1$ gilt.

(iii) X heißt *aperiodisch*, falls sie irreduzibel ist und für ein $i \in \mathcal{S}$ (und damit für alle $i \in \mathcal{S}$, siehe [5]) $\text{ggt}\{k \geq 1 : P_i(X_k = i) > 0\} = 1$ gilt.

Irreduzibilität bedeutet, dass jeder Zustand j von jedem Zustand i aus in endlich vielen Schritten mit positiver Wahrscheinlichkeit erreicht werden kann. Wenn *zusätzlich* jeder Zustand i P-f.s. nach endlich vielen Schritten wieder erreicht wird, falls wir zu Beginn in i gestartet haben, dann nennen wir diese Markov-Kette rekurrent. Es ist intuitiv klar, dass jede irreduzible DMK im Falle eines endlichen Zustandraumes auch rekurrent ist. Ein formaler Beweis befindet sich in [5]. Es kann gezeigt werden, dass, falls eine irreduzible DMK X *nicht* aperiodisch ist, ein $d \in \mathbb{N}$, $d \geq 2$ existiert, so dass

$$\text{ggT}\{k \geq 1 : P_i(X_k = i) > 0\} = d$$

für alle $i \in \mathcal{S}$ (d-Periodizität ist eine Solidaritätseigenschaft, siehe [5]) gilt. In diesem Fall existiert dann eine Zerlegung $\mathcal{S} = \sum_{r=0}^{d-1} \mathcal{S}_r$ mit $P_i(X_1 \in \mathcal{S}_{r+1}) = 1$ für alle $i \in \mathcal{S}_r$, $0 \leq r < d$ und $\mathcal{S}_d \stackrel{\text{def}}{=} \mathcal{S}_0$. Es handelt sich m.a.W. um eine disjunkte Zerlegung des Zustandraums $\mathcal{S}$, die von dem stochastischen Prozess X zyklisch durchlaufen wird.

Wir sind im Folgenden an dem Konvergenzverhalten von X interessiert, wobei die Art der Konvergenz noch näher spezifiziert werden muss. Hierzu wird zunächst eine Metrik auf der Menge der Verteilungen auf $\mathcal{S}$ definiert:

Definition 1.2.4. Seien Q_1 und Q_2 zwei Verteilungen auf $\mathcal{S}$. Dann heißt

$$\|Q_1 - Q_2\| \stackrel{\text{def}}{=} \sup_{A \subseteq \mathcal{S}} |Q_1(A) - Q_2(A)| \tag{1.2}$$

der *Variationsabstand* zwischen Q_1 und Q_2.

Satz 1.2.5.

$$d(Q_1, Q_2) \stackrel{\text{def}}{=} \|Q_1 - Q_2\| \tag{1.3}$$

definiert eine Metrik, und es gilt

$$\|Q_1 - Q_2\| = \frac{1}{2} \sum_{s \in \mathcal{S}} |Q_1(s) - Q_2(s)|. \tag{1.4}$$

Beweis. Wir weisen für d nur die Dreiecksungleichung nach. Für drei Verteilungen Q_1, Q_2, Q_3 auf $\mathcal{S}$ gilt

$$\begin{aligned} d(Q_1, Q_2) + d(Q_2, Q_3) &= \sup_{A \subseteq \mathcal{S}} |Q_1(A) - Q_2(A)| + \sup_{B \subseteq \mathcal{S}} |Q_2(B) - Q_3(B)| \\ &\geq \sup_{A \subseteq \mathcal{S}} (|Q_1(A) - Q_2(A)| + |Q_2(A) - Q_3(A)|) \\ &\geq d(Q_1, Q_3). \end{aligned}$$

Für (1.4) betrachte man die Menge $A_0 \stackrel{\text{def}}{=} \{s \in \mathcal{S} : Q_1(s) \geq Q_2(s)\}$. Es gilt dann

$$\begin{aligned} 2\|Q_1 - Q_2\| &= (Q_1(A_0) - Q_2(A_0)) + (Q_2(A_0^c) - Q_1(A_0^c)) \\ &= \sum_{s \in \mathcal{S}} |Q_1(s) - Q_2(s)|, \end{aligned}$$

da das Supremum in (1.3) offenbar in A_0 und A_0^c angenommen wird. □

Obige Metrik ist die Formalisierung eines Abstandbegriffs zwischen zwei Wahrscheinlichkeitsverteilungen. Diese reagiert, anschaulich gesprochen, offenbar sehr sensibel auf Unterschiede zwischen den betrachteten Verteilungen, da das Supremum, das hier ein Maximum ist, gezielt die größte Diskrepanz der beiden Verteilungen hervorhebt.

Wir sind nun an der Frage nach der Existenz einer Verteilung ξ^* interessiert, so dass

$$\|P_\lambda^{X_k} - \xi^*\| \to 0, \quad k \to \infty \tag{1.5}$$

für alle Startverteilungen λ auf $\mathcal{S}$ gelten soll. Es lässt sich zeigen, dass ξ^* in diesem Fall eine *stationäre Verteilung* sein muss, d.h. es gilt $P_{\xi^*}^{X_1} = \xi^*$. Da (1.5) für beliebiges λ gelten soll, ist ξ^* ferner eindeutig bestimmt. Bevor wir das in diesem Kontext entscheidende Theorem niederschreiben, muss noch eine kleine Erweiterung von Definition 1.2.3 (ii) gegeben werden. Sei hierzu

$$\tau(i) \stackrel{\text{def}}{=} \inf\{k \geq 1 : X_k = i\}, \quad \inf \emptyset \stackrel{\text{def}}{=} \infty.$$

X ist genau dann rekurrent, falls X irreduzibel ist und $P_i(\tau(i) < \infty) = 1$, $i \in \mathcal{S}$ gilt, was man sich mit einer kleinen Überlegung oder durch Blick in [5] klarmacht. Wir nennen X *positiv rekurrent*, falls X rekurrent ist und zusätzlich $E_i\tau(i) < \infty$ für ein (und damit für alle $i \in \mathcal{S}$, siehe [5]) gilt. Falls $\mathcal{S}$ endlich ist, so sind die Begriffe Rekurrenz und positive Rekurrenz äquivalent, siehe [5]. Ein wichtiges Theorem in diesem Zusammenhang ist der sogenannte *Ergodensatz für aperiodische, positiv rekurrente* DMK:

Theorem 1.2.6. *Sei X eine aperiodische, positiv rekurrente* DMK. *Dann gilt für eine eindeutig bestimmte Verteilung ξ^**

$$\|P_\lambda^{X_k} - \xi^*\| \to 0, \quad k \to \infty \tag{1.6}$$

für jede Anfangsverteilung λ.

Für einen Beweis siehe z.B. [5, S.76ff]. Falls X irreduzibel, aber *nicht* aperiodisch ist, so folgt aufgrund des zyklischen Verhaltens der DMK (siehe Prosa unter Definition 1.2.3), dass X unmöglich im Sinne von (1.6) konvergieren kann. Falls ferner X nicht irreduzibel ist, so müssten wir (1.6) in Bezug auf den *beliebigen* Anfangsverteilungen λ einschränken und könnten auch nicht mehr die Eindeutigkeit von ξ^* garantieren. X muss ferner positiv rekurrent sein, da nur so eine stationäre Verteilung ξ^* existiert (siehe [5]). Theorem 1.2.6 besagt quasi, dass, wenn alle zu einer potentiellen Konvergenz notwendigen Bedingungen erfüllt sind, auch tatsächlich eine Konvergenz vorliegt. Allerdings ist die Frage nach der Konvergenzrate, also der Geschwindigkeit der Konvergenz, durch Theorem 1.2.6 noch nicht beantwortet. Bevor wir hierauf eingehen, sei noch angemerkt, dass für ξ^* in Theorem 1.2.6 stets $\xi_i^* > 0$, $i \in \mathcal{S}$ gilt (siehe [5, S.63]).

Falls $\mathcal{S}$ *endlich* ist, kann gezeigt werden, dass eine *gleichmäßige, geometrische Konvergenz* vorliegt. Von besonderer Bedeutung ist hierbei der zweitgrößte Eigenwert der Übergangsmatrix. Wir notieren hierzu ein Theorem, das letztendlich im Zusammenhang mit der *Theorie nichtnegativer Matrizen* bewiesen wird. Insbesondere spielt der Satz von *Perron Frobenius* eine entscheidende Rolle. Für Details sei der Leser wieder auf [5] verwiesen.

Sei X eine irreduzible, aperiodische DMK mit Zustandsraum $\mathcal{S} = \{1, \dots, r\}$ und Übergangsmatrix $\mathbb{P}$. Es ist 1 offenbar ein positiver Eigenwert von $\mathbb{P}$, da $\mathbb{P}1_\mathcal{S} = 1_\mathcal{S}$ mit $1_\mathcal{S} = (1, \dots, 1)^t \in \mathbb{R}^\mathcal{S}$ gilt. Dieser besitzt nach dem Satz von *Perron Frobenius* ferner die algebraische Vielfachheit 1 und dominiert alle anderen Eigenwerte von $\mathbb{P}$ strikt. Sei

$$\rho \stackrel{\text{def}}{=} \max\{|\lambda| \,:\, \lambda \text{ ist Eigenwert von } \mathbb{P} \text{ und } \lambda \neq 1\} < 1$$

und

$$d^* \stackrel{\text{def}}{=} \max\{d \,:\, d \text{ ist die Größe eines Jordanblocks zu } \lambda \in \mathcal{E}_\rho\},$$

wobei $\mathcal{E}_\rho$ die Menge der Eigenwerte vom Betrag ρ bezeichne. Wir haben dann

Theorem 1.2.7. *Für zwei Konstanten $c, C \in (0, \infty)$ gilt unter obigen Voraussetzungen*

$$ck^{d^*-1}\rho^k \leq \max_{1 \leq i \leq r} \|P_i^{X_k} - \xi^*\| \leq Ck^{d^*-1}\rho^k$$

für alle $k \geq 1$ (gleichmäßige geometrische Ergodizität).

Die konkrete Berechnung von ρ, d^*, c, C ist allerdings i.Allg. sehr aufwendig. Die interessante Aussage des Theorems ist daher eher qualitativer Natur, obwohl verschiedene Verfahren zur Analyse obiger Parameter existieren (siehe [5]). Begnügen wir uns mit der Abschätzung

$$\max_{1 \leq i \leq r} \|P_i^{X_k} - \xi^*\| \leq \alpha\beta^k, \quad k \geq 0 \tag{1.7}$$

für ein $\alpha > 0$ und $0 < \beta < 1$, so lässt sich diese Schranke auch ohne Verwendung der Theorie nichtnegativer Matrizen nachweisen. Genauer basiert der Beweis in [5]

von Theorem 1.2.6 auf einem sehr eleganten Kopplungsargument, das hier

$$\max_{1\le i\le r} \|P_i^{X_k} - \xi^*\| \le a_k, \quad k \ge 0$$

für eine Nullfolge $(a_k)_{k\ge 0}$ liefert. Da $\mathcal{S}$ endlich ist, können wir, wie in [5] ausführlich gezeigt wird, die Konvergenzrate von $(a_k)_k$ durch $a_k \le \alpha\beta^k$, $k \ge 0$ abschätzen, woraus (1.7) folgt. Dies liegt letztendlich daran, dass die sogenannte *Doeblin Bedingung*

$$\exists i_0 \in \mathcal{S},\, k_0 \ge 1 \,:\, \inf_{i\in\mathcal{S}} P_i(X_{k_0} = i_0) > 0$$

wegen

$$P_i(X_k = i_0) \to \xi^*_{i_0} > 0, \quad k \to \infty \quad \forall i, i_0 \in \mathcal{S},$$

siehe Theorem 1.2.6, und der Endlichkeit von $\mathcal{S} = \{1, \dots, r\}$ immer erfüllt ist.

Wir haben den *Variationsabstand* als naheliegendes formales Mittel zur Messung der Abweichung zweier Verteilungen eingeführt. Eine alternative Methode liegt in der Betrachtung des *Separationsabstandes*.

Definition 1.2.8. Seien Q_1, Q_2 zwei Verteilungen auf $\mathcal{S}$. Dann heißt

$$\mathrm{sep}(Q_1, Q_2) \stackrel{\mathrm{def}}{=} \sup_{s\in\mathcal{S}} \left\{1 - \frac{Q_1(s)}{Q_2(s)}\right\}$$

der *Separationsabstand* zwischen Q_1 und Q_2, wobei die Konvention $\frac{x}{0} \stackrel{\mathrm{def}}{=} \infty$, $x \ge 0$ benutzt werde.

Es muss allerdings erwähnt werden, dass *sep* die Axiome einer Metrik verletzt und damit die grundsätzlichen Prinzipien, die eine jede Abstandsmessung erfüllen sollte, *nicht* erfüllt. Offenbar genügt der Separationsabstand aber noch den Aussagen

$$0 \le \mathrm{sep}(Q_1, Q_2) \le 1 \quad \text{und} \quad \mathrm{sep}(Q_1, Q_2) = 0 \iff Q_1 = Q_2$$

für beliebige Verteilungen Q_1, Q_2 auf $\mathcal{S}$. Erstere folgt, da für zwei Verteilungen Q_1, Q_2 wegen deren Normiertheit unmöglich $Q_2(s) < Q_1(s)$ für alle $s \in \mathcal{S}$ erfüllt sein kann. Zweitere ist klar, da $Q_1 \ne Q_2$ die Existenz eines $s \in \mathcal{S}$ impliziert mit $Q_1(s) < Q_2(s)$. Allerdings ist schon die Symmetrie $\mathrm{sep}(Q_1, Q_2) = \mathrm{sep}(Q_2, Q_1)$ i.Allg. nicht mehr erfüllt. Betrachte hierzu z.B.

$$\mathcal{S} = \{1, 2\}, \quad Q_1(1) = Q_1(2) = \frac{1}{2}, \quad Q_2(1) = \frac{1}{3},\, Q_2(2) = \frac{2}{3}.$$

Es folgt dann

$$\mathrm{sep}(Q_1, Q_2) = \frac{1}{4} \ne \frac{1}{3} = \mathrm{sep}(Q_2, Q_1).$$

Folgendes Lemma verschafft etwas mehr Einsicht in die Bedeutung des Separationsabstandes.

Lemma 1.2.9. *Seien Q_1, Q_2 zwei Verteilungen auf $\mathcal{S}$ und*

$$K \stackrel{\mathrm{def}}{=} \{\alpha \in \mathbb{R}^{\ge 0} \,:\, Q_1 = (1-\alpha)Q_2 + \alpha V_\alpha \text{ für eine Verteilung } V_\alpha \text{ auf } \mathcal{S}\}.$$

Dann gilt

$$\mathrm{sep}(Q_1, Q_2) = \inf K = \min K. \tag{1.8}$$

Beweis. Sei $\alpha \in K$. Wir haben dann

$$Q_1(s) - (1-\alpha)Q_2(s) \geq 0 \;\Rightarrow\; \alpha \geq 1 - \frac{Q_1(s)}{Q_2(s)}, \quad s \in \mathcal{S},$$

woraus $\alpha \geq \mathrm{sep}(Q_1, Q_2)$ folgt. Umgekehrt ergibt

$$\mathrm{sep}(Q_1, Q_2) \geq 1 - \frac{Q_1(s)}{Q_2(s)}, \quad s \in \mathcal{S}$$

sofort

$$Q_1(s) - (1 - \mathrm{sep}(Q_1, Q_2))Q_2(s) \geq 0,\ s \in \mathcal{S},$$

was $\mathrm{sep}(Q_1, Q_2) \in K$ nach sich zieht, womit das Lemma bewiesen ist. □

Als Vergleich zwischen dem Separationsabstand und Variationsabstand haben wir die Ungleichungskette

$$0 \leq \|Q_1 - Q_2\| \leq \mathrm{sep}(Q_1, Q_2) \leq 1, \tag{1.9}$$

wobei die mittlere Ungleichung wegen

$$\begin{aligned} \|Q_1 - Q_2\| &= \sum_{\{Q_2 \geq Q_1\}} (Q_2(s) - Q_1(s)) \\ &\leq \sum_{\{Q_2 \geq Q_1\}} Q_2(s) \sup_{i \in \mathcal{S}} \left(1 - \frac{Q_1(i)}{Q_2(i)}\right) \\ &\leq \mathrm{sep}(Q_1, Q_2) \end{aligned}$$

gilt. Der Separationsabstand lässt sich zumindest in den Beispielen dieser Arbeit wesentlich leichter auswerten als der Variationsabstand. Im Zusammenhang seiner Analyse spielt das Prinzip der *stark stationären Zeiten* eine wichtige Rolle. Dieses werden wir nun in einem leicht verallgemeinerten Rahmen (siehe auch [5]) näher betrachten: Sei X eine DMK bzgl. einer Filtration $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$, d.h. X ist $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$-adaptiert und es gilt

$$P^{X_{k+1}|\mathcal{F}_k} = P^{X_{k+1}|X_k} = P^{X_1|X_0}, \quad k \in \mathbb{N}_0,$$

wobei die zweite Gleichung die Homogenität fordert. Diese Verallgemeinerung wird sich in Abschnitt 2.5 als hilfreich herausstellen. Falls nicht explizit erwähnt, meinen wir mit einer DMK X immer X bzgl. der kanonischen Filtration. Für anschließende Definition wird weiter die Existenz einer stationären Verteilung ξ^* gefordert.

Definition 1.2.10. Für eine DMK X bzgl. $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$ ist eine *stark stationäre Zeit* T eine P-f.s. endliche $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$-Stoppzeit, so dass folgende drei äquivalente Bedingungen gelten:

(i) $P(X_k = i | T = k) = \xi_i^*, \quad k \in \mathbb{N}_0, i \in \mathcal{S},$

(ii) $P(X_k = i | T \leq k) = \xi_i^*, \quad k \in \mathbb{N}_0, i \in \mathcal{S},$

(iii) $P^{X_T} = \xi^*$ und X_T, T sind unabhängige Zufallsvariablen.

Bemerkung 1.2.11. Wir haben hier die Existenz einer stationären Verteilung ξ^* vorausgesetzt. Falls X zusätzlich positiv rekurrent ist, so kann gefolgert werden, dass genau eine stationäre Verteilung ξ^* existiert und dann $\xi_s^* > 0$, $\forall s \in \mathcal{S}$ gilt (siehe [5]).

Lemma 1.2.12. *Die Aussagen in Definition 1.2.10 sind äquivalent.*

Beweis. (i) $\Rightarrow$ (ii) : Zunächst wird mit vollständiger Induktion nach u

$$P(X_{k+u} = i|T = k) = \xi_i^*, \quad k, u \in \mathbb{N}_0, i \in \mathcal{S}$$

gezeigt. Für $u = 0$ ist dieses gerade die Voraussetzung (i). Sei $u \geq 1$. Es gilt

$$\begin{aligned} & P(X_{k+u} = i, X_{k+u-1} = s, T = k) \\ = & \int_{\{X_{k+u-1}=s, T=k\}} P(X_{k+u} = i|\mathcal{F}_{k+u-1})\, dP \\ = & \int_{\{X_{k+u-1}=s, T=k\}} P(X_{k+u} = i|X_{k+u-1})\, dP, \quad p_{s,i} \stackrel{\text{def}}{=} P(X_1 = i|X_0 = s) \\ = & P(X_{k+u-1} = s, T = k)\, p_{s,i}. \end{aligned}$$

Damit folgt

$$\begin{aligned} & P(X_{k+u} = i|T = k), \quad k, u \in \mathbb{N}_0 \\ = & \frac{1}{P(T = k)} \sum_{s\in\mathcal{S}} P(X_{k+u} = i, X_{k+u-1} = s, T = k) \\ = & \frac{1}{P(T = k)} \sum_{s\in\mathcal{S}} P(X_{k+u-1} = s, T = k)\, p_{s,i} \\ = & \sum_{s\in\mathcal{S}} P(X_{k+u-1} = s|T = k)\, p_{s,i} \\ = & \sum_{s\in\mathcal{S}} \xi_s^*\, p_{s,i} \\ = & \xi_i^*, \end{aligned}$$

wobei die Induktionsvoraussetzung in der vorletzten Zeile benutzt wurde. Hieraus ergibt sich für $k \in \mathbb{N}_0, i \in \mathcal{S}$

$$P(X_k = i|T \leq k) = \frac{1}{P(T \leq k)} \sum_{j=0}^{k} P(X_k = i|T = j)P(T = j) = \xi_i^*.$$

(ii) $\Rightarrow$ (i) : Für $k \in \mathbb{N}, i \in \mathcal{S}$ haben wir

$$\begin{aligned} P(X_k = i, T = k) &= P(X_k = i, T \leq k) - P(X_k = i, T \leq k - 1) \\ &= \xi_i^* P(T \leq k) - \xi_i^* P(T \leq k - 1) \\ &= \xi_i^* P(T = k), \end{aligned}$$

da analog wie oben $P(X_k = i|T \leq k - 1) = \xi_i^*$ gezeigt werden kann.
(iii) $\Rightarrow$ (i) : Für $k \in \mathbb{N}_0, i \in \mathcal{S}$ folgt

$$P(X_k = i|T = k) = P(X_T = i|T = k) = P(X_T = i) = \xi_i^*.$$

(i) $\Rightarrow$ (iii) : Für $k \in \mathbb{N}_0, i \in \mathcal{S}$ gilt

$$P(X_T = i|T = k) = \xi_i^*.$$

Daraus folgt

$$P(X_T = i) = \sum_{k \in \mathbb{N}_0} P(X_T = i | T = k) P(T = k) = \xi_i^*, \quad i \in \mathcal{S}.$$

Wegen

$$P(X_T = i | T = k) = P(X_T = i), \quad k \in \mathbb{N}_0, i \in \mathcal{S}$$

folgt weiter die Unabhängigkeit von X_T und T. □

Die entscheidende Aussage lautet

Satz 1.2.13. *Sei T eine stark stationäre Zeit zu der $(\mathcal{F}_k)_{k \in \mathbb{N}_0}$ - DMK X und ξ^* eine stationäre Verteilung. Dann gilt*

$$\text{sep}(P^{X_k}, \xi^*) \leq P(T > k), \quad k \geq 0. \tag{1.10}$$

Beweis. Es gilt für $i \in \mathcal{S}, k \geq 0$

$$P(X_k = i) \geq P(X_k = i, T \leq k) = P(T \leq k)\xi_i^* = (1 - P(T > k))\xi_i^*,$$

woraus

$$1 - \frac{P(X_k = i)}{\xi_i^*} \leq P(T > k)$$

folgt und damit wegen

$$\text{sep}(P^{X_k}, \xi^*) = \sup_{i \in \mathcal{S}} \left\{ 1 - \frac{P(X_k = i)}{\xi_i^*} \right\}$$

die Behauptung. □

Die Stoppzeit T beschreibt einen Zeitpunkt, ab dem der Prozess stationär wird. Gelingt es uns, eine solche Stoppzeit zu finden, so erhalten wir über (1.10) eine obere Schranke von $\text{sep}(P^{X_k}, \xi^*)$. Wir werden hiervon am Ende von Abschnitt 2.3 und in Abschnitt 2.5 Gebrauch machen.

Zum Schluss sei noch bemerkt, dass für eine *positiv rekurrente* DMK X mit stationärer Verteilung ξ^* und

$$d(k) \stackrel{\text{def}}{=} \|P^{X_k} - \xi^*\|, \quad s(k) \stackrel{\text{def}}{=} \text{sep}(P^{X_k}, \xi^*) \tag{1.11}$$

folgendes Monotonieverhalten besteht:

Lemma 1.2.14. *Für alle $k \geq 0$ gilt*

(i) $d(k+1) \leq d(k)$,

(ii) $s(k+1) \leq s(k)$.

Beweis. (i) Es gilt

$$\begin{aligned} 2d(k+1) &= \sum_{i\in\mathcal{S}} |P(X_{k+1}=i) - \xi_i^*| \\ &= \sum_{i\in\mathcal{S}} |\sum_{j\in\mathcal{S}} P(X_k=j)p_{j,i} - \sum_{j\in\mathcal{S}} \xi_j^* p_{j,i}| \\ &\leq \sum_{i\in\mathcal{S}}\sum_{j\in\mathcal{S}} p_{j,i}|P(X_k=j) - \xi_j^*| \\ &= \sum_{j\in\mathcal{S}} \left(\sum_{i\in\mathcal{S}} p_{j,i}\right) |P(X_k=j) - \xi_j^*| \\ &= \sum_{j\in\mathcal{S}} |P(X_k=j) - \xi_j^*| \\ &= 2d(k), \end{aligned}$$

wobei wieder $p_{j,i} \stackrel{\text{def}}{=} P(X_1=i|X_0=j)$ gesetzt wurde.

(ii) Offenbar reicht es

$$\inf_{i\in\mathcal{S}} \left\{ \frac{P(X_{k+1}=i)}{\xi_i^*} \right\} \geq \inf_{i\in\mathcal{S}} \left\{ \frac{P(X_k=i)}{\xi_i^*} \right\}$$

zu zeigen. Wir werden nachweisen, dass zu jedem fixierten $i \in \mathcal{S}$ ein $j = j(i) \in \mathcal{S}$ existiert mit

$$\frac{P(X_{k+1}=i)}{\xi_i^*} \geq \frac{P(X_k=j)}{\xi_j^*}.$$

Das ist äquivalent zu

$$P(X_{k+1}=i)\xi_j^* \geq P(X_k=j)\xi_i^* \quad \text{für ein } j \in \mathcal{S}.$$

Angenommen, es gilt

$$P(X_{k+1}=i)\xi_j^* < P(X_k=j)\xi_i^*, \quad \forall j \in \mathcal{S}.$$

Wegen der Rekurrenz von X existiert ein $j_0 \in \mathcal{S}$ mit $P(X_1=i|X_0=j_0) > 0$. Dies implizierte

$$P(X_{k+1}=i)\xi_j^* P(X_1=i|X_0=j) \leq P(X_k=j)\xi_i^* P(X_1=i|X_0=j)$$

mit strikter Ungleichheit für mindestens ein $j \in \mathcal{S}$. Durch Summation erhielten wir daher

$$P(X_{k+1}=i)\sum_{j\in\mathcal{S}} \xi_j^* P(X_1=i|X_0=j) < \xi_i^* \sum_{j\in\mathcal{S}} P(X_k=j)P(X_1=i|X_0=j),$$

was

$$P(X_{k+1}=i)\xi_i^* < \xi_i^* P(X_{k+1}=i)$$

implizierte, also einen Widerspruch. □

1.3 Random Walks auf endlichen Gruppen

In diesem Abschnitt werden speziell einige Aussagen über die endlichen, homogenen Markov-Ketten in Satz 1.1.2 getroffen. Genauer werden hier Markov-Ketten X auf endlichen Gruppen G, d.h. $\mathcal{S} = G$ untersucht, wobei die Übergangsmatrix $\mathbb{P}$ von der Form

$$\mathbb{P}(\tau, \sigma) = Q(\tau^{-1} \circ \sigma), \quad \tau, \sigma \in G \tag{1.12}$$

für eine Verteilung Q auf G ist. X wird in diesem Fall auch ein *Random Walk* auf G genannt.

Lemma 1.3.1. *Es gilt*

(i) $\sum_{\tau \in G} \mathbb{P}(\tau, \sigma) = 1, \quad \sigma \in G$, *d.h.* $\mathbb{P}$ *ist eine doppelt stochastische Matrix.*

(ii) *Die Gleichverteilung* U *auf* $\mathfrak{S}_n$ *ist eine stationäre Verteilung.*

Beweis. (i) $\sum_{\tau \in G} \mathbb{P}(\tau, \sigma) = \sum_{\tau \in G} Q(\tau) = 1, \quad \sigma \in G$

(ii) $\sum_{\tau \in G} \mathbb{P}(\tau, \sigma) U(\tau) = \frac{1}{n!} \sum_{\tau \in G} \mathbb{P}(\tau, \sigma) = \frac{1}{n!} = U(\sigma), \quad \sigma \in G.$

□

Aus Theorem 1.2.6 folgt sofort, dass, falls X aperiodisch ist, X im Sinne von (1.6) gegen U konvergiert. Wie wir unter diesem Theorem konstatiert haben, liegt dann sogar *gleichmäßige geometrische Ergodizität* vor.

Beispiel 1.3.2. Der *Top-to-Random-Shuffle* (siehe Einleitung zu Kapitel 2) generiert eine aperiodische Markov-Kette. Jeder Zustand $\pi \in \mathfrak{S}_n$ ist aperiodisch, d.h. erfüllt die Bedingung in Definition 1.2.3 (iii), da der T1TRS mit positiver Wahrscheinlichkeit wieder $\pi \in \mathfrak{S}_n$ erzeugt. Das ist der Fall, wenn die erste Karte abgehoben wird und wieder oben auf das Deck zurückgelegt wird, d.h. wenn nichts passiert ist. Ferner sind alle Zustände miteinander *verbunden*, d.h. jeder Zustand kann von jedem Zustand aus in endlich vielen Schritten mit positiver Wahrscheinlichkeit erreicht werden: Angenommen, das Kartendeck befindet sich in der Reihenfolge $\tau \in \mathfrak{S}_n$. Es wird nun gezeigt, wie mit endlich vielen T1TRS ein beliebiges, fixiertes $\pi \in \mathfrak{S}_n$ erreicht werden kann: Sei $-1 \leq i \leq n-1$ maximal, so dass $\pi(n-j) = \tau(n-j)$, $j = 0, \ldots, i$ gilt, wobei $i = -1$ für $\pi(n) \neq \tau(n)$ stehe. Falls $i = n-1$, so ist nichts zu zeigen. Anderenfalls existiert ein $1 \leq k \leq n-i-2$ mit $\tau(k) = \pi(n-i-1)$. Falls wir k mal die oberste Karte abheben und jeweils zur Position $n-i-1$ zurückstecken, so gelangen wir zu einem Zustand $\widetilde{\tau} \in \mathfrak{S}_n$ mit $\pi(n-j) = \widetilde{\tau}(n-j)$, $j = 0, \ldots, i+1$. Nach endlich vielen Wiederholungen dieses Verfahrens geht das Deck schließlich in $\pi \in \mathfrak{S}_n$ über. Alle obigen Schritte besitzen eine *positive Wahrscheinlichkeit*, weshalb die zugehörige Markov-Kette auch irreduzibel ist, insgesamt also aperiodisch ist.

Dass der T1TRS gegen die Gleichverteilung konvergiert, ist daher eine triviale Konsequenz aus der allgemeinen Theorie endlicher Markov-Ketten. Ferner wird in dieser Arbeit auch der zweitgrößte Eigenwert (siehe Abschnitt 2.4) der T1TRS Übergangsmatrix ermittelt, so dass wegen Theorem 1.2.7 auch die *geometrische Konvergenzrate* ermittelt ist. Dennoch ist gerade der in der Einleitung beschriebene *Cutoff-Effekt*, der *nicht* aus der allgemeinen Theorie ableitbar ist, von besonderem Interesse und muss deshalb separat analysiert werden (siehe Abschnitt 2.3).

Wir wollen an dieser Stelle noch eine äquivalente Charakterisierung für Irreduzibilität und Aperiodizität angeben. Hierzu zunächst einige Definitionen.

Definition 1.3.3. (i) Für $T \subseteq G$ und $k \in \mathbb{N}$ sei

$$T^{\circ k} \stackrel{\text{def}}{=} \{\tau_1 \circ \ldots \circ \tau_k \,:\, \tau_i \in T,\, i = 1, \ldots, k\}.$$

(ii) Für zwei Verteilungen Q_1, Q_2 auf G sei die *Faltung* dieser Verteilungen durch

$$Q_2 * Q_1(\sigma) \stackrel{\text{def}}{=} \sum_{\tau \in G} Q_1(\tau) Q_2(\tau^{-1} \circ \sigma), \quad \sigma \in G$$

definiert. Ferner ist

$$Q_1^{*k} \stackrel{\text{def}}{=} \overbrace{Q_1 * \ldots * Q_1}^{k\text{-mal}}, \quad k \in \mathbb{N}$$

wohldefiniert, da diese Verknüpfung offenbar assoziativ ist.

(iii) Für eine Verteilung Q auf G sei $\operatorname{supp}(Q) \stackrel{\text{def}}{=} \{\tau \in G \,:\, Q(\tau) > 0\}$ der *Träger* von Q.

Offensichtlich ist $\langle T \rangle \stackrel{\text{def}}{=} \bigcup_{k \in \mathbb{N}} T^{\circ k} \subseteq G$ die kleinste Untergruppe von G, die T enthält. Die Existenz des Inversen folgt wegen der Endlichkeit von G direkt aus $\tau^{|G|} = \mathrm{id}$, $\tau \in G$. Letzteres kann z.B. aus dem in der Algebra bekannten Satz von *Lagrange* (siehe Lang [15]) leicht gefolgert werden.

Lemma 1.3.4. *Es gilt für jede Verteilung Q auf G*

$$\operatorname{supp}(Q^{*k}) = (\operatorname{supp}(Q))^{\circ k}, \quad k \in \mathbb{N}.$$

Beweis. Sei o.E. $k \geq 2$.
„$\subseteq$ ":
Für $\tau \in \operatorname{supp}(Q^{*k})$ gilt

$$Q^{*k}(\tau) = \sum_{\sigma_1, \ldots, \sigma_{k-1} \in G} Q(\sigma_1) \cdot \ldots \cdot Q(\sigma_{k-1}) \cdot Q(\sigma_{k-1}^{-1} \circ \ldots \circ \sigma_1^{-1} \circ \tau) > 0.$$

Es müssen daher $\gamma_1, \ldots, \gamma_{k-1} \in \operatorname{supp}(Q)$ existieren mit $\gamma_{k-1}^{-1} \circ \ldots \circ \gamma_1^{-1} \circ \tau \in \operatorname{supp}(Q)$. Hieraus folgt

$$\tau = \gamma_1 \circ \ldots \circ \gamma_{k-1} \circ (\gamma_{k-1}^{-1} \circ \ldots \circ \gamma_1^{-1} \circ \tau) \in (\operatorname{supp}(Q))^{\circ k}.$$

„$\supseteq$ ":
Sei $\tau \in (\operatorname{supp}(Q))^{\circ k}$. Dann existieren $\gamma_1, \ldots, \gamma_k \in \operatorname{supp}(Q)$ mit $\tau = \gamma_1 \circ \ldots \circ \gamma_k$. Dies ergibt

$$\begin{aligned} Q^{*k}(\tau) &= \sum_{\sigma_1, \ldots, \sigma_{k-1} \in G} Q(\sigma_1) \cdot \ldots \cdot Q(\sigma_{k-1}) \cdot Q(\sigma_{k-1}^{-1} \circ \ldots \circ \sigma_1^{-1} \circ \tau) \\ &\geq Q(\gamma_1) \cdot \ldots \cdot Q(\gamma_{k-1}) \cdot Q(\gamma_{k-1}^{-1} \circ \ldots \circ \gamma_1^{-1} \circ \tau) \\ &= \prod_{i=1}^{k} Q(\gamma_i) > 0, \end{aligned}$$

d.h. $\tau \in \operatorname{supp}(Q^{*k})$. □

Wir erhalten insbesondere $\langle \operatorname{supp}(Q) \rangle = \bigcup_{k \in \mathbb{N}} \operatorname{supp}(Q^{*k})$. Hieraus folgt sofort

Korollar 1.3.5. *Sei X eine wie zu Beginn dieses Abschnitts definierte Markov-Kette. Es ist X genau dann irreduzibel, wenn $\langle \mathrm{supp}(Q)\rangle = G$ gilt.*

Beweis. Wir haben für $\tau, \sigma \in G$, $k \in \mathbb{N}$

$$P(X_k = \tau | X_0 = \sigma) = \sum_{\substack{\sigma_1,\ldots,\sigma_k \in G,\\ \sigma_1 \circ \ldots \circ \sigma_k = \sigma^{-1}\circ\tau}} Q(\sigma_1)\cdot\ldots\cdot Q(\sigma_k) = P(X_k = \sigma^{-1}\circ\tau | X_0 = \mathrm{id}).$$

X ist daher genau dann irreduzibel, falls zu jedem $\gamma \in G$ ein $k \in \mathbb{N}$ existiert, so dass

$$P(X_k = \gamma | X_0 = \mathrm{id}) = Q^{*k}(\gamma) > 0 \iff \gamma \in \langle \mathrm{supp}(Q)\rangle$$

gilt, womit das Korollar bewiesen ist. □

Satz 1.3.6. *Sei X irreduzibel. X ist genau dann aperiodisch, falls supp(Q) keine Teilmenge einer Nebenklasse eines nichttrivialen Normalteilers von G ist.*

Beweis. Siehe Woess [19]. □

Die Normalteiler der *symmetrischen Gruppe* $\mathfrak{S}_\mathrm{n}$ sind bekannt: Sei

$$\begin{array}{rccl} \mathrm{sgn}: & \mathfrak{S}_\mathrm{n} & \to & \{-1,1\} \\ & \pi & \mapsto & \prod_{i<j} \frac{\pi(j)-\pi(i)}{j-i}. \end{array}$$

Wir wissen, dass sgn ein Gruppenhomomorphismus ist und daher

$$\mathfrak{A}_n \stackrel{\mathrm{def}}{=} \ker(\mathrm{sgn}) = \mathrm{sgn}^{-1}(\{1\}) \subseteq \mathfrak{S}_\mathrm{n}$$

ein Normalteiler ist. $\mathfrak{A}_n$ heißt die *alternierende Gruppe*. Für $n \in \mathbb{N} - \{4\}$ ist $\mathfrak{A}_n \subseteq \mathfrak{S}_\mathrm{n}$ *der einzige* nicht triviale Normalteiler von $\mathfrak{S}_\mathrm{n}$. Falls $n = 4$, so ist neben $\mathfrak{A}_4$ zusätzlich die *Kleinsche Vierergruppe* $\mathfrak{V}_4 \subseteq \mathfrak{S}_4$ ein Normalteiler. $\mathfrak{S}_4$ besitzt genau diese beiden nicht trivialen Normalteiler. Hierbei ist

$$\mathfrak{V}_4 \stackrel{\mathrm{def}}{=} \{\mathrm{id}, (1,2)\circ(3,4), (1,3)\circ(2,4), (1,4)\circ(2,3)\},$$

wobei

$$\begin{array}{rccll} (i,j): & \{1,\ldots,4\} & \to & \{1,\ldots,4\} & \in \mathfrak{S}_4, \quad 1 \le i < j \le 4 \\ & k & \mapsto & \begin{cases} k & \text{falls } k \notin \{i,j\}, \\ j & \text{falls } k = i, \\ i & \text{falls } k = j \end{cases} & \end{array}$$

eine sogenannte *Transposition* ist. Siehe auch Lang [15] für weitere gruppentheoretische Resultate.

Zum Abschluss dieses Kapitels werden wir noch etwas genauer auf den Separationsabstand (Definition 1.2.8) eingehen. Es ist trivial, dass, falls der Separationsabstand minimiert wird, ebenso der Variationsabstand entsprechend minimiert wird (siehe (1.9)). Interessant ist, dass auch eine umgekehrte Aussage für *Random Walks auf endlichen Gruppen* wahr ist. Grob gesprochen wird der Separationsabstand im Sinne von (1.11) nach höchstens $2k$ Schritten minimiert, falls der Variationsabstand hierzu k Schritte benötigt. Diese Aussage werden wir nun formal präzisieren und beweisen. Folgendes ist eine Ausarbeitung von Proposition 5.13 in [2].

Satz 1.3.7. *Sei X ein Random Walk auf G, der in einem Zustand $\pi_0 \in G$ starte, d.h. $P(X_0 = \pi_0) = 1$. Seien weiter $d(k), s(k)$ wie in (1.11) mit $\xi_\pi^* \stackrel{\text{def}}{=} \frac{1}{|G|}$, $\pi \in G$ (Gleichverteilung) definiert, siehe (1.3.1). Es gilt dann mit*

$$\phi(\epsilon) \stackrel{\text{def}}{=} 1 - (1 - 2\epsilon^{1/2})(1 - \epsilon^{1/2})^2, \quad 0 \le \epsilon \le \frac{1}{4}$$

die Ungleichungskette

$$d(2k) \le s(2k) \le \phi(2d(k)), \quad k \ge 1, \text{ falls } d(k) \le \frac{1}{8}. \tag{1.13}$$

Wir benötigen für den Beweis zunächst ein Lemma.

Lemma 1.3.8. *Sei $(q_1, \ldots, q_u) \in [0,1]^u$ eine Wahrscheinlichkeitsverteilung, d.h. $\sum_{i=1}^u q_i = 1$. Falls*

$$\sum_{i=1}^u \left| q_i - \frac{1}{u} \right| \le \epsilon \le \frac{1}{4}$$

für ein $\epsilon \ge 0$, so gilt

$$\sum_{i=1}^u q_i q_{\pi(i)} \ge \frac{1 - \phi(\epsilon)}{u}, \quad \pi \in \mathfrak{S}_u. \tag{1.14}$$

Beweis. Falls $\epsilon = 0$, so folgt $q_i = \frac{1}{u}$, $i = 1, \ldots, u$ und $\phi(\epsilon) = 0$, woraus unmittelbar (1.14) folgt. Sei also o.E. $\epsilon > 0$. Wähle $2 \le \alpha \le \frac{1}{\epsilon}$ und sei

$$I \stackrel{\text{def}}{=} \left\{ 1 \le i \le u \,:\, \left| q_i - \frac{1}{u} \right| > \frac{\alpha\epsilon}{u} \right\} \subseteq \{1, \ldots, u\}.$$

Es folgt dann

$$\sum_{i \in I} \left| q_i - \frac{1}{u} \right| \ge \frac{\alpha\epsilon |I|}{u},$$

woraus wegen unserer Voraussetzung $|I| \le \frac{u}{\alpha}$ gelten muss. Hieraus ergibt sich für $\pi \in \mathfrak{S}_u$

$$\begin{aligned} |I^c \cap \pi^{-1}(I^c)| &= |I^c| + |\pi^{-1}(I^c)| - |I^c \cup \pi^{-1}(I^c)| \\ &= 2(u - |I|) - |I^c \cup \pi^{-1}(I^c)| \\ &\ge \left\lceil 2\left(u - \frac{u}{\alpha}\right) - u \right\rceil \\ &= \left\lceil u - \frac{2u}{\alpha} \right\rceil. \end{aligned}$$

Für mindestens $\lceil u - \frac{2u}{\alpha} \rceil$ Elemente i aus $\{1, \ldots, u\}$ gilt daher $i \notin I$ und $\pi(i) \notin I$. Wegen

$$\frac{1}{u} - q_z \le \left| q_z - \frac{1}{u} \right| \le \frac{\alpha\epsilon}{u}, \quad z = i, \pi(i), \quad i \in I^c \cap \pi^{-1}(I^c)$$

folgt für *solche* Elemente

$$q_i \ge \frac{1}{u} - \frac{\alpha\epsilon}{u} \ge 0 \text{ und } q_{\pi(i)} \ge \frac{1}{u} - \frac{\alpha\epsilon}{u} \ge 0.$$

Insgesamt haben wir daher

$$\begin{aligned}\sum_{i=1}^{u} q_i q_{\pi(i)} &\geq \left\lceil u - \frac{2u}{\alpha} \right\rceil \left(\frac{1}{u} - \frac{\alpha\epsilon}{u}\right)^2 \\ &\geq \left(u - \frac{2u}{\alpha}\right)\left(\frac{1}{u} - \frac{\alpha\epsilon}{u}\right)^2 \\ &= \frac{1}{u}\left(1 - \frac{2}{\alpha}\right)(1-\alpha\epsilon)^2.\end{aligned}$$

Mit $\alpha \stackrel{\text{def}}{=} \epsilon^{-1/2}$ erhalten wir die Aussage des Lemmas. Diese Wahl ist zulässig, da $0 < \epsilon \leq \frac{1}{4}$ die Abschätzung $\epsilon^{-1/2} \geq 2$ impliziert und $\epsilon^{-1/2} \leq \epsilon^{-1}$ aus $\epsilon \leq 1$ folgt. □

Mit Hilfe dieses Lemmas können wir Satz 1.3.7 leicht beweisen.

Beweis. Sei $k \geq 1$ und $\pi \in G$. Es gilt dann

$$\begin{aligned}& P(X_{2k} = \pi) \\ = & \sum_{\tau\in G} P(X_{2k} = \pi | X_k = \tau) P(X_k = \tau) \\ = & \sum_{\tau\in G} P(X_{2k} = \pi | X_k = \tau) P(X_k = \tau | X_0 = \pi_0) \\ = & \sum_{\tau\in G} Q^{*k}(\tau^{-1}\circ\pi) Q^{*k}(\pi_0^{-1}\circ\tau), \quad \gamma \stackrel{\text{def}}{=} \pi_0^{-1}\circ\tau \\ = & \sum_{\gamma\in G} Q^{*k}(\gamma^{-1}\circ\pi_0^{-1}\circ\pi) Q^{*k}(\gamma).\end{aligned}$$

$\Lambda : \gamma \mapsto \gamma^{-1}\circ\pi_0^{-1}\circ\pi$ ist eine Bijektion in G, da $\Lambda^{-1} : \gamma \mapsto \pi_0^{-1}\circ\pi\circ\gamma^{-1}$ die Inverse ist, d.h. $\Lambda\circ\Lambda^{-1} = \Lambda^{-1}\circ\Lambda = \mathrm{id}_G$ erfüllt. Für

$$d(k) \leq \frac{1}{8} \iff 2d(k) = \sum_{\gamma\in G} \left| Q^{*k}(\gamma) - \frac{1}{|G|} \right| \leq \frac{1}{4}$$

erhalten wir mit Lemma 1.3.8 somit

$$P(X_{2k} = \pi) \geq \frac{1 - \phi(2d(k))}{|G|}, \quad \pi \in G.$$

Wir haben also

$$\phi(2d(k)) \geq 1 - \frac{P(X_{2k} = \pi)}{\frac{1}{|G|}}, \quad \pi \in G,$$

was $\phi(2d(k)) \geq s(2k)$ impliziert. Damit ist insgesamt der Beweis erbracht, da die erste Ungleichung aus (1.13) schon aus (1.9) bekannt ist. □

Eine einfache Rechnung zeigt

$$\phi(\epsilon) = 4\epsilon^{\frac{1}{2}}(1 + o(1)), \quad \epsilon \to 0,$$

also $\phi(\epsilon) \cong 4\epsilon^{\frac{1}{2}}$ für $\epsilon \to 0$. Insbesondere folgt hieraus $\phi(\epsilon) \to 0$, $\epsilon \to 0$. Gerade dies macht Satz 1.3.7 interessant. Wir geben zum Abschluss noch einen mittels *Maple* erstellten Plot von $\epsilon \to \phi(\epsilon)$, $0 \leq \epsilon \leq \frac{1}{4}$ an.

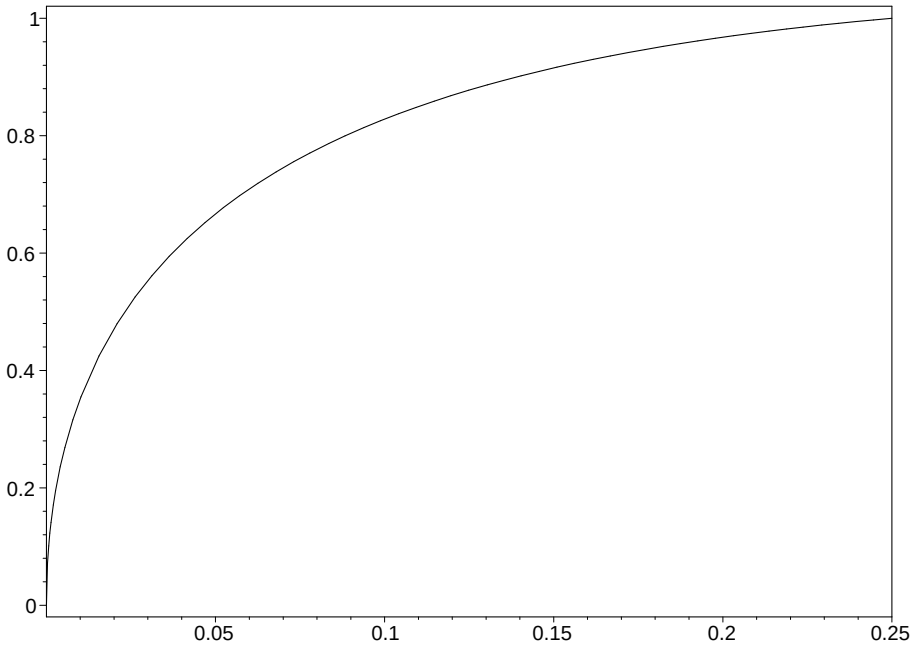

Abbildung 1.1: Darstellung der Funktion $\epsilon \mapsto \phi(\epsilon)$

Kapitel 2

Top-to-Random-Shuffles

Wir werden uns in diesem Kapitel mit einer speziellen Methode des Kartenmischens beschäftigen. Bei dieser Methode wird die oberste *(Top)* Karte eines gegebenen Kartendecks abgehoben und an eine zufällige *(Random)* Position des verbleibenden Decks wieder zurückgesteckt. Diese Mischvariante besitzt die natürliche Namensgebung *Top-to-Random-Shuffle*. Unser Hauptanliegen wird in der Ermittlung der nötigen unabhängigen Wiederholungen obiger Mischvariante bestehen, derer es bedarf, um ein sortiertes Kartendeck nahezu perfekt zu durchmischen. Es wird sich herausstellen, dass n sortierte Karten nach einer Größenordnung von $n \log n$ *Top-to-Random-Shuffles* nahezu perfekt durchmischt sind. Diese verbale Beschreibung muss natürlich noch formalisiert werden, so dass die nicht zu vermeidenden Ungenauigkeiten obiger Prosa durch den üblichen mathematischen Formalismus eliminiert werden.

In dem ursprünglichen Mischverfahren wird immer nur die oberste Karte an eine Zufallsposition zurückgesteckt. Wir werden im Sinne einer Verallgemeinerung auch die Situation untersuchen, in der $0 \leq m \leq n$ Karten abgehoben werden und dann an jeweils zufällige Positionen unabhängig voneinander in das verbleibende Deck zurückgesteckt werden. Hierbei werden wir auch die Variante untersuchen, bei der m zwischen jeder Mischiteration nach einer vorgegebenen Wahrscheinlichkeitsverteilung variiert. Besonders erwähnenswert ist hierbei die Tatsache, dass bei dem Zurückstecken der m Karten die relative Ordnung zwischen diesen *nicht* erhalten bleiben muss, sie dem Deck also völlig willkürlich wieder zurückgeführt werden. Das ist ein wesentlicher Unterschied zu dem sogenannten *Riffle-Shuffle*, bei dem das Kartendeck zuerst durch einfaches Abheben in zwei etwa gleich große Decks gespalten wird und diese dann in bekannter Weise quasi fächerförmig wieder zusammengefügt werden. Hierbei bleibt die relative Ordnung der beiden ursprünglichen Decks natürlich erhalten. Als naheliegendes Pendant zu dem *Top-to-Random-Shuffle* ist es recht interessant, einige Aspekte des *Riffle-Shuffles* auch in dieser Arbeit zu betrachten, da durchaus einige Verbindungen zu den *Top-to-Random-Shuffles* bestehen. Bei allen hier untersuchten Mischverfahren liegt unser Hauptinteresse in dem Nachweis des in der Einleitung schon erwähnten *Cutoff-Effekts*.

Die Resulate werden meistens durch ein *direktes* Vorgehen bewiesen. Hierbei halten wir uns im Wesentlichen an Diaconis, Fill und Pitman [8]. An einigen Stellen wird aber auch das Prinzip der *stark stationären Zeiten* verwendet. Diese Technik wird im Kontext mit Kartenmischsystemen in Aldous und Diaconis [1] beschrieben, woraus hier vereinzelte Passagen diskutiert werden. In [8] wird noch auf eine Verbindung zwischen gewissen halbeinfachen Unteralgebren und Kartenmischsystemen eingegangen, was in dieser Arbeit ebenfalls diskutiert werden wird.

Zur besseren Übersicht werden wir an dieser Stelle unser weiteres Vorgehen skiz-

zieren, wobei naturgemäß einige Begrifflichkeiten verwendet werden, die erst zu einem späteren Zeitpunkt rigoros eingeführt werden.
In *Abschnitt 2.1* und in Abschnitt 2.2 wird ein Kartendeck mit einer fixierten Anzahl von n Karten betrachtet. Wir stellen uns hierbei in Abschnitt 2.1 die Frage, welche Verteilung auf $\mathfrak{S}_n$ nach k Top-to-Random-Shuffles (T1TRS) besteht. Korollar 2.1.8 gibt uns hierauf mit der geschlossenen Formel (2.19) eine Antwort.

Mit Hilfe von *Abschnitt 2.2* wird Satz 2.1.9 bewiesen. Dieser stellt eine Verallgemeinerung von Korollar 2.1.8 dar und liefert u.a. eine geschlossene Formel für die Verteilung nach k Top-m-to-Random-Shuffles (TmTRS).

In *Abschnitt 2.3* ist n nicht mehr fixiert, und es wird die Situation $n \to \infty$ betrachtet, d.h. anschaulich ein Kartendeck mit einer immer größer werdenden Anzahl von Karten. Die Anzahl der Mischschritte k sollte hierbei sinnigerweise mit n ansteigen, weshalb wir uns genauer in der Situation

$$k = k(n) = k_n, \quad n \to \infty$$

befinden. Die zentrale Frage lautet hier, welche Relation zwischen k und n zu bestehen hat, m.a.W. wie die Funktion $n \mapsto k(n) = k$ auszusehen hat, damit $k(n)$ T1TRS ausreichen / nicht ausreichen, ein Kartendeck mit n Karten für großes $n\,(n \to \infty)$ zu durchmischen. Hierbei sind wir besonders an der Existenz und der Form einer Grenzvariation interessiert, gegen die

$$\|Q_1^{*k_n} - U\|_{n\geq 1} \tag{2.1}$$

konvergiert. (2.1) bezeichnet genauer den Variationsabstand (siehe (1.2)) zwischen k_n Faltungen des T1TRS Q_1 (siehe (2.3)) und der Gleichverteilung U auf $\mathfrak{S}_n$. Theorem 2.3.7 beantwortet diese Frage und ist damit eine zentrale Aussage dieses Abschnitts und auch der gesamten Arbeit, weshalb wir es an dieser Stelle noch einmal notieren wollen:

Sei $k_n \overset{\text{def}}{=} \lfloor n \log n + cn \rfloor$ *für ein fixiertes* $c \in \mathbb{R}$. *Es gilt dann*

$$\|Q_1^{*k_n} - U\| = f(c) + o(1), \quad n \to \infty$$

mit

$$f(c) \overset{\text{def}}{=} \frac{1}{2}(1 - e^{-e^{-c}}(1 + e^{-c})), \quad c \geq 0$$

und

$$f(c) \overset{\text{def}}{=} 1 - e^{-e^{-c}} \sum_{u=0}^{l^*} e^{-uc} \left(\frac{1}{u!} - \frac{1}{(l^*+1)!} \right) - \frac{1}{(l^*+1)!}, \quad c < 0,$$

wobei

$$l^* = l^*(c) = \left\lfloor \frac{\log(e^{e^{|c|}}(e^{|c|} - 1) + 1)}{|c|} \right\rfloor - 1.$$

Obiges Theorem ist eine Folgerung aus den allgemeiner gehaltenen Aussagen in Satz 2.3.1 und Theorem 2.3.4. Es ist klar, dass wir im Folgenden besonders an dem Verhalten von $f(c)$ für $c \to \pm\infty$ interessiert sind. Es stellt sich diesbezüglich heraus, dass

$$\begin{aligned} f(c) &= \tfrac{1}{4}(1 + o(1))e^{-2c}, \quad c \to \infty && \text{(Bemerkung 2.3.13)}, \\ f(c) &= 1 - e^{-(1+o(1))e^{-c}}, \quad c \to -\infty && \text{(Bemerkung 2.3.11)} \end{aligned}$$

gilt. Wir haben daher die Konvergenz $f(c) \to 1, c \to -\infty$ mit *doppelt exponentieller* und $f(c) \to 0$, $c \to \infty$ mit *einfach* exponentieller Konvergenzrate nachgewiesen. Aufgrund dieses Konvergenzverhaltens von $f(c)$, $|c| \to \infty$ sprechen wir von einem *Cutoff-Effekt* und können sagen, dass ein Kartendeck mit n Karten für großes n nach einer Größenordnung von $n \log n$ Mischschritten durchmischt ist.

Im Anschluss hieran sind wir wieder an der Verallgemeinerung interessiert, in welcher Relation k und n zueinander stehen müssen, damit $k(n)$ TmTRS, $m \geq 1$ ausreichen / nicht ausreichen, ein Kartendeck mit n Karten ($n \to \infty$) zu durchmischen. Diese Frage beantwortet Satz 2.3.15, der wieder wesentlich auf Satz 2.3.1 und Theorem 2.3.4 basiert. Es stellt sich heraus, dass eine Größenordnung von $k_n^m \stackrel{\text{def}}{=} \frac{n}{m} \log n$ TmTRS zur Durchmischung nötig ist, wobei sich wieder dieselbe Grenzvariation wie bei dem T1TRS ergibt, so dass auch hier wieder ein *Cutoff-Effekt* vorliegt. Es besteht daher *asymptotisch gesehen* offenbar kein Unterschied zwischen k_n^m TmTRS und $k_n^m \cdot m$ T1TRS. Vergleichen wir auf einem n Kartendeck den Variationsabstand zwischen *einem* TnTRS bzw. n T1TRS zur Gleichverteilung U, so sehen wir allerdings sofort, dass zwischen beiden Vorgehensweisen i.d.R. ein deutlicher Unterschied besteht. Korollar 2.3.20 liefert in diesem Kontext die allgemeinere Aussage, dass für beliebige k, m, n mit $m \leq n$ die Verteilung von k TmTRS bzgl. des Variationsabstandes nicht weiter von der Gleichverteilung als die Verteilung von $k \cdot m$ T1TRS entfernt ist.

Zum Abschluss weisen wir in Theorem 2.3.22 mit Hilfe des Prinzips der *stark stationären Zeiten* noch einmal einen *Cutoff-Effekt* bei dem T1TRS nach. Besonders erwähnenswert ist hierbei, dass wir nicht nur eine obere Schranke des Variationsabstandes erhalten, sondern es uns auch gelingt, eine aussagekräftige untere Schranke anzugeben.

In *Abschnitt 2.4* stellen die Sätze 2.4.3 und 2.4.6 die Hauptaussagen dar. Satz 2.4.3 berechnet von der Übergangsmatrix $\mathbb{P}$ des T1TRS die Eigenwerte und deren Multiplizität. Entscheidend ist hierfür die Charakterisierung von Eigenwerten im Sinne des technischen Lemmas 2.4.1, das hier ausführlich bewiesen wird. Die besondere Bedeutung des zweitgrößten Eigenwertes von $\mathbb{P}$, der $\frac{n-2}{n}$ beträgt, wird in Theorem 1.2.7 behandelt.

Satz 2.4.6 liefert einen Zusammenhang zwischen der Algebra $\mathcal{B}$, die von den formalen Summen

$$\left(\sum_{\pi \,:\, L(\pi)=l} \pi \right)_{l=1,\ldots,n} \qquad (L(\pi) \text{ aus Lemma 2.1.1})$$

erzeugt wird, und den Top-m-to-Random-Shuffles ($0 \leq m \leq n$). Es wird hier die Kommutativität obiger Algebra und eine im Sinne der TmTRS interessante Basis angegeben. Korollar 2.4.10 besagt schließlich, dass die formalen Summen

$$\left(\sum_{\pi \,:\, G(\pi)=j} \pi \right)_{j=0,\ldots,n-1} \text{ und } \left(\sum_{\pi \,:\, F(\pi)=k} \right)_{k=1,\ldots,n} \qquad (G(\pi), F(\pi) \text{ aus Definition 2.4.9})$$

eine zu $\mathcal{B}$ isomorphe Algebra generieren, woraus u.a. folgt, dass diese ebenso kommutieren.

Wir untersuchen in *Abschnitt 2.5* zuerst eine Klasse von Mischverfahren, die eine Abänderung des TmTRS und eine vereinfachte Variante des Riffle-Shuffles als Spezialfälle enthält. Bei der Abänderung des TmTRS müssen die m abgehobenen

Karten beim Zurücklegen in das Kartendeck ihre relative Ordnung beibehalten, und bei der vereinfachten Variante des Riffle-Shuffles wird das Kartendeck vor dem Zusammenfügen immer *genau* halbiert. Auch hier interessiert uns wieder der Variationsabstand zur Gleichverteilung U, für den wir mit Hilfe des Prinzips *stark stationärer Zeiten* in Satz 2.5.6 eine obere Schranke angeben können. Es stellt sich heraus, dass die abgeänderte Variante des TmTRS asymptotisch nicht mehr als eine Größenordnung von $\frac{n}{m}\log n$ Mischschritten benötigt und die vereinfachte Form des Riffle-Shuffles nach einer Größenordnung von nicht mehr als $2\log_2 n$ Mischschritten nahe der Gleichverteilung ist. Erstere Aussage ist hierbei besonders interessant, da die Abänderung des TmTRS durch die fehlende Permutation der m Karten zumindest asymptotisch offenbar keine Verlangsamung bewirkt.

Satz 2.5.13 weist einen *Cutoff-Effekt* bei dem unter Bemerkung 2.5.10 definierten Riffle-Shuffle (Gilber, Shannon, Reeds Verteilung) bzgl. des *Separationsabstandes* nach. Dieser tritt erst nach einer Größenordnung von $2\log_2 n$ Mischschritten auf, während ein *Cutoff-Effekt* bzgl. des Variationsabstandes, siehe Theorem 2.5.16, schon nach einer Größenordung von $\frac{3}{2}\log_2 n$ Mischschritten erscheint. Falls wir den Abstand zweier Maße bzgl. des Separationsabstandes und nicht bzgl. des Variationsabstandes messen, werden wir dieses auch zukünfig immer explizit erwähnen.

Zum Abschluss wird noch eine weitere Abänderung des Riffle-Shuffles untersucht, die sich von diesem in einem wesentlichen Charakteristikum unterscheidet: Vor dem Zusammenfügen beider Kartenstapel wird der zuvor abgehobene gründlich durchmischt. Ein *Cutoff-Effekt* tritt hier nach einer Größenordnung von $\log_2 n$ Mischschritten auf, wie Satz 2.5.18 zeigt. Grob formuliert ist der Riffle-Shuffle nach Gilbert, Shannon, Reeds daher 50% langsamer als diese Variante.

2.1 Modellierung der Top-to-Random-Shuffles

Gegeben sei ein Kartendeck mit n Karten, wobei wir jede Karte mit einer natürlichen Zahl in $\{1,\dots,n\}$ identifizieren. Jede konkrete Anordnung in dem Kartendeck entspricht folglich genau einem Element $\pi \in \mathfrak{S}_n$ aus der symmetrischen Gruppe $\mathfrak{S}_n$. Mit

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 4 & 5 & 2 & 6 \end{pmatrix} \in \mathfrak{S}_6$$

ist z.B. die Realisierung eines Kartendecks mit 6 Karten gemeint, bei der Karte 3 oben liegt, gefolgt von den Karten $1, 4, 5, 2$ und 6. An Position $1 \le j \le n$ liegt also Karte $\pi(j)$, währenddessen Karte $1 \le j \le n$ sich an Position $\pi^{-1}(j)$ befindet. Mit einem sortierten n Kartendeck ist immer $\mathrm{id} \in \mathfrak{S}_n$ gemeint. Ein Mischsystem ist in diesem Sinne eine Wahrscheinlichkeitsverteilung auf $\mathfrak{S}_n$, was in Abschnitt 1.1 detailliert diskutiert wird. Wir wollen diese für den *Top-m-to-Random-Shuffle* (TmTRS), bei dem eine fixierte Anzahl von m Karten abgehoben wird und dem verbleibenden Stapel willkürlich untergemischt wird, angeben.

Lemma 2.1.1. *Sei $L(\pi) \in \{1,\dots,n\}$ die Länge der aufsteigenden Sequenz von π, die n enthält, d.h. die eindeutig bestimmte Zahl $L(\pi)$ mit der Eigenschaft*

$$\pi^{-1}(n) > \pi^{-1}(n-1) > \ldots > \pi^{-1}(n-L(\pi)+1) < \pi^{-1}(n-L(\pi)). \tag{2.2}$$

Dann ist

$$Q_m(\pi) \overset{\text{def}}{=} \begin{cases} 0 & \text{falls } L(\pi) < n-m, \\ \frac{(n-m)!}{n!} & \text{falls } L(\pi) \ge n-m \end{cases} \tag{2.3}$$

die Verteilung zu dem TmTRS.

Beweis. Da genau m Karten abgehoben werden, gilt

$$\pi^{-1}(n) > \pi^{-1}(n-1) > \ldots > \pi^{-1}(m+1),$$

woraus $Q_m(L < n-m) = 0$ nach (2.2) folgt. Direkt aus der definierenden Eigenschaft des TmTRS folgt ferner, dass jede Permutation π mit $L(\pi) \geq n-m$ mit *derselben positiven* Wahrscheinlichkeit erreicht werden kann. Wir haben daher

$$Q_m(\pi) = \frac{1}{|\{L \geq n-m\}|}, \quad \text{falls } L(\pi) \geq n-m,$$

wobei der Nenner gleich der Anzahl der Blattkombinationen ist, die bei einem TmTRS Durchgang erreicht werden können. Das entspricht offenbar der Anzahl, m unterscheidbare Kugeln in n Zellen ohne Doppelbelegungen zu legen. Es ist bekannt, dass es hierfür $\frac{n!}{(n-m)!}$ Möglichkeiten gibt. Formaler ist für $1 \leq l \leq n$ die Abbildung

$$\begin{array}{rcl} \mathfrak{S}_\mathrm{n} \supseteq \{L \geq l\} & \longrightarrow & M \\ \pi & \mapsto & (\pi^{-1}(1), \ldots, \pi^{-1}(n-l)) \end{array}$$

mit

$$M \stackrel{\text{def}}{=} \{(x_1, \ldots, x_{n-l}) \in \{1, \ldots, n\}^{n-l} \,:\, x_i \neq x_j \,\forall i \neq j\}$$

eine Mengenbijektion, und es gilt bekanntlich $|M| = \frac{n!}{l!}$, woraus

$$|\{L \geq l\}| = \frac{n!}{l!} \tag{2.4}$$

folgt. □

Ist $0 \leq m \leq n$ nicht fixiert, sondern hängt über einer Verteilung μ auf $\{0 \ldots n\}$ ebenfalls vom Zufall ab, so ergibt sich offenbar als korrespondierende Verteilung

$$Q_\mu(\pi) \stackrel{\text{def}}{=} \sum_{m=0}^{n} \mu(m) Q_m(\pi), \quad \pi \in \mathfrak{S}_\mathrm{n}\,. \tag{2.5}$$

Man stelle sich dieses Mischsystem etwa als zweistufiges Experiment vor. In einem ersten Schritt wird die Anzahl m der abzuhebenden Karten ermittelt und in einem zweiten, an welchen Stellen diese m Karten eingefügt werden. Wir interessieren uns *vorerst* für den Fall, dass die Kartendeckgröße n fixiert ist und dasselbe Mischverfahren mehrfach hintereinander auf dieses Kartendeck angewandt wird. Da diese Iterationen als unabhängig voneinander angenommen werden, müssen wir folglich die Faltung (siehe Definition 1.3.3 (ii)) der korrespondierenden Verteilungen berechnen. Hierzu definieren wir auf $\mathfrak{S}_\mathrm{n}$ die Verknüpfung $\sigma\pi \stackrel{\text{def}}{=} \pi \circ \sigma$ für alle $\sigma, \pi \in \mathfrak{S}_\mathrm{n}$, wobei mit der rechten Seite die gewöhnliche Abbildungskomposition gemeint ist. Dass diese inverse Vorgehensweise für unsere Zwecke von Vorteil ist, verdeutlicht ein Beispiel am besten.

Beispiel 2.1.2. Sei ein Deck mit 6 Karten gegeben. Zuerst(π) wird die erste Karte an Stelle 3 eingefügt und dann(σ) die erste Karte an Position 5.

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 5 & 6 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 1 & 6 \end{pmatrix}$$

Insgesamt ergibt sich offenbar

$$\begin{aligned}\sigma\pi &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ \pi(\sigma(1)) & \pi(\sigma(2)) & \pi(\sigma(3)) & \pi(\sigma(4)) & \pi(\sigma(5)) & \pi(\sigma(6)) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 1 & 4 & 5 & 2 & 6 \end{pmatrix} \\ &\neq \sigma \circ \pi.\end{aligned}$$

Nach beiden Mischvorgängen befindet sich an Position k die Karte $\sigma(k)$ des zuvor mit π gemischten Kartendecks, d.h. $\pi(\sigma(k))$, $1 \leq k \leq 6$. Die Faltung zweier Mischungen P, Q ist mit obiger Notation folglich

$$(Q * P)(\sigma) = \sum_{\pi \in \mathfrak{S}_n} Q(\sigma\pi^{-1})P(\pi) \quad \text{für alle } \sigma \in \mathfrak{S}_n\,. \tag{2.6}$$

Seien $(\mu_i)_{i=1,\ldots,k}$ Verteilungen auf $\{0,\ldots,n\}$. Es wird sich als erstes tieferliegendes Ergebnis herausstellen, dass die Faltung $Q_{\mu_k} * \ldots * Q_{\mu_1}$ wieder von der Form Q_μ ist für eine Verteilung μ auf $\{0,\ldots,n\}$. Hierzu gehen wir wie folgt vor: Für zwei Verteilungen μ, ν auf $\{0,\ldots,n\}$ sei $\nu\sharp\mu$ definiert als die Verteilung der Anzahl der besetzten Zellen, wenn zuerst i Kugeln zufällig auf n verschiedene Zellen *ohne Doppelbelegungen* verteilt werden und anschließend weitere j Kugeln *unter völliger Ignoranz* der vorherigen i Kugeln zufällig in j verschiedene Zellen derselben Zellenanordnung gelegt werden. i bzw. j werden hierfür in einem *vorgeschalteten unabhängigen* Experiment, dessen Ausgang die Realisierung einer μ bzw. ν verteilten Zufallsvariablen ist, *unabhängig voneinander* ermittelt.

Lemma 2.1.3. *Es gilt*

$$\nu\sharp\mu(k) = \sum_{i,j=0,\ldots,n} \mu(i)\nu(j)\binom{n-i}{k-i}\binom{i}{j-(k-i)}\Big/\binom{n}{j}, \tag{2.7}$$

und diese Verknüpfung ist assoziativ und kommutativ.

Beweis. Wegen der geforderten Unabhängigkeit reicht es, (2.7) für $\mu = \delta_i$ und $\nu = \delta_j$ mit fixierten $0 \leq i, j \leq n$ nachzuweisen. Wir stellen uns hierzu eine Urne mit $n-i$ roten und i schwarzen Kugeln vor. Offensichtlich entspricht die Wahrscheinlichkeit, bei j Zügen ohne Zurücklegen genau $k-i$ rote Kugeln zu ziehen, der Wahrscheinlichkeit von genau k besetzten Zellen im Sinne des vor diesem Lemma beschriebenen Experimentes. Es ergibt sich folglich eine *hypergeometrische Verteilung*, womit (2.7) bewiesen ist. Siehe auch [3, S.123].

Die Verknüpfung ist per Konstruktion assoziativ und kommutativ, da $\mu_1\sharp\ldots\sharp\mu_n$ (man beachte die fehlende Klammerung) nichts anderes als die Verteilung der besetzten Zellen angibt, wenn i_1 Kugeln, i_2 Kugeln, …, i_n Kugeln sukzessive unabhängig voneinander auf n Zellen, wie oben beschrieben, verteilt werden. □

Satz 2.1.4. *Seien $\mu_1,\ldots,\mu_k$ Verteilungen auf $\{0,\ldots,n\}$. Dann gilt*

$$Q_{\mu_k} * \ldots * Q_{\mu_1} = Q_\mu \quad \textit{mit} \quad \mu \stackrel{\text{def}}{=} \mu_k\sharp\ldots\sharp\mu_1. \tag{2.8}$$

Beweis. Wir setzen vorerst $k = 2$. Q_{μ_i} entspricht dem i-ten Mischschritt, $i = 1, 2$. Sei M_1 eine μ_1 verteilte Zufallsvariable. Zu Beginn werden M_1 Karten abgehoben

und dem Kartendeck zurückgeführt. Diese Karten versehen wir mit einer deutlichen Markierung auf ihren Rücken (zur Anschauung) und nummerieren sie von 1 bis M_1 durch. Per Definition des TmTRS befinden sich diese M_1 Karten an zufälligen Positionen. Für den zweiten Schritt sei eine von M_1 unabhängige, μ_2 verteilte Zufallsvariable M_2 gegeben. Es werden als Nächstes M_2 Karten abgehoben und dem Deck wieder zurückgeführt. Wir markieren wieder jede noch nicht markierte Karte von diesen M_2 Karten und nummerieren sie mit $M_1 + 1$ bis N durch, wobei $M_1 \leq N \leq M_1 + M_2$ gilt. Offensichtlich besitzt N die Verteilung $\mu_2 \sharp \mu_1$. Ferner besitzen alle markierten Karten zufällige Positionen innerhalb des Decks, womit der Satz für $k = 2$ bewiesen ist. Für beliebiges k iteriere man obige Idee. □

Da Verteilungen der Form Q_μ in dieser Arbeit eine große Rolle spielen, werden wir der naheliegenden Abbildung $\mu \mapsto Q_\mu$ ein Lemma widmen. Hierzu sei zur Erinnerung erwähnt, dass ein *Monoid* eine Verallgemeinerung einer Gruppe ist, bei der nicht jedes Element ein Inverses besitzen muss.

Lemma 2.1.5. *Sei*

$$\begin{array}{rcl} \varphi : \{\text{Verteilungen auf } \{0,\dots,n\}\} & \longrightarrow & \{\text{Verteilungen auf } \mathfrak{S}_\mathrm{n}\} \\ \mu & \mapsto & Q_\mu. \end{array}$$

Dann gilt

(i) *φ ist ein Monoidhomomorphismus, wobei wir die Menge der Verteilungen auf $\mathfrak{S}_\mathrm{n}$ unter der Faltung und die Menge der Verteilungen auf $\{0,\dots,n\}$ unter unserer neuen Verknüpfung $\sharp$ als Monoid betrachten.*

(ii) *φ ist nicht injektiv, aber*

$$\begin{array}{rcl} \widetilde{\varphi} : \{\text{Verteilungen auf } \{0,\dots,n\}\} & \longrightarrow & \{\text{Verteilungen auf } \mathfrak{S}_\mathrm{n}\} \times [0,1] \\ \mu & \mapsto & (Q_\mu, \mu(n)) \end{array}$$

ist injektiv.

(iii) *φ ist für $n \geq 2$ nicht surjektiv, aber*

$$\begin{array}{rcl} \widehat{\varphi} : \{\text{Verteilungen auf } \{0,\dots,n\}\} & \longrightarrow & M \\ \mu & \mapsto & Q_\mu \end{array}$$

mit

$$\begin{aligned} M \overset{\text{def}}{=} \{&Q \text{ Verteilung auf } \mathfrak{S}_\mathrm{n} : L(\pi) \leq L(\tau) \Rightarrow Q(\pi) \leq Q(\tau) \text{ und} \\ &L(\pi) = L(\tau) \Rightarrow Q(\pi) = Q(\tau),\ \pi, \tau \in \mathfrak{S}_\mathrm{n}\} \end{aligned}$$

ist surjektiv.

(iv) *Sei π_1 wie in (2.9) definiert und*

$$N \overset{\text{def}}{=} \{(Q,x) \in M \times [0,1] : x \leq n! Q(\pi_1)\}.$$

Dann ist die Abbildung

$$\begin{array}{rcl} \psi : \{\text{Verteilungen auf } \{0,\dots,n\}\} & \longrightarrow & N \\ \mu & \mapsto & (Q_\mu, \mu(n)) \end{array}$$

eine Bijektion.

Beweis. (i): Aus Satz 2.1.4 folgt

$$\varphi(\nu \sharp \mu) = Q_{\nu \sharp \mu} = Q_\nu * Q_\mu = \varphi(\nu) * \varphi(\mu), \quad \text{für alle } \nu,\, \mu.$$

Ferner werden die neutralen Elemente δ_0 bzw. δ_{id} der beiden Monoide aufeinander abgebildet, d.h. $\varphi(\delta_0) = \delta_{\mathrm{id}}$.

(ii): φ ist nicht injektiv, da $\varphi(\delta_{n-1}) = \varphi(\delta_n)$ gilt, denn $Q_{n-1} = Q_n = U$, wobei U die Gleichverteilung auf $\mathfrak{S}_\mathrm{n}$ bezeichne. Wir zeigen nun, dass $\widetilde{\varphi}$ injektiv ist. Hierzu seien π_l, $l = 1, \ldots, n-1$ durch

$$\pi_l \stackrel{\text{def}}{=} \begin{pmatrix} 1 & \ldots & n-l-1 & n-l & n-l+1 & n-l+2 & \ldots & n \\ 1 & \ldots & n-l-1 & n-l+1 & n-l & n-l+2 & \ldots & n \end{pmatrix} \tag{2.9}$$

definiert und $\pi_n \stackrel{\text{def}}{=} \mathrm{id}$. Wegen $L(\pi_l) = l$, $l = 1, \ldots, n$ gilt

$$\begin{aligned} Q_\mu(\pi_l) &= \sum_{m=0}^{n} \mu(m) Q_m(\pi_l) \\ &= \sum_{m=0}^{n} \mu(m) \frac{(n-m)!}{n!} \mathbb{1}_{\{L \geq n-m\}}(\pi_l) \\ &= \sum_{m=0}^{n} \mu(m) \frac{(n-m)!}{n!} \mathbb{1}_{\{n-m,\ldots,n\}}(l) \\ &= \sum_{m=n-l}^{n} \mu(m) \frac{(n-m)!}{n!}, \quad j \stackrel{\text{def}}{=} n-m \\ &= \sum_{j=0}^{l} \mu(n-j) \frac{j!}{n!}, \quad l = 1, \ldots, n. \end{aligned} \tag{2.10}$$

Hieraus folgt

$$\mu(n-l) = \frac{n!}{l!} \left(Q_\mu(\pi_l) - \sum_{j=0}^{l-1} \mu(n-j) \frac{j!}{n!} \right), \quad l = 1, \ldots, n. \tag{2.11}$$

Ist die Verteilung Q_μ, d.h. insbesondere auch $Q_\mu(\pi_l)$, $l = 1, \ldots, n$ bekannt und kennen wir noch zusätzlich $\mu(n)$, so können wir mittels (2.11) iterativ

$$\mu(n-1), \mu(n-2), \ldots, \mu(0),$$

also die Verteilung μ ermitteln. Für zwei Verteilungen μ, ν auf $\{0, 1, \ldots, n\}$ folgt daher

$$\widetilde{\varphi}(\mu) = \widetilde{\varphi}(\nu) \implies Q_\mu = Q_\nu,\, \mu(n) = \nu(n) \implies \mu = \nu.$$

$\widetilde{\varphi}$ ist folglich injektiv.

(iii): Für $n \geq 2$ folgt aus (2.5) $Q_\mu(\mathrm{id}) \geq Q_\mu(\pi)$, $\pi \in \mathfrak{S}_\mathrm{n}$. Falls Q eine Verteilung auf $\mathfrak{S}_\mathrm{n}$ mit $Q(\pi) < Q(\mathrm{id})$ für ein $\pi \in \mathfrak{S}_\mathrm{n}$ ist, kann daher unmöglich ein μ existieren mit $Q = Q_\mu$, d.h. φ ist nicht surjektiv.

Wir zeigen nun die Surjektivität von $\widehat{\varphi}$. Sei hierzu $Q \in M$. Es wird als Nächstes eine Verteilung μ auf $\{0, \ldots, n\}$ konstruiert mit $Q = Q_\mu$: Nach (2.10) und der Voraussetzung $L(\pi) = L(\tau) \Rightarrow Q(\pi) = Q(\tau)$, $\pi, \tau \in \mathfrak{S}_\mathrm{n}$ reicht es offenbar, dass μ

die Gleichungen

$$Q(\pi_l) = \sum_{j=0}^{l} \mu(n-j)\frac{j!}{n!}, \quad l = 1, \ldots, n \tag{2.12}$$

erfüllt. (2.12) wird im Folgenden gelöst: Aus

$$Q(\pi_1) = \frac{1}{n!}(\mu(n-1) + \mu(n))$$

folgt $\mu(n-1)$ nach Vorgabe von $\mu(n)$, wobei zusätzlich

$$0 \leq \mu(n) \leq n!Q(\pi_1) \tag{2.13}$$

erfüllt sei. (2.13) garantiert $\mu(n-1) \geq 0$. Wir definieren $\mu(n-l)$, $l = 2, \ldots, n$ genau wie in Teil (ii) iterativ mittels (2.11) mit Q anstelle von Q_μ. Es folgt dann (2.12) mit einer Funktion $\mu : \{0, \ldots, n\} \to \mathbb{R}$, $\mu(n-1), \mu(n) \geq 0$. Als Nächstes wird gezeigt, dass μ eine Verteilung ist: Mit (2.12) erhalten wir

$$Q(\pi_l) = Q(\pi_{l-1}) + \mu(n-l)\frac{l!}{n!}, \quad l = 2, \ldots, n.$$

Die Voraussetzung $L(\pi) \leq L(\tau) \Rightarrow Q(\pi) \leq Q(\tau)$ ergibt unmittelbar

$$\mu(n-l) \geq 0,\ l = 2, \ldots, n.$$

Wir haben hiermit ein endliches Maß μ konstruiert, das (2.12) erfüllt. Offenbar bleiben allgemeiner die Umformungsschritte in (2.10) mit einem endlichen Maß μ anstelle einer Verteilung μ immer noch gültig. Insbesondere folgt zusammen mit (2.12) daraus

$$\sum_{m=0}^{n} \mu(m)Q_m(\pi_l) = \sum_{j=0}^{l} \mu(n-j)\frac{j!}{n!} = Q(\pi_l), \quad l = 1, \ldots, n.$$

Hieraus erhalten wir

$$\begin{aligned}
\sum_{m=0}^{n} \mu(m) &= \sum_{m=0}^{n} \mu(m)Q_m(\mathfrak{S}_\mathrm{n}) = \sum_{m=0}^{n} \mu(m) \sum_{l=1}^{n} |\{L = l\}|Q_m(\pi_l) \\
&= \sum_{l=1}^{n} |\{L = l\}| \sum_{m=0}^{n} \mu(m)Q_m(\pi_l) = \sum_{l=1}^{n} |\{L = l\}|Q(\pi_l) \\
&= Q(\mathfrak{S}_\mathrm{n}) = 1,
\end{aligned}$$

weshalb μ ist eine Verteilung ist. Insgesamt haben wir

$$M \subseteq \varphi(\{\text{Verteilungen auf } \{0, \ldots, n\}\})$$

gezeigt. Da umgekehrt wegen (2.3) und (2.5) $Q_\mu \in M$ für alle Verteilungen μ auf $\{0, \ldots, n\}$ gilt, ist somit die Surjektivität von $\widehat{\varphi}$ gezeigt.

(iv): Wegen $Q_\mu(\pi_1) = \frac{1}{n!}(\mu(n-1) + \mu(n))$ folgt $\mu(n) \leq n!Q_\mu(\pi_1)$, so dass ψ wohldefiniert ist. ψ ist wegen (ii) injektiv. Andererseits ist ψ auch surjektiv, da zu gegebenen (Q, x) immer ein μ konstruiert werden kann mit $Q_\mu = Q$ und $\mu(n) = x$, vgl. hierzu Teil (iii). □

Unter Beachtung der Interpretation von $\mu_1 \sharp \ldots \sharp \mu_k$ als Verteilung der besetzten Zellen in dem über Satz 2.1.4 beschriebenen Experiment, erhalten wir eine erste Aussage bzgl. des Variationsabstandes von $Q_{\mu_1 \sharp \ldots \sharp \mu_k}$ zur Gleichverteilung U.

Lemma 2.1.6. *Seien $n \in \mathbb{N}$ und $\epsilon \in (0,1)$ fixiert. Ferner sei eine Familie $(\mu_i)_{i \in \mathbb{N}}$ von Verteilungen auf $\{0, \ldots, n\}$ gegeben, wobei $\mu_i(0) \leq 1 - \epsilon$ für unendlich viele $i \in \mathbb{N}$ gelte. Dann gilt*

(i) $\mu_1 \sharp \ldots \sharp \mu_k(j) \to \delta_n(j), \quad k \to \infty, \quad j = 0, \ldots, n,$

(ii) $\|Q_{\mu_1 \sharp \ldots \sharp \mu_k} - U\| \to 0, \quad k \to \infty.$

Beweis. (i) Es reicht offenbar $\mu_1 \sharp \ldots \sharp \mu_k(n) \to 1$, $k \to \infty$ zu zeigen, wobei wir o.E. $\mu_i(0) \leq 1 - \epsilon$, $\forall i \in \mathbb{N}$ annehmen dürfen. Die Wahrscheinlichkeit α, innerhalb von $k = n$ Schritten alle n Zellen aufzufüllen, beträgt mindestens $(\frac{\epsilon}{n})^n$, da es hierzu ausreicht, jeweils in dem j-ten Experiment ($j = 1, \ldots, n$) *mindestens eine* Kugel zu wählen ($1 - \mu_i(0) \geq \epsilon$) *und* diese in die j-te Zelle zu legen ($\frac{1}{n}$). Falls es uns nicht gelungen ist, in den ersten n Experimenten alle Zellen zu besetzen, so werden wir es in den nächsten n Experimenten wieder versuchen, et.c. *(geometric trials argument)*. Hieraus resultiert wegen der Unabhängigkeit der Versuche die Abschätzung

$$\mu_1 \sharp \ldots \sharp \mu_{nr}(n) \geq \sum_{u=0}^{r-1} \alpha(1-\alpha)^u \to 1, \quad r \to \infty,$$

denn

$$\alpha \geq \left(\frac{\epsilon}{n}\right)^n > 0.$$

Da $k \mapsto \mu_1 \sharp \ldots \sharp \mu_k(n)$ offenbar monoton wachsend ist, folgt schließlich

$$\mu_1 \sharp \ldots \sharp \mu_k(n) \to 1, \, k \to \infty.$$

(ii) Dies ist eine einfache Folgerung aus (i). Es gilt

$$\begin{aligned} \|Q_{\mu_1 \sharp \ldots \sharp \mu_k} - U\| &= \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} |Q_{\mu_1 \sharp \ldots \sharp \mu_k}(\pi) - U(\pi)| \\ &= \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} \left| \sum_{m=0}^{n} \mu_1 \sharp \ldots \sharp \mu_k(m) Q_m(\pi) - \frac{1}{n!} \right| \\ &\to \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} \left| \sum_{m=0}^{n} \delta_n(m) Q_m(\pi) - \frac{1}{n!} \right| = 0, \quad k \to \infty. \end{aligned}$$

In der letzten Zeile haben wir $U = Q_n$, (i) und die Stetigkeit der Funktion

$$(x_0, \ldots, x_n) \mapsto \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} \left| \sum_{m=0}^{n} x_m Q_m(\pi) - \frac{1}{n!} \right|$$

auf $[0,1]^{n+1}$ ausgenutzt. □

Für den TmTRS, $m \geq 1$ mit $\mu_i \stackrel{\text{def}}{=} \delta_m$ für alle $i \in \mathbb{N}$ ergibt sich somit auch ohne Kenntnis der allgemeinen Theorie DMK, wie zu erwarten, sofort die Konvergenz gegen die Gleichverteilung. Allerdings wird überhaupt nichts über die Geschwindigkeit der Konvergenz ausgesagt, wofür wir uns hier besonders interessieren: Es wird sich,

wie schon erwähnt, herausstellen, dass für große n ein *Cutoff-Effekt* existiert. Bis zu einer bestimmten Anzahl von Mischungen ist das Deck von der Gleichverteilung auf $\mathfrak{S}_n$ noch weit entfernt, geht bei weiterem Mischen dann aber sehr schnell in diese über (approximativ). Nur wegen dieses Effektes ist eine Aussage wie „Das Deck benötigt $n \log n$ T1TRS, um durchmischt zu werden“ überhaupt sinnvoll. Ergäbe sich nur eine sehr langsame, kontinuierliche Konvergenz gegen die Gleichverteilung, wären wir gezwungen, die Prägnanz obiger Aussage durch mathematisch formale Zusätze zu verunschönen.

Wir werden im Korollar 2.1.8 eine geschlossene Formel für die Verteilung k nacheinander ausgeführter T1TRS angeben. Hierzu zunächst folgender Satz.

Satz 2.1.7. *Sei μ eine Verteilung auf $\{0, \dots, n\}$. Dann gilt*

$$Q_\mu(\pi) = \frac{1}{n!} \sum_{u=0}^{L(\pi)} u! \mu(n-u), \quad \pi \in \mathfrak{S}_n \,. \tag{2.14}$$

Ferner erhalten wir für $1 \le l \le n$

$$Q_\mu(L \ge l) = \sum_{u=0}^{l-1} \mu(n-u) \frac{u!}{l!} + \sum_{u \ge l} \mu(n-u). \tag{2.15}$$

Es ist Q_μ bedingt auf $\{L = l\}$ die Gleichverteilung auf dieser Menge.

Beweis. Es gilt (vgl. (2.10))

$$\begin{aligned} Q_\mu(\pi) &= \sum_{m=0}^{n} \mu(m) Q_m(\pi) \\ &= \sum_{m=0}^{n} \mu(m) \frac{(n-m)!}{n!} \mathbb{1}_{\{n-L(\pi),\dots,n\}}(m), \quad u \overset{\text{def}}{=} n-m \\ &= \sum_{u=0}^{L(\pi)} \mu(n-u) \frac{u!}{n!}. \end{aligned}$$

Dies zeigt (2.14) und impliziert ferner (2.15) für $l = n$, da $\{L \ge n\} = \{\mathrm{id}\}$.

Sei $f(l)$ definiert als die rechte Seite von (2.15). Es reicht offenbar

$$f(l) - f(l+1) = Q_\mu(L = l), \quad 1 \le l \le n-1 \tag{2.16}$$

zu zeigen. Es ist

$$\begin{aligned} & f(l) - f(l+1) \\ = \; & \sum_{u=0}^{l-1} \mu(n-u) \frac{u!}{l!} + \sum_{u \ge l} \mu(n-u) - \sum_{u=0}^{l} \mu(n-u) \frac{u!}{l!(l+1)} - \sum_{u \ge l+1} \mu(n-u) \\ = \; & \sum_{u=0}^{l-1} \mu(n-u) \frac{u!}{l!} \left(1 - \frac{1}{l+1}\right) - \frac{\mu(n-l)}{l+1} + \mu(n-l) \\ = \; & \sum_{u=0}^{l} \mu(n-u) \frac{u!}{l!} \left(1 - \frac{1}{l+1}\right), \quad l = 1, \dots, n-1. \end{aligned} \tag{2.17}$$

Andererseits gilt wegen (2.4)

$$\begin{aligned} |\{L=l\}| &= |\{L\geq l\}| - |\{L\geq l+1\}| \\ &= \frac{n!}{l!}\left(1-\frac{1}{l+1}\right), \quad l=1,\ldots,n-1, \end{aligned} \tag{2.18}$$

woraus unter Beachtung von (siehe (2.14))

$$Q_\mu(L=l) = |\{L=l\}| \sum_{u=0}^{l} \mu(n-u)\frac{u!}{n!}$$

mit (2.17) schließlich (2.16) folgt. Da (2.14) offenbar von π nur über $L(\pi)$ abhängt, ist (2.14) bedingt auf $\{L=l\}$ die Gleichverteilung auf dieser Menge. □

Korollar 2.1.8. *Nach k unabhängigen T1TRS ist die Wahrscheinlichkeit einer Permutation π durch*

$$Q_1^{*k}(\pi) = \frac{1}{n!}\sum_{u=0}^{L(\pi)} u! P_k^n(u) \tag{2.19}$$

gegeben, wobei $P_k^n(u)$ die Wahrscheinlichkeit ist, genau u unbesetzte Zellen bei k zufälligen Besetzungen einer n Zellenanordnung zu erhalten. Es gilt

$$P_k^n(u) = \sum_{\nu=u}^{n} (-1)^{\nu-u}\binom{n}{\nu}\binom{\nu}{u}\left(1-\frac{\nu}{n}\right)^k. \tag{2.20}$$

Ferner erhalten wir für $1\leq l\leq n$

$$Q_1^{*k}(L\geq l) = \sum_{u=0}^{l-1} P_k^n(u)\frac{u!}{l!} + \sum_{u\geq l} P_k^n(u). \tag{2.21}$$

Letzteres ist die Wahrscheinlichkeit der Existenz einer n enthaltenden, aufsteigenden Sequenz, die mindestens die Länge l besitzt, nach k T1TRS Mischvorgängen aufzufinden. Schließlich ist Q_1^{*k} bedingt auf $\{L=l\}$ die Gleichverteilung auf dieser Menge.

Beweis. Mit

$$\mu \stackrel{\text{def}}{=} \overbrace{\delta_1 \sharp \ldots \sharp \delta_1}^{k\text{-mal}}$$

gilt offenbar

$$P_k^n(u) = \mu(n-u), \quad u=0,\ldots,n.$$

Dies folgt direkt aus der Definition der Verknüpfung $\sharp$. Nach Satz 2.1.4 gilt $Q_1^{*k} = Q_\mu$. Wegen Satz 2.1.7 bleibt daher nur noch (2.20) nachzuweisen. Hierzu werden die Zellen der n Zellenanordnung mit $1,\ldots,n$ durchnummeriert, und A_i, $1\leq i\leq n$ bezeichne das Ereignis, dass Zelle i nach k Besetzungen *unbesetzt* ist. Wir definieren weiter

$$\begin{aligned} s_0 &\stackrel{\text{def}}{=} 1, \\ s_\nu &\stackrel{\text{def}}{=} \sum_{1\leq i_1<\ldots<i_\nu\leq n} P(A_{i_1}\cap\ldots\cap A_{i_\nu}), \quad 1\leq\nu\leq n. \end{aligned}$$

Wegen

$$P(A_{i_1} \cap \ldots \cap A_{i_\nu}) = \left(1 - \frac{\nu}{n}\right)^k, \quad 1 \leq \nu \leq n, \quad 1 \leq i_1 < \ldots < i_\nu \leq n$$

folgt

$$s_\nu = \binom{n}{\nu} \left(1 - \frac{\nu}{n}\right)^k, \quad \nu = 0, \ldots, n.$$

Hieraus erhalten wir durch Anwendung der Siebformel und Übergang zum Komplementärereignis die Wahrscheinlichkeit, ausschließlich besetzte Zellen vorzufinden, nämlich

$$P_k^n(0) = 1 - P\left(\bigcup_{i=1}^n A_i\right) = \sum_{\nu=0}^n (-1)^\nu s_\nu = \sum_{\nu=0}^n (-1)^\nu \binom{n}{\nu} \left(1 - \frac{\nu}{n}\right)^k. \tag{2.22}$$

Das Ereignis, *genau* u unbesetzte Zellen vorzufinden, ergibt sich, indem wir zuerst diese u Zellen aus der n Zellenanordnung aussuchen und anschließend die verbleibenden $n - u$ Zellen mit den k Kugeln vollständig ausfüllen. In Formeln wird dies wegen

$$P_k^{n-u}(0) = \frac{|\{\text{vollständig gefüllte } n-u \text{ Zellenbesetzungen}\}|}{|\{n-u \text{ Zellenbesetzungen}\}|},$$

wobei der Nenner gleich $(n-u)^k$ ist, zu

$$\begin{aligned}
P_k^n(u) &= \binom{n}{u} \frac{(n-u)^k}{n^k} \cdot P_k^{n-u}(0) \qquad (2.23)\\
&= \sum_{\nu=0}^{n-u} (-1)^\nu \binom{n}{u} \binom{n-u}{\nu} \left(1 - \frac{u+\nu}{n}\right)^k, \quad \widetilde{\nu} \stackrel{\text{def}}{=} \nu + u\\
&= \sum_{\widetilde{\nu}=u}^{n} (-1)^{\widetilde{\nu}-u} \binom{n}{u} \binom{n-u}{\widetilde{\nu}-u} \left(1 - \frac{\widetilde{\nu}}{n}\right)^k, \quad \nu \stackrel{\text{def}}{=} \widetilde{\nu}\\
&= \sum_{\nu=u}^{n} (-1)^{\nu-u} \frac{n!}{(n-u)!u!} \frac{(n-u)!}{(n-\nu)!(\nu-u)!} \frac{\nu!}{\nu!} \left(1 - \frac{\nu}{n}\right)^k\\
&= \sum_{\nu=u}^{n} (-1)^{\nu-u} \binom{n}{\nu} \binom{\nu}{u} \left(1 - \frac{\nu}{n}\right)^k,
\end{aligned}$$

womit (2.20) und damit das ganze Korollar bewiesen ist. □

Obiges lässt sich noch verallgemeinern. Hierzu konstatieren wir an dieser Stelle ein allgemeineres Resultat, dem der nächste Abschnitt gewidmet wird.

Satz 2.1.9. *Es gilt*

$$Q_{m_k} * Q_{m_{k-1}} * Q_{m_1}(\pi) = \frac{1}{n!} \sum_{u=0}^{L(\pi)} u! P_{m_1,\ldots,m_k}(u), \tag{2.24}$$

wobei mit

$$\begin{aligned} P_{m_1,\ldots,m_k}(u) &\overset{\text{def}}{=} \delta_{m_1} \sharp \ldots \sharp \delta_{m_k}(n-u) \\ &= \sum_{\nu=u}^{n} (-1)^{\nu-u} \binom{n}{\nu} \binom{\nu}{u} \prod_{j=1}^{k} \frac{\binom{n-\nu}{m_j}}{\binom{n}{m_j}} \end{aligned} \tag{2.25}$$

die Verteilung der unbesetzten Zellen einer n Zellenanordnung gemeint ist, wenn in diese sukzessive zuerst m_1 bis schließlich m_k Kugeln unabhängig voneinander rein zufällig hineingelegt werden. Innerhalb der einzelnen Schritte sind dabei keine Mehrfachbesetzungen von Zellen erlaubt.

Wegen Satz 2.1.7 muss lediglich (2.25) nachgewiesen werden. Es handelt sich hierbei um das sogenannte Komiteeproblem. Damit haben wir insbesondere eine geschlossene Darstellung für die Verteilung eines Kartendecks nach k TmTRS gefunden. Allerdings sollte an dieser Stelle eher gesagt werden, dass das Problem auf das Komiteeproblem zurückgeführt wurde. Wir machen darauf aufmerksam, dass (2.20) als Spezialfall von (2.25) ($m_1 = \ldots = m_k = 1$) auch durch diesen bewiesen ist. Dennoch wird ein Beweis diese Spezialfalls explizit in Korollar 2.1.8 angegeben, da zum einen eine elementarere Vorgehensweise als in Abschnitt 2.2 aufgezeigt werden soll und zum anderen Zwischenergebnisse dieses Beweises im Beweis vom Satz 2.3.1 benötigt werden.

Schließlich wird noch ein Lemma notiert, das später im Zusammenhang mit dem *Riffle-Shuffle* eine recht interessante Aussage ermöglicht (siehe Satz 2.5.18).

Lemma 2.1.10. *Für die zuvor definierte Verknüpfung $\sharp$ gilt für $0 < p_1, p_2 < 1$*

$$\mathrm{Bin}(n, p_1) \sharp \mathrm{Bin}(n, p_2) = \mathrm{Bin}(n, 1 - (1-p_1)(1-p_2)), \tag{2.26}$$

wobei mit Bin *die Binomialverteilung gemeint ist.*

Beweis. $I_j^{(k)}$, $j = 1, \ldots, n$, $k = 1, 2$ seien Indikatorzufallsvariablen mit

$$I_j^{(k)} = \begin{cases} 0 & \text{falls im } k\text{-ten Schritt } \textit{keine} \text{ Kugel in Zelle } j \text{ gelegt wird,} \\ 1 & \text{sonst.} \end{cases}$$

Diese sind per Voraussetzung unabhängig. $\sum_{j=1}^{n} \max\left\{I_j^{(1)}, I_j^{(2)}\right\}$ ist offensichtlich $\mathrm{Bin}(n, p_1) \sharp \mathrm{Bin}(n, p_2)$ verteilt. Wegen

$$P\left(\max\{I_j^{(1)}, I_j^{(2)}\} = 0\right) = P\left(I_j^{(1)} = I_j^{(2)} = 0\right) = (1-p_1)(1-p_2)$$

folgt die Behauptung. □

2.2 Das Komiteeproblem

Wie weiter oben bereits angedeutet, ist das Ziel dieses Abschnitts, (2.25) zu beweisen.

Als Nächstes werden einige grundlegende Eigenschaften über Hilberträume wiederholt.

Intermezzo über Hilberträume
Ein komplexer Vektorraum H mit Skalarprodukt $\langle \cdot, \cdot \rangle$ heißt Hilbertraum, falls er als normierter Vektorraum vollständig ist, also jede Cauchy Folge konvergiert. Die Norm wird hierbei wie gewohnt von dem Skalarprodukt induziert.

Eine Menge $\{u_\alpha : \alpha \in A\} \subseteq H$ wird orthonormale Menge in H genannt, falls $\langle u_\alpha, u_\beta \rangle = \delta_{\alpha,\beta}$, wobei A eine beliebige Indexmenge bezeichne und mit δ das Kronecker Symbol gemeint ist. Ist der von obiger Teilmenge erzeugte Vektorraum sogar dicht in H, so heißt diese ein vollständiges Orthonormalsystem (VONS) von H.

Der $\mathbb{C}$ Vektorraum $H \stackrel{\text{def}}{=} L^2([-\pi,\pi]^d)$ aller quadratisch integrierbaren komplexen Funktionen von $[-\pi,\pi]^d$ nach $\mathbb{C}$ ist ein Hilbertraum. $[-\pi,\pi]^d$ wird hierbei kanonisch als Maßraum betrachtet, indem wir $\mathfrak{B}^d_{[-\pi,\pi]^d}$ (Borelsche Spur σ-Algebra) als σ-Algebra und $\lambda^d_{[-\pi,\pi]^d}$ (Lebesgue Maß) als Maß wählen. Als Skalarprodukt wird

$$\langle f, g \rangle \stackrel{\text{def}}{=} \frac{1}{(2\pi)^d} \int f(t)\overline{g(t)} \lambda^d_{[-\pi,\pi]^d}(dt)$$

gewählt. H wird wie üblich als Quotientenvektorraum modulo $\lambda^d_{[-\pi,\pi]^d}$-f.ü verschwindende Funktionen betrachtet (sonst implizierte $\langle f, f \rangle$ nicht unbedingt $f = 0$). Weiter ist bekannt, dass

$$u_{n_1,\ldots,n_d}(t_1,\ldots,t_d) \stackrel{\text{def}}{=} \exp(i \sum_{j=1}^{d} n_j t_j), \quad n_1,\ldots,n_d \in \mathbb{Z} \tag{2.27}$$

ein VONS für H bildet (vgl. Rudin [17, Kapitel 4]). Das ist die Grundlage einer jeden diskreten Fourieranalyse.

Für das Folgende entscheidend ist allerdings nur, dass obiges System orthonormal ist, was leicht nachzurechnen ist. Dennoch sollte der übliche abstrakte Hintergrund von (2.27) noch kurz erwähnt werden, was hiermit geschehen ist.

Zum Komiteeproblem
Die folgenden Ausführungen orientieren sich an Holst [13]. N Personen sollen derart auf R Komitees verteilt werden, dass zu jedem Komitee genau $n_j, j = 1,\ldots,R$ Personen gehören. Hierbei kann eine Person auch zu mehreren oder gar keinem Komitee gehören. Bis auf die obige Bedingung an die Anzahl n_j der Komiteemitglieder erfolgt die Zuordnung rein zufällig. Die Bestimmung der Verteilung der Anzahl der Personen, die zu *keinem* Komitee gehören, wird das Komiteeproblem genannt. Offensichtlich können wir das Komiteeproblem mit einer N Zellenanordnung identifizieren, in die wir sukzessive $n_1, n_2, \ldots, n_R$ Kugeln legen, wobei in jedem Schritt keine Mehrfachbesetzungen erlaubt sind. Im ersten Schritt wird die Zellenanordnung rein zufällig mit genau n_1 Kugeln bestückt. In unserer Analogiebetrachtung sind dies die n_1 Personen, die dem ersten Komitee angehören. Im zweiten Schritt werden unabhängig vom ersten entsprechend n_2 Kugeln verteilt u.s.w. Auf diese Weise wird der Bezug zum vorherigen Abschnitt klar. Obiges wird im Folgenden nun leicht verallgemeinert formalisiert.

Gegeben sei eine $R \times N$ Matrix mit genau n_j Einsen und sonst Nullen in der j-ten Zeile. Jede Zeilenkonstellation wird rein zufällig unabhängig von jeder anderen z.B. mit einem Urnenexperiment via Ziehen ohne Zurücklegen ermittelt. Sei

$$\mathbf{I} = (I_{j,k})_{\substack{j=1,\ldots,R \\ k=1,\ldots,N}} \tag{2.28}$$

obige $R \times N$ Matrix ($\mathbf{I}$ ist eine Zufallsvariable). Für eine gegebene Funktion

$$f : \{0,1\}^R \to \mathbb{R}$$

definieren wir die Zufallsvariable

$$N_0 \stackrel{\text{def}}{=} \sum_{k=1}^{N} f(I_{1k}, I_{2k}, \ldots, I_{Rk}). \tag{2.29}$$

Das Komiteeproblem ergibt sich mit

$$f(x_1, \ldots, x_R) \stackrel{\text{def}}{=} \mathbb{1}_{\{x_1=\ldots=x_R=0\}}. \tag{2.30}$$

Zufallsvariablen N_0 für allgemeineres f werden in [13] studiert. Als technisches Hilfmittel werden unabhängig, identisch Bernoulli verteilte Zufallsvariablen

$$\mathbf{X} \stackrel{\text{def}}{=} (X_{jk})_{\substack{j=1,\ldots,R \\ k=1,\ldots,N}}, \quad p_j \stackrel{\text{def}}{=} P(X_{jk}=1) = 1 - P(X_{jk}=0) = 1 - q_j \in (0,1) \tag{2.31}$$

eingeführt, wobei q_j wie üblich über letztere Gleichung definiert wird. Es gilt offensichtlich

$$P^{\mathbf{I}} = P^{\mathbf{X}|\sum_{k=1}^{N} X_{jk}=n_j,\, j=1,\ldots,R}. \tag{2.32}$$

Obige Gleichung gilt unabhängig von der Wahl der $p_i \in (0,1)$. Daher bildet die Abbildung $x \mapsto \left(\sum_{k=1}^{N} x_{jk}\right)_{1\le j\le R}$ im Sinne von [4, S. 34] eine *suffiziente Statistik* für die Verteilungsfamilie $\left(P^{\mathbf{X}}\right)_{(p_1,\ldots,p_R)\in(0,1)^R}$. Aus (2.32) folgt für jede reelle Funktion g auf $\{0,1\}^{R\times N}$

$$E\left(e^{ivg(\mathbf{I})}\right) = E\left(e^{ivg(\mathbf{X})} \,\middle|\, \sum_{k=1}^{N} X_{jk} = n_j,\ j=1,\ldots,R\right), \quad v \in \mathbb{C}. \tag{2.33}$$

Um das VONS (2.27) nutzen zu können, schreiben wir

$$\begin{aligned}
&E\left(e^{ivg(\mathbf{X})+i\sum_{j=1}^{R} u_j \sum_{k=1}^{N} X_{jk}}\right) \\
&\quad= \sum_{n_1',\ldots,n_R'} E\left(e^{ivg(\mathbf{X})+i\sum_{j=1}^{R} u_j \sum_{k=1}^{N} X_{jk}} \,\middle|\, \sum_{k=1}^{N} X_{jk} = n_j', \forall j\right) \cdot P\left(\sum_{k=1}^{N} X_{jk} = n_j', \forall j\right) \\
&\quad= \sum_{n_1',\ldots,n_R'} E\left(e^{ivg(\mathbf{X})} \,\middle|\, \sum_{k=1}^{N} X_{jk} = n_j', \forall j\right) \cdot P\left(\sum_{k=1}^{N} X_{jk} = n_j', \forall j\right) \cdot e^{i\sum_{j=1}^{R} u_j n_j'}.
\end{aligned} \tag{2.34}$$

Wir definieren zur Abkürzung

$$h(n_1', \ldots, n_R') \stackrel{\text{def}}{=} E\left(e^{ivg(\mathbf{X})} \,\middle|\, \sum_{k=1}^{N} X_{jk} = n_j', \forall j\right) \cdot P\left(\sum_{k=1}^{N} X_{jk} = n_j', \forall j\right).$$

Dann gilt mit der abkürzenden Symbolik $du \stackrel{\text{def}}{=} du_1 \ldots du_R$

$$(2\pi)^{-R} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} E\left(e^{ivg(\mathbf{X})+i\sum_{j=1}^{R} u_j \sum_{k=1}^{N} X_{jk}}\right) e^{-i\sum_{j=1}^{R} u_j n_j}\, du \tag{2.35}$$

$$= \sum_{n'_1,\dots n'_R} h(n'_1,\dots n'_R) \cdot (2\pi)^{-R} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} e^{i\sum_{j=1}^{R} u_j n'_j} e^{-i\sum_{j=1}^{R} u_j n_j}\, du \tag{2.36}$$

$$= \sum_{n'_1,\dots n'_R} h(n'_1,\dots n'_R) \cdot \delta_{(n_1,\dots,n_R),(n'_1,\dots,n'_R)} \tag{2.37}$$

$$= h(n_1,\dots,n_R) \tag{2.38}$$

$$= E\left(e^{ivg(\mathbf{I})}\right) \cdot P\left(\sum_{k=1}^{N} X_{jk} = n_j, \forall j\right). \tag{2.39}$$

Der Übergang von (2.35) zu (2.36) folgt aus (2.34). (2.37) ergibt sich mittels des VONS (2.27) aus (2.36). Schließlich erhalten wir (2.39) wegen (2.33) aus (2.38). Zusammenfassend ist

$$\begin{aligned} E\left(e^{ivg(\mathbf{I})}\right) &= (2\pi)^{-R} \left(P\left(\sum_{k=1}^{N} X_{jk} = n_j, \forall j\right)\right)^{-1} \cdot \\ &\quad \cdot \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} E\left(e^{ivg(\mathbf{X})+i\sum_{j=1}^{R} u_j \sum_{k=1}^{N} X_{jk}}\right) e^{-i\sum_{j=1}^{R} u_j n_j}\, du \end{aligned} \tag{2.40}$$

gezeigt. Wir haben hiermit eine spezielle Darstellung der charakteristischen Funktion von $g(\mathbf{I})$ gefunden. Im Folgenden werde g so gewählt, dass es sich um das Komiteeproblem handelt, d.h.

$$g(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{k=1}^{N} f(x_{1k},\dots,x_{Rk}) = \sum_{k=1}^{N} \mathbb{1}_{\{x_{1k}=\dots=x_{Rk}=0\}}, \quad \mathbf{x} \in \{0,1\}^{R\times N}. \tag{2.41}$$

Wie schon erwähnt, sind wir an der Verteilung von $N_0 = g(\mathbf{I})$ interessiert, die wir indirekt über die Bestimmung von $E\left(e^{ivN_0}\right)$ ermitteln werden. Der Vorteil dieser Vorgehensweise liegt in der Multiplikativität des Erwartungswertes von Produkten unabhängiger Zufallsvariablen (hier die Komponenten von $\mathbf{X}$) und der Tatsache, dass wir die Erwartungswerte der entstehenden Faktoren leicht angeben können. Es folgt nun eine längere Rechnung:

$$\begin{aligned} E\left(e^{ivN_0}\right) &= (2\pi)^{-R} \left(P\left(\sum_{k=1}^{N} X_{jk} = n_j, \forall j\right)\right)^{-1} \cdot \\ &\quad \cdot \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} E\left(e^{\sum_{k=1}^{N} iv\mathbb{1}_{\{X_{1k}=\dots=X_{Rk}=0\}} + i\sum_{j=1}^{R} u_j \sum_{k=1}^{N} X_{jk}}\right) \cdot e^{-i\sum_{j=1}^{R} u_j n_j}\, du. \end{aligned} \tag{2.42}$$

Da insbesondere die Spalten von $\mathbf{X}$ unabhängig und identisch verteilt sind, ist dies gleich

$$(2\pi)^{-R} \left(P\left(\sum_{k=1}^{N} X_{jk} = n_j, \forall j\right)\right)^{-1} \cdot \tag{2.43}$$

$$\cdot \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \left[E\left(e^{iv\mathbb{1}\{X_{11}=\ldots=X_{R1}=0\}+i\sum_{j=1}^{R} u_j X_{j1}} \right) \right]^N \cdot e^{-i\sum_{j=1}^{R} u_j n_j}\, du.$$

Es gilt

(i) $P\left(\sum_{k=1}^{N} X_{jk} = n_j, \forall j\right) = \prod_{j=1}^{R} \binom{N}{n_j} p_j^{n_j} q_j^{N-n_j}$

(ii) $E\left(e^{iv\mathbb{1}\{X_{11}=\ldots=X_{R1}=0\}+i\sum_{j=1}^{R} u_j X_{j1}}\right) = \pi_0\left(e^{iv}-1\right) + \prod_{j=1}^{R}\left(p_j e^{iu_j} + q_j\right), \quad \pi_0 \stackrel{\text{def}}{=} \prod_{j=1}^{R} q_j.$

Setzen wir dies in (2.43) ein, so ergibt sich

$$E\left(e^{ivN_0}\right) = (2\pi)^{-R}\left(\prod_{j=1}^{R}\binom{N}{n_j} p_j^{n_j} q_j^{N-n_j}\right)^{-1} \cdot \tag{2.44}$$

$$\cdot \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \sum_{m=0}^{N} \binom{N}{m} \pi_0^m \left(e^{iv}-1\right)^m \prod_{j=1}^{R}\left[\left(p_j e^{iu_j}+q_j\right)^{N-m} e^{-iu_j n_j}\right] du$$

$$= \sum_{m=0}^{N}\binom{N}{m}\pi_0^m\left(e^{iv}-1\right)^m \left(\prod_{j=1}^{R}\binom{N}{n_j} p_j^{n_j} q_j^{N-n_j}\right)^{-1} \cdot \tag{2.45}$$

$$\cdot (2\pi)^{-R} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \prod_{j=1}^{R} \sum_{k=0}^{N-m}\left[\binom{N-m}{k} p_j^k e^{iu_j k} q_j^{N-m-k} e^{-iu_j n_j}\right] du$$

$$= \sum_{m=0}^{N}\binom{N}{m}\pi_0^m\left(e^{iv}-1\right)^m \left(\prod_{j=1}^{R}\binom{N}{n_j} p_j^{n_j} q_j^{N-n_j}\right)^{-1} \cdot \tag{2.46}$$

$$\cdot \prod_{j=1}^{R}\binom{N-m}{n_j} p_j^{n_j} q_j^{N-m-n_j}$$

$$= \sum_{m=0}^{N}\binom{N}{m}\left(e^{iv}-1\right)^m \prod_{j=1}^{R}\left[\binom{N-m}{n_j} \Big/ \binom{N}{n_j}\right] \tag{2.47}$$

$$= \sum_{k=0}^{N}\sum_{m=k}^{N}\binom{N}{m}\binom{m}{k}(-1)^{m-k} \prod_{j=1}^{R}\left[\binom{N-m}{n_j} \Big/ \binom{N}{n_j}\right] e^{ikv}. \tag{2.48}$$

Das Integral verschwindet von (2.45) nach (2.46) wegen der schon mehrfach erwähnten Orthonormalität von (2.27). Von (2.46) nach (2.47) wurden nur einige Kürzungen vorgenommen. Durch eine Anwendung des Binomischen Theorems erhalten wir in (2.48) schließlich die Fourierdarstellung von $E\left(e^{ivN_0}\right)$ als Funktion von v. Aus der Eindeutigkeit dieser Darstellung und $E(e^{ivN_0}) = \sum_{k=0}^{N} P(N_0 = k) e^{ivk}$ folgt sofort

$$P(N_0 = y) = \sum_{m=y}^{N} (-1)^{m-y}\binom{N}{m}\binom{m}{y}\prod_{j=1}^{R}\left[\binom{N-m}{n_j} \Big/ \binom{N}{n_j}\right]. \tag{2.49}$$

Hiermit ist insbesondere (2.25) bewiesen. Indem $v \stackrel{\text{def}}{=} -i\log(s)$, $s \in (0,1]$ gesetzt wird, bekommt man weiter die erzeugende Funktion $s \mapsto E(s^{N_0})$. Durch iteratives linksseitiges Differenzieren an der Stelle 1 erhalten wir mit der Notation

$$x^{(k)} \stackrel{\text{def}}{=} \prod_{i=0}^{k-1}(x-i), \quad x \in \mathbb{R}, \quad k = 1, 2, \ldots \tag{2.50}$$

die faktoriellen Momente

$$E(N_0^{(k)}) = N^{(k)} \prod_{j=1}^{R} \left[(N-n_j)^{(k)} \Big/ N^{(k)} \right], \quad k = 1, 2, \ldots, \tag{2.51}$$

d.h. insbesondere

$$E(N_0) = N \prod_{j=1}^{R} \frac{N-n_j}{N} = N \prod_{j=1}^{R} \left(1 - \frac{n_j}{N}\right). \tag{2.52}$$

Die Rechnung hierzu ist kanonisch und kann in [13] betrachtet werden.

Es sei zum Schluss noch darauf hingewiesen, dass für fixiertes $N \in \mathbb{N}$ die Konvergenz

$$P(N_0 \geq 1) \leq E(N_0) \to 0 \tag{2.53}$$

besteht, falls $R \to \infty$ und $n_j > 0$ für unendlich viele $j \in \mathbb{N}$ gilt. Dies folgt unmittelbar aus (2.52), da wir dann

$$\left(1 - \frac{n_j}{N}\right) \leq \left(1 - \frac{1}{N}\right) < 1 \quad \text{für unendlich viele } j \in \mathbb{N}$$

haben und daher

$$E(N_0) \leq N \left(1 - \frac{1}{N}\right)^{a_R}$$

für eine Folge $(a_R)_{R\in\mathbb{N}}$ mit $a_R \to \infty$, $R \to \infty$ gilt, woraus (2.53) folgt. Bei sukzessiven Kugelbesetzungen mit jeweils mindestens einer Kugel ist folglich irgendwann jede Zeile besetzt. Das ist ein offensichtliches Faktum, das wir schon im Lemma 2.1.6 diskutiert haben.

2.3 Variationsabstand und Cutoff-Effekt

Wir betrachten als Erstes eine n-Zellenkonstellation, in die k mal eine Kugel gelegt werde. Diese k Bestückungen erfolgen *rein zufällig* und *unabhängig* voneinander, d.h. es sind auch Mehrfachbesetzungen möglich. In unserer üblichen Analogiebetrachtung entspricht dies k unabhängig voneinander ausgeführten sukzessiven T1TRS. Von Interesse ist im Folgenden die Verteilung der Anzahl der unbesetzten Zellen für $n \to \infty$, wobei hierbei $k = k(n)$ eine Funktion von n ist, weshalb wir zur Verdeutlichung zukünftig k_n anstelle von k schreiben werden. Falls der Mittelwert $\frac{k_n}{n}$ von Kugeln pro Zelle sehr groß ist, so können keine leeren Zellen erwartet werden. $P_{k_n}^n(0)$ in der Notation von Korollar 2.1.8 wird nahe Eins sein, und $P_{k_n}^n(u)$ wird nahe Null sein für alle $1 \leq u \leq n$. Falls andererseits $\frac{k_n}{n}$ gegen Null strebt, so gilt $P_{k_n}^n(u) \to 0$ für jedes fixierte $u \in \mathbb{N}_0$. Es stellt sich daher auf natürlichem Wege die Frage, welche Relation zwischen k_n und n zu bestehen hat, damit $P_{k_n}^n$ im Grenzübergang $n \to \infty$ nicht degeneriert. Der folgende Satz (vgl. Feller [9, S. 101ff]) beantwortet diese Frage.

Satz 2.3.1. *Falls* $\lambda_n \stackrel{\text{def}}{=} ne^{-\frac{k_n}{n}}$ *für* $n \to \infty$ *beschränkt ist, d.h.* $\limsup_{n\to\infty} \lambda_n < \infty$ *gilt, so folgt*

$$P_{k_n}^n(u) - e^{-\lambda_n}\frac{\lambda_n^u}{u!} \to 0, \quad n \to \infty, \quad \textit{für jedes fixierte } u \in \mathbb{N}_0. \tag{2.54}$$

Beweis. Der Leser lasse noch einmal in Kürze den Beweis von Korollar 2.1.8 Revue passieren, da wir uns auf diesen mehrfach beziehen werden. Insbesondere verwenden wir wieder die Notation

$$s_\nu^n \stackrel{\text{def}}{=} s_\nu = \binom{n}{\nu}\left(1-\frac{\nu}{n}\right)^{k_n},$$

wobei wir hier die Abhängigkeit von n durch entsprechende Indizierung kenntlich gemacht haben. Der folgende Beweis gliedert sich in zwei Schritte. In dem Ersten wird (2.54) für $u = 0$ bewiesen, und in dem Zweiten beweisen wir (2.54) für beliebiges, aber fixiertes $u \in \mathbb{N}_0$. Der zweite Schritt ist eine Zurückführung auf den ersten, wobei (2.23) als Verbindungsglied dienen wird.

Schritt 1. Unter der Konvention $s_\nu^n = 0$, $\nu > n$ folgt mittels (2.22)

$$e^{-\lambda_n} - P_{k_n}^n(0) = \sum_{\nu=0}^{\infty}(-1)^\nu\left(\frac{\lambda_n^\nu}{\nu!} - s_\nu^n\right). \tag{2.55}$$

Wir werden zeigen, dass die rechte Seite in (2.55) für $n \to \infty$ gegen Null konvergiert: Wegen

$$(n-\nu)^\nu \le n^{(\nu)} \le n^\nu, \quad \nu = 0, \dots, n$$

in der Notation von (2.50) ergibt sich

$$n^\nu\left(1-\frac{\nu}{n}\right)^{\nu+k_n} \le \nu!\, s_\nu^n = n^{(\nu)}\left(1-\frac{\nu}{n}\right)^{k_n} \le n^\nu\left(1-\frac{\nu}{n}\right)^{k_n}, \quad \nu = 0, \dots, n. \tag{2.56}$$

Aus der bekannten Ungleichung $e^t \ge 1 + t$, $t \in \mathbb{R}$ folgt einerseits

$$e^{-t} \ge 1 - t,\ t \in \mathbb{R} \quad \Longrightarrow \quad -t \ge \log(1-t),\ t < 1$$

und andererseits durch Einsetzen von $t = \frac{-y}{1-y}$, $y < 1\ (\Rightarrow t < 1)$

$$\frac{y}{1-y} \ge \log\left(1+\frac{y}{1-y}\right) = \log\left(\frac{1}{1-y}\right) = -\log(1-y),$$

was zusammen

$$\frac{-t}{1-t} \le \log(1-t) \le -t, \quad t < 1$$

ergibt. Unter Benutzung dieser Ungleichungen erhalten wir für $\nu = 0, \dots, n-1$ durch Anwendung der Identität $x = e^{\log x}$, $x > 0$ einerseits

$$\left(1-\frac{\nu}{n}\right)^{\nu+k_n} = e^{(\nu+k_n)\log\left(1-\frac{\nu}{n}\right)} \ge e^{(\nu+k_n)\frac{-\frac{\nu}{n}}{1-\frac{\nu}{n}}} = \left(e^{-\frac{\nu+k_n}{n-\nu}}\right)^\nu$$

und andererseits

$$\left(1-\frac{\nu}{n}\right)^{k_n} = e^{k_n\log\left(1-\frac{\nu}{n}\right)} \le \left(e^{-\frac{k_n}{n}}\right)^\nu.$$

Insgesamt folgen aus (2.56) damit die analytisch leichter zugänglichen Ungleichungen

$$\left(ne^{-\frac{\nu+k_n}{n-\nu}}\right)^\nu \leq \nu!\, s_\nu^n \leq \left(ne^{-\frac{k_n}{n}}\right)^\nu, \quad \nu = 0, \ldots, n-1. \tag{2.57}$$

Mit $\eta_n \stackrel{\text{def}}{=} ne^{-\frac{\nu+k_n}{n-\nu}}$ und $n > \nu$ ergibt dies in prägnanterer Notation

$$\eta_n^\nu \leq \nu!\, s_\nu^n \leq \lambda_n^\nu, \quad n > \nu. \tag{2.58}$$

Wir nehmen zunächst an, dass $0 < a < \lambda_n < b$ für ein fixiertes Intervall $[a, b]$ für alle n gilt. Als Nächstes wird gezeigt, dass unter dieser Annahme für jedes fixierte $\nu \in \mathbb{N}_0$ der Quotient der Schranken in (2.58) gegen Eins konvergiert. Wegen

$$\frac{\eta_n}{\lambda_n} = \frac{ne^{-\frac{\nu+k_n}{n-\nu}}}{ne^{-\frac{k_n}{n}}} = e^{\frac{k_n}{n} - \frac{\nu+k_n}{n-\nu}}$$

genügt es hierfür

$$\frac{k_n}{n} - \frac{\nu + k_n}{n-\nu} \to 0, \quad n \to \infty$$

nachzuweisen. Substituieren wir $k_n = n \log(\frac{n}{\lambda_n})$, so ergibt sich unter Benutzung von $0 < a < \lambda_n < b$, $n \in \mathbb{N}$ und $\frac{\log n}{n} \to 0$, $n \to \infty$

$$\begin{aligned} & \log\left(\frac{n}{\lambda_n}\right) - \frac{\nu}{n-\nu} - \frac{n}{n-\nu}\log\left(\frac{n}{\lambda_n}\right) \\ = & -\frac{\nu}{n-\nu} - \frac{\nu}{n-\nu}\log\left(\frac{n}{\lambda_n}\right) \\ = & \frac{\nu}{n-\nu}(\log\lambda_n - 1) - \nu\frac{n}{n-\nu}\frac{\log n}{n} \to 0, \quad n \to \infty. \end{aligned}$$

Damit ist

$$\frac{\eta_n}{\lambda_n} \to 1, \quad n \to \infty \quad (\nu \in \mathbb{N}_0 \text{ fixiert}) \tag{2.59}$$

gezeigt. Hieraus folgt mit (2.58)

$$\overbrace{\frac{\lambda_n^\nu}{\nu!} - s_\nu^n}^{\geq 0} + \overbrace{\frac{\lambda_n^\nu}{\nu!}\left(\frac{\eta_n^\nu}{\lambda_n^\nu} - 1\right)}^{\to 0} = \frac{\eta_n^\nu}{\nu!} - s_\nu^n \leq 0, \quad n > \nu,$$

woraus sich

$$0 \leq \limsup_{n\to\infty}\left(\frac{\lambda_n^\nu}{\nu!} - s_\nu^n\right) = \limsup_{n\to\infty}\left(\frac{\eta_n^\nu}{\nu!} - s_\nu^n\right) \leq 0$$

und daraus unter erneuter Beachtung von (2.58)

$$\frac{\lambda_n^\nu}{\nu!} - s_\nu^n \to 0, \quad n \to \infty \quad (\nu \in \mathbb{N}_0 \text{ fixiert}) \tag{2.60}$$

ergibt. Letzteres ist auch richtig, falls wir auf die obige Einschränkung $\lambda_n > a$, $n \geq 1$ verzichten. Denn angenommen (o.E. $\nu > 0$)

$$\frac{\lambda_n^\nu}{\nu!} - s_\nu^n \not\to 0, \quad n \to \infty.$$

Dann existierte wegen der Beschränktheit von $n \mapsto \frac{\lambda_n^\nu}{\nu!} - s_\nu^n \geq 0$, $n > \nu$ eine Teilfolge

$(n_r)_r$ mit

$$\frac{\lambda_{n_r}^\nu}{\nu!} - s_\nu^{n_r} \to \alpha > 0, \quad r \to \infty. \tag{2.61}$$

Wegen

$$0 \leq \frac{\lambda_{n_r}^\nu}{\nu!} - s_\nu^{n_r} \leq \frac{\lambda_{n_r}^\nu}{\nu!}$$

folgte notwendig $\lambda_{n_r} \not\to 0$, $r \to \infty$, da sonst (2.61) verletzt wäre. Es existierte daher ein $a > 0$ und eine Teilfolge $(n_z)_z$ von $(n_r)_r$ mit $\lambda_{n_z} > a$, $z \geq 1$. Für $(n_z)_z$ anstelle von n ergäbe sich nach dem bisher Bewiesenen wieder (2.59), woraus wieder

$$\frac{\lambda_{n_z}^\nu}{\nu!} - s_\nu^{n_z} \to 0, \quad z \to \infty$$

folgte, was ein Widerspruch zu (2.61) ist. Daher ist (2.60) stets erfüllt, falls λ_n beschränkt ist. Wir werden nun mittels majorisierter Konvergenz zeigen, dass die Summe in (2.55) für $n \to \infty$ gegen Null konvergiert: Da $s_n^n \in \{0, 1\}$,

$$s_n^n = 1 \iff k_n = 0 \iff \lambda_n = n$$

und $\frac{n^n}{n!} \geq 1$ gilt, folgt $\frac{\lambda_n^n}{n!} \geq s_n^n$, d.h. in (2.58) gilt die obere Schranke auch für $\nu = n$. Wir folgern nun mit $\beta \stackrel{\text{def}}{=} \sup_n \lambda_n < \infty$ und $s_\nu^n = 0$, $\nu > n$ aus (2.58) die entscheidende Abschätzung

$$\left|(-1)^\nu \left(\frac{\lambda_n^\nu}{\nu!} - s_\nu^n\right)\right| = \frac{\lambda_n^\nu}{\nu!} - s_\nu^n \leq \frac{\lambda_n^\nu}{\nu!} \leq \frac{\beta^\nu}{\nu!}, \quad n \in \mathbb{N},\, \nu \in \mathbb{N}_0,$$

woraus wegen (2.60) und

$$\sum_{\nu=0}^{\infty} \frac{\beta^\nu}{\nu!} = e^\beta < \infty$$

wie gewünscht

$$e^{-\lambda_n} - P_{k_n}^n(0) \to 0, \quad n \to \infty$$

folgt.

Schritt 2. Sei nun $u \in \mathbb{N}_0$ fixiert. Nach (2.23) gilt $P_{k_n}^n(u) = s_u^n \cdot P_{k_n}^{n-u}(0)$. Mit

$$\widetilde{\lambda_n} \stackrel{\text{def}}{=} (n-u)e^{-\frac{k_n}{n-u}}, \quad n > u$$

folgt nach näherer Betrachtung der Indizes aus dem bisher Bewiesenen

$$P_{k_n}^{n-u}(0) - e^{-\widetilde{\lambda_n}} \to 0, \quad n \to \infty, \tag{2.62}$$

sofern $\limsup_{n\to\infty} \widetilde{\lambda_n} < \infty$ gilt, was wegen $\widetilde{\lambda_n} \leq \lambda_n$, $n > u$ und der Voraussetzung $\limsup_{n\to\infty} \lambda_n < \infty$ der Fall ist. Es wird im Folgenden

$$\lambda_n - \widetilde{\lambda_n} \to 0, \quad n \to \infty \tag{2.63}$$

gezeigt. Es ist

$$\begin{aligned} \widetilde{\lambda_n} &= (n-u)e^{-\frac{k_n}{n-u}}, \quad n > u \\ &= ne^{-\frac{k_n}{n-u}} - ue^{-\frac{k_n}{n-u}} \end{aligned}$$

$$\begin{aligned} &= ne^{-\frac{k_n}{n-u}} - \delta_n, \quad \delta_n \stackrel{\text{def}}{=} ue^{-\frac{k_n}{n-u}} \\ &= n^{-\frac{u}{n-u}} n^{\frac{n}{n-u}} \left(e^{-\frac{k_n}{n}}\right)^{\frac{n}{n-u}} - \delta_n \\ &= n^{-\frac{u}{n-u}} \lambda_n^{\frac{n}{n-u}} - \delta_n. \end{aligned} \tag{2.64}$$

Offenbar gilt $\delta_n \to 0$, $n \to \infty$, da

$$0 \le \delta_n \le \frac{u}{n} \cdot ne^{-\frac{k_n}{n}} = \frac{u}{n}\lambda_n \to 0, \quad n \to \infty.$$

Ferner haben wir

$$n^{-\frac{u}{n-u}} \to 1, \quad n \to \infty$$

und

$$\lambda_n - \lambda_n^{\frac{n}{n-u}} \to 0, \quad n \to \infty.$$

Die letzte Konvergenz wollen wir noch etwas genauer begründen. Sei $0 < \epsilon < 1$. Betrachte die Zerlegung

$$\begin{aligned} &\left|\lambda_n - \lambda_n^{\frac{n}{n-u}}\right| \\ = \;& \mathbb{1}_{(0,\epsilon]}(\lambda_n) \left|\lambda_n - \lambda_n^{1+\frac{u}{n-u}}\right| + \mathbb{1}_{(\epsilon,\infty)}(\lambda_n) \left|1 - \lambda_n^{\frac{u}{n-u}}\right| \lambda_n \\ \le \;& \underbrace{\mathbb{1}_{(0,\epsilon]}(\lambda_n) \left(\lambda_n + \lambda_n^{1+\frac{u}{n-u}}\right)}_{\le 2\epsilon} + \underbrace{\mathbb{1}_{(\epsilon,\infty)}(\lambda_n) \left|1 - e^{\frac{u}{n-u}} \overbrace{\mathbb{1}_{(\epsilon,\infty)}(\lambda_n)\log(\lambda_n)}^{\text{beschränkt}}\right|}_{\to 0} \lambda_n. \end{aligned}$$

Hieraus folgt

$$\limsup_{n\to\infty} \left|\lambda_n - \lambda_n^{\frac{n}{n-u}}\right| \le 2\epsilon, \quad \epsilon > 0$$

und damit die Konvergenz gegen Null. (2.63) folgt jetzt aus (2.64) und den darunter getroffenen Konvergenzaussagen, denn

$$\begin{aligned} \lambda_n - \widetilde{\lambda_n} &= \lambda_n - n^{-\frac{u}{n-u}} \lambda_n^{\frac{n}{n-u}} + \delta_n \\ &= \overbrace{\lambda_n - \lambda_n^{\frac{n}{n-u}}}^{\to 0} + \overbrace{\left(1 - n^{-\frac{u}{n-u}}\right)}^{\to 0} \overbrace{\lambda_n^{\frac{n}{n-u}}}^{\text{beschränkt}} + \overbrace{\delta_n}^{\to 0} \\ &\to 0, \quad n \to \infty. \end{aligned}$$

Aus $\lambda_n - \widetilde{\lambda_n} \to 0$, $n \to \infty$ und der Tatsache, dass eine stetige Funktion auf einem kompakten Intervall gleichmäßig stetig ist, folgt

$$e^{-\widetilde{\lambda_n}} - e^{-\lambda_n} \to 0, \quad n \to \infty. \tag{2.65}$$

Durch Addition von (2.62) und (2.65) erhalten wir zunächst

$$P_{k_n}^{n-u}(0) - e^{-\lambda_n} \to 0, \quad n \to \infty. \tag{2.66}$$

Hieraus folgt schließlich

$$P_{k_n}^n(u) - \frac{\lambda_n^u}{u!} e^{-\lambda_n} = s_u^n P_{k_n}^{n-u}(0) - \frac{\lambda_n^u}{u!} e^{-\lambda_n}$$

$$= \left(s_u^n - \frac{\lambda_n^u}{u!}\right) e^{-\lambda_n} + s_u^n \left(P_{k_n}^{n-u}(0) - e^{-\lambda_n}\right) \to 0, \quad n \to \infty,$$

wobei der erste Term wegen (2.60) und $\lambda_n > 0$, $n \geq 1$ gegen Null konvergiert. Für den zweiten Term beachte man (2.66) und die Beschränktheit von $n \mapsto s_u^n$, denn

$$s_u^n = \overbrace{s_u^n - \frac{\lambda_n^u}{u!}}^{\to 0} + \overbrace{\frac{\lambda_n^u}{u!}}^{\text{beschränkt}}.$$

□

Bemerkung 2.3.2. λ ist insbesondere beschränkt, wenn es von n unabhängig ist. Hierzu muss mit $c \stackrel{\text{def}}{=} -\log \lambda$ folgende Relation gelten:

$$\lambda = ne^{-\frac{\widetilde{k_n}}{n}} \iff \log \lambda = \log n - \frac{\widetilde{k_n}}{n} \iff \widetilde{k_n} = n \log n + cn.$$

Da $\widetilde{k_n} \in \mathbb{R} - \mathbb{N}$ keinen Sinn ergibt, betrachten wir $k_n \stackrel{\text{def}}{=} \lfloor \widetilde{k_n} \rfloor$ oder $k_n \stackrel{\text{def}}{=} \lceil \widetilde{k_n} \rceil$. Wegen

$$\begin{aligned} o(n) &= \lfloor n \log n + cn \rfloor - (n \log n + cn), \quad c \in \mathbb{R}, \\ o(n) &= \lceil n \log n + cn \rceil - (n \log n + cn), \quad c \in \mathbb{R} \end{aligned}$$

ergibt in beiden Fällen eine Anwendung von Korollar 2.3.3, dass die Verteilungskonvergenz $P_{k_n}^n \to \mathrm{Poi}(e^{-c})$ vorliegt. Der Leser beachte weiter, dass, um $k_n < 0$ bei $c < 0$ auszuschließen, $(P_{k_n}^n)_{n \geq N_c}$ natürlich für ein passendes $N_c \geq 1$ betrachtet werden muss.

Korollar 2.3.3. *Sei $(k_n)_{n \geq 1}$ eine Folge, so dass $k_n \in \mathbb{N}_0$, $n \geq 1$ und*

$$k_n = n \log n + cn + o(n)$$

für ein fixiertes $c \in \mathbb{R}$ gilt. Wir haben dann die Konvergenz

$$P_{k_n}^n(u) \to e^{-e^{-c}} \frac{(e^{-c})^u}{u!} = \mathrm{Poi}(e^{-c})(u), \quad n \to \infty, \quad u \in \mathbb{N}_0 \text{ fixiert.}$$

Beweis. Mit λ_n wie im Satz 2.3.1 gilt

$$\lambda_n = ne^{-\frac{k_n}{n}} = ne^{-\frac{n \log n + cn}{n}} \cdot e^{-\frac{o(n)}{n}} = e^{-c}(1 + o(1)), \quad n \in \mathbb{N},$$

folglich $\lambda_n \to e^{-c}$, $n \to \infty$ und damit

$$e^{-\lambda_n} \frac{\lambda_n^u}{u!} - e^{-e^{-c}} \frac{(e^{-c})^u}{u!} \to 0. \tag{2.67}$$

Da insbesondere $\limsup_{n \to \infty} \lambda_n < \infty$ ist, ergibt eine Anwendung von Satz 2.3.1 durch Addition von (2.54) und (2.67) schließlich

$$P_{k_n}^n(u) - e^{-e^{-c}} \frac{(e^{-c})^u}{u!} \to 0, \quad n \to \infty, \quad \text{für jedes fixierte } u \in \mathbb{N}_0.$$

□

Mit diesem Hintergrundwissen erscheint nächstes Theorem interessant. Genauer handelt es sich um eines der wichtigsten technischen Hilfsmittel bei der Untersuchung von *Top-to-Random-Shuffles* in Bezug auf den Variationsabstand zur Gleichverteilung. Es kann daher als eines der *Hauptresultate* dieser Arbeit angesehen werden.

Theorem 2.3.4. *Sei $\lambda > 0$ fixiert. Für jedes $n \geq 1$ sei eine Verteilung μ_n auf $\{0, \ldots, n\}$ gegeben, so dass die Konvergenz $\mu_n(n-j) \to e^{-\lambda}\frac{\lambda^j}{j!}$, $n \to \infty$ für jedes fixierte $j \in \mathbb{N}_0$ besteht. Mit Q_{μ_n} wie in (2.5) gilt dann*

$$\|Q_{\mu_n} - U\| = 1 - e^{-\lambda}\sum_{u=0}^{l^*}\lambda^u\left(\frac{1}{u!} - \frac{1}{(l^*+1)!}\right) - \frac{1}{(l^*+1)!} + o(1), \quad n \to \infty, \tag{2.68}$$

wobei

$$l^* \stackrel{\text{def}}{=} \begin{cases} 1 & \text{falls } 0 < \lambda \leq 1, \\ \left\lfloor \frac{\log(e^\lambda(\lambda-1)+1)}{\log\lambda} \right\rfloor - 1 & \text{falls } \lambda > 1. \end{cases} \tag{2.69}$$

Beweis. Wir zeigen als Erstes

$$\|Q_{\mu_n} - U\| = \|Q^L_{\mu_n} - U^L\|,$$

wobei $L(\pi)$ die Länge der aufsteigenden Sequenz in π bezeichne, die n enthält (siehe Lemma 2.1.1). Q_{μ_n} (siehe Satz 2.1.7) und U sind gleichverteilt auf $\{L = l\}$, falls wir unter diesem Ereignis bedingen. Hieraus folgt für beliebiges $A \in \mathfrak{S}_n$

$$\begin{aligned} |Q_{\mu_n}(A) - U(A)| &= \left|\sum_{l=1}^{n}\big[Q_{\mu_n}(A|L=l)Q_{\mu_n}(L=l) - U(A|L=l)U(L=l)\big]\right| \\ &= \left|\sum_{l=1}^{n}U(A|L=l)\big[Q_{\mu_n}(L=l) - U(L=l)\big]\right|. \end{aligned}$$

Offenbar maximiert $A \stackrel{\text{def}}{=} L^{-1}(B)$ mit $B \stackrel{\text{def}}{=} \{l : Q_{\mu_n}(L=l) \geq U(L=l)\}$ oberen Ausdruck. Das bedeutet

$$\|Q_{\mu_n} - U\| = |Q_{\mu_n}(A) - U(A)| = |Q^L_{\mu_n}(B) - U^L(B)| = \|Q^L_{\mu_n} - U^L\|,$$

also das Gewünschte.

Wir fahren nun mit der Berechnung von $\|Q^L_{\mu_n} - U^L\|$ fort. Setzen wir in Satz 2.1.7 $\mu \stackrel{\text{def}}{=} \mu_n$, so folgt aus diesem

$$\begin{aligned} Q_{\mu_n}(L=l) &= \frac{|\{L=l\}|}{n!}\sum_{u=0}^{l}\mu_n(n-u)u! \\ &= U(L=l)\sum_{u=0}^{l}\mu_n(n-u)u!, \quad l = 1, \ldots, n. \end{aligned} \tag{2.70}$$

$Q_{\mu_n}(L=l)/U(L=l)$ ist daher offensichtlich in l monoton wachsend. Sei

$$\begin{aligned} l^*_n &\stackrel{\text{def}}{=} \max\{l \geq 1 : U(L=l) \geq Q_{\mu_n}(L=l)\} \\ &= \max\{l \geq 1 : \sum_{u=0}^{l}\mu_n(n-u)u! \leq 1\}. \end{aligned} \tag{2.71}$$

Wir erhalten folglich

$$\begin{aligned} \|Q_{\mu_n} - U\| &= \sum_{l=1}^{l_n^*} (U(L=l) - Q_{\mu_n}(L=l)) \\ &= \sum_{l=1}^{l_n^*} U(L=l) \left(1 - \sum_{u=0}^{l} \mu_n(n-u)u!\right). \end{aligned} \tag{2.72}$$

Es wird als Nächstes

$$\begin{aligned} &\sum_{l=1}^{l_n^*} U(L=l) \left(1 - \sum_{u=0}^{l} \mu_n(n-u)u!\right) \\ = &\sum_{l=1}^{l^*} U(L=l) \left(1 - e^{-\lambda} \sum_{u=0}^{l} \lambda^u\right) + o(1), \quad n \to \infty \end{aligned} \tag{2.73}$$

mit

$$l^* \stackrel{\text{def}}{=} \max\{l \geq 1 \,:\, e^{-\lambda} \sum_{u=0}^{l} \lambda^u \leq 1\}$$

gezeigt: Der Leser beachte zunächst, dass $U(L=l), l = 1, \ldots, l^*$ für hinreichend großes $n \in \mathbb{N}$ *nicht* von n abhängt, da nach (2.18)

$$U(L=l) = \frac{1}{l!} \left(1 - \frac{1}{l+1}\right), \quad l = 1, \ldots, n-1$$

gilt. Weiter haben wir für fixiertes $l > l^*$ per Voraussetzung

$$\lim_{n\to\infty} \sum_{u=0}^{l} \mu_n(n-u)u! = e^{-\lambda} \sum_{u=0}^{l} \lambda^u > 1. \tag{2.74}$$

Hieraus folgt

$$\limsup_{n\to\infty} l_n^* \leq l^*,$$

da sonst eine Teilfolge $(l_{n_k}^*)_k$ existierte mit $l_{n_k}^* > l^*$, $k \in \mathbb{N}$. Dies ergäbe wegen (2.71)

$$\sum_{u=0}^{l^*+1} \mu_{n_k}(n_k-u)u! \leq \sum_{u=0}^{l_{n_k}^*} \mu_{n_k}(n_k-u)u! \leq 1, \quad k \in \mathbb{N}$$

und andererseits wegen (2.74)

$$\lim_{k\to\infty} \sum_{u=0}^{l^*+1} \mu_{n_k}(n_k-u)u! > 1,$$

was zusammen einen Widerspruch liefert. Zusätzlich gilt

$$\liminf_{n\to\infty} l_n^* \geq l^* - 1,$$

da sonst eine Teilfolge $(l_{m_k}^*)_k$ existierte mit $l_{m_k}^* \leq l^* - 2$, $k \in \mathbb{N}$. Aufgrund der

Voraussetzungen des Theorems und der Definition von l^* folgte

$$\begin{aligned}\limsup_{k\to\infty} \sum_{u=0}^{l^*_{m_k}+1} \mu_{m_k}(m_k-u)u! &\leq \lim_{k\to\infty} \sum_{u=0}^{l^*-1} \mu_{m_k}(m_k-u)u! \\ &= e^{-\lambda}\sum_{u=0}^{l^*} \lambda^u - e^{-\lambda}\lambda^{l^*} \\ &< 1,\end{aligned}$$

was

$$\sum_{u=0}^{l^*_{m_k}+1} \mu_{m_k}(m_k-u)u! < 1$$

für fast alle $k \in \mathbb{N}$ nach sich zöge. Das ist ein Widerspruch zur Definition von $l^*_{m_k}$ für solche $k \in \mathbb{N}$. Wir haben also insgesamt

$$l^* - 1 \leq \liminf_{n\to\infty} l^*_n \leq \limsup_{n\to\infty} l^*_n \leq l^*$$

gezeigt. Hieraus folgt

$$l^*_n \in \{l^*-1, l^*\}, \quad \text{für fast alle } n \in \mathbb{N}. \tag{2.75}$$

Falls

$$e^{-\lambda}\sum_{u=0}^{l^*} \lambda^u = 1 \implies U(L=l^*)\left(1 - e^{-\lambda}\sum_{u=0}^{l^*}\lambda^u\right) = 0,$$

so folgt (2.73) offenbar sofort aus (2.75) und den Voraussetzungen dieses Theorems, weshalb wir o.E. $e^{-\lambda}\sum_{u=0}^{l^*}\lambda^u < 1$ annehmen können. Es gilt dann $l^*_n = l^*$ für fast alle $n \in \mathbb{N}$, da sonst eine Teilfolge $(l^*_{r_k})_k$ existierte mit $l^*_{r_k} = l^* - 1$, $\forall k \in \mathbb{N}$, woraus sich der Widerspruch

$$1 \leq \lim_{k\to\infty} \sum_{u=0}^{\overbrace{l^*-1}^{l^*_{r_k}}+1} u_{r_k}(r_k-u)u! = \sum_{u=0}^{l^*} e^{-\lambda}\lambda^u < 1$$

ergibt. Hiermit ist (2.73) gezeigt. Zusammen mit (2.70), (2.72) und den Voraussetzungen des Theorems folgt hieraus weiter

$$\begin{aligned}\|Q_{\mu_n} - U\| &= \sum_{l=1}^{l^*} U(L=l)\left(1 - e^{-\lambda}\sum_{u=0}^{l}\lambda^u\right) + o(1) \\ &= \sum_{l=1}^{l^*} U(L=l)\left(1 - \sum_{u=0}^{l}\mu_n(n-u)u!\right) + o(1) \\ &= \sum_{l=1}^{l^*}(U(L=l) - Q_{\mu_n}(L=l)) + o(1), \quad n\to\infty. \end{aligned}\tag{2.76}$$

Wir können somit l^*_n in (2.72) durch das von n unabhängige, besser greifbare l^* ersetzen, müssen allerdings einen Fehler der Ordnung $o(1)$ (Nullfolge in n) hinzuaddieren. Wir werden nun (2.76) weiter umformen. Setzen wir in Satz 2.1.7 wieder

$\mu \stackrel{\text{def}}{=} \mu_n$, so erhalten wir nach einer einfachen Umformung von (2.15)

$$Q_{\mu_n}(L \geq l) = 1 - \sum_{u=0}^{l-1} \mu_n(n-u) \left(1 - \frac{u!}{l!}\right), \quad 1 \leq l \leq n. \tag{2.77}$$

Ferner gilt nach (2.4)

$$U(L \geq l) = \frac{|\{L \geq l\}|}{n!} = \frac{n!}{l!} \cdot \frac{1}{n!} = \frac{1}{l!}, \quad 1 \leq l \leq n. \tag{2.78}$$

Schließlich formen wir (2.76) mit Hilfe von (2.77) und (2.78) weiter um:

$$\begin{aligned}
\|Q_{\mu_n} - U\| &= Q_{\mu_n}(L \geq l^* + 1) - U(L \geq l^* + 1) + o(1) \\
&= 1 - \sum_{u=0}^{l^*} \mu_n(n-u) \left(1 - \frac{u!}{(l^*+1)!}\right) - \frac{1}{(l^*+1)!} + o(1) \\
&= 1 - \sum_{u=0}^{l^*} \left(e^{-\lambda}\frac{\lambda^u}{u!} + o_u(1)\right) \left(1 - \frac{u!}{(l^*+1)!}\right) - \frac{1}{(l^*+1)!} + o(1) \\
&= 1 - e^{-\lambda} \sum_{u=0}^{l^*} \lambda^u \left(\frac{1}{u!} - \frac{1}{(l^*+1)!}\right) - \frac{1}{(l^*+1)!} + o(1), \quad n \to \infty,
\end{aligned}$$

wobei die letzte Gleichung trivial ist, da l^* nicht von n abhängt. Hiermit ist (2.68) bewiesen.

Es bleibt noch die Darstellung (2.69) von l^* zu beweisen. Dies erledigen wir in einer Fallunterscheidung.

Fall 1. $0 < \lambda \leq 1$
Zunächst folgt wegen $e^\lambda \geq 1 + \lambda$ stets $1 \in \{l \geq 1 : e^{-\lambda} \sum_{u=0}^{l} \lambda^u \leq 1\} \neq \emptyset$. Ferner liefert ein Spezialfall der Restgliedabschätzung in der Reihendarstellung von e^λ in Forster [11, S.72]

$$e^\lambda \leq 1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3}, \quad \forall\, 0 \leq \lambda \leq 2.$$

Hieraus ergibt sich sofort

$$e^\lambda < 1 + \lambda + \lambda^2, \quad \forall\, 0 < \lambda \leq 1,$$

weshalb in diesem Fall $l^* = 1$ gilt.

Fall 2. $\lambda > 1$
Wir betrachten die Umformungen

$$\sum_{u=0}^{l} \lambda^u = \frac{\lambda^{l+1} - 1}{\lambda - 1} \leq e^\lambda \iff \lambda^{l+1} \leq e^\lambda(\lambda - 1) + 1 \iff l \leq \frac{\log(e^\lambda(\lambda-1)+1)}{\log \lambda} - 1.$$

Da l^* ganzzahlig sein muss, folgt unmittelbar (2.69), womit das Theorem vollständig bewiesen ist. □

Als Nächstes wird die Abbildung $\lambda \mapsto l^*(\lambda)$, $\lambda > 0$ mit $l^* = l^*(\lambda)$ aus (2.69) genauer untersucht. $l^*(\lambda)$ ist die größte ganze Zahl mit $e^\lambda \geq \sum_{u=0}^{l^*} \lambda^u$, wie im Beweis von Theorem 2.3.4 gezeigt wurde.

Lemma 2.3.5. *Für $l \geq 2$ besitzt die Gleichung $1 + \lambda + \ldots + \lambda^l = e^\lambda$ genau eine positive Lösung λ_l. Es gilt $\lambda_1 < \lambda_2 < \lambda_3 < \ldots \uparrow \infty$, wobei $\lambda_1 \stackrel{\text{def}}{=} 0$. Numerische Berechnungen zeigen:*

l	1	2	3	4	5	6	7
λ_l	0	1.8	5.1	8.8	12.8	17.1	21.5

.

Tabelle 1.1. Numerische Approximationen von λ_l

$\lambda \mapsto l^*(\lambda)$ ist eine monoton wachsende, rechtsseitig stetige Treppenfunktion auf $(0, \infty)$ mit den Sprungstellen $(\lambda_l)_{l\geq 2}$ und $l^*((0,\infty)) = \mathbb{N}$. Präziser haben wir

$$l^*(\lambda) = \sum_{l=1}^{\infty} l \cdot \mathbb{1}_{[\lambda_l, \lambda_{l+1})}(\lambda), \quad \lambda > 0. \tag{2.79}$$

Beweis. Es wird die über dem Lemma definierende Eigenschaft von l^* und nicht die geschlossene Formel (2.69) genutzt werden. Wir betrachten hierzu die Ungleichung

$$e^x \geq \sum_{u=0}^{l} x^u, \quad x > 0,\, l \geq 1$$

und stellen nach Division durch x^l und Nutzung der Exponentialreihe fest, dass diese äquivalent ist zu

$$\begin{aligned}
& e^x \geq \sum_{u=0}^{l} x^u, \quad x > 0,\, l \geq 1 && (2.80)\\
\iff \quad & \sum_{u=0}^{l} \frac{x^u}{u!} + \sum_{u=l+1}^{\infty} \frac{x^u}{u!} \geq \sum_{u=0}^{l} x^u \\
\iff \quad & \sum_{u=l+1}^{\infty} \frac{x^u}{u!} \geq \sum_{u=0}^{l} x^u \left(1 - \frac{1}{u!}\right) \\
\iff \quad & \sum_{u=l+1}^{\infty} \frac{1}{u!} x^{u-l} \geq \sum_{u=0}^{l} x^{u-l} \left(1 - \frac{1}{u!}\right). && (2.81)
\end{aligned}$$

Wird in obigen Ungleichungen die Relation $\geq$ durch $\leq, >, <$ oder $=$ ersetzt, so gelten immer noch die Äquivalenzen. Wegen

$$e^x < 1 + x + x^2 \leq \sum_{u=0}^{l} x^u, \quad \forall\, 0 < x \leq 1,\, l \geq 2$$

(siehe Fall 1 im Beweis von Theorem 2.3.4) und wegen $e^x > \sum_{u=0}^{l} x^u$ für alle hinreichend großen x mit fixiertem $l \geq 2$ liefert der Zwischenwertsatz für jedes $l \geq 2$ die Existenz eines $\lambda_l > 0$, so dass in (2.80) bzw. (2.81) Gleichheit besteht. Die linke Seite von (2.81) ist in x *streng* monoton wachsend, währenddessen die rechte in x monoton fallend ist. Gleichheit besteht in (2.81) bzw. (2.80) daher genau für ein $\lambda > 0$, woraus die Eindeutigkeit der $(\lambda_l)_{l\geq 2}$ folgt.

Offenbar gilt $e^{\lambda_l} < 1 + \lambda_l + \ldots + \lambda_l^l + \lambda_l^{l+1}$. Dieselbe strikte Ungleichung folgt daher mit $x \stackrel{\text{def}}{=} \lambda_l$ ebenso in (2.81), wobei dort l durch $l+1$ ersetzt werde. Für eine

Gleichheit in (2.81) und damit in (2.80) muss aufgrund der Monotonieeigenschaften von (2.81) deshalb $\lambda_{l+1} > \lambda_l$ gelten. Wegen $\lambda_l \geq \lambda_2 > 1$, $l \geq 2$ folgt

$$e^{\lambda_l} = 1 + \lambda_l + \ldots + \lambda_l^l \to \infty, \quad l \to \infty,$$

was $\lambda_l \uparrow \infty$ impliziert.

Zur letzten Aussage bzgl. der Treppenfunktion notieren wir

$$\sum_{u=0}^{l+1} \lambda_l^u > e^{\lambda_l} = \sum_{u=0}^{l} \lambda_l^u, \quad l \geq 2,$$

woraus

$$l^*(\lambda_l) = l, \quad l \geq 2 \tag{2.82}$$

folgt. Für $\lambda_l < \lambda < \lambda_{l+1}$, $l \geq 2$ erhalten wir einmal mehr mit dem Monotonieverhalten der beiden Terme in (2.81)

$$\sum_{u=0}^{l} \lambda^u < e^{\lambda} < \sum_{u=0}^{l+1} \lambda^u,$$

woraus $l^*(\lambda) = l$ folgt. Für $0 < \lambda < \lambda_2$ schließen wir analog $e^{\lambda} < 1 + \lambda + \lambda^2$. Wegen $e^{\lambda} > 1 + \lambda$, $\lambda > 0$ folgt daher für $0 = \lambda_1 < \lambda < \lambda_2$ ebenso $l^*(\lambda) = 1$. Insgesamt erhalten wir zusammen mit (2.82) leicht die Behauptungen bzgl. l^*.

Obere Tabelle wurde von Diaconis, Fill und Pitman [8] übernommen und kann mit einem Computeralgebrasystem leicht verifiziert werden. □

Beispiel 2.3.6. Für $\lambda \in [5.2, 8.7]$ besitzt nach obigem Lemma die Summe in (2.68) genau vier Terme. Hierbei haben wir absichtlich nicht $\lambda \in (5.1, 8.8]$ geschrieben, um evtl. Rundungsfehler in Tabelle 1.1 zu berücksichtigen.

Wir können mit unserem jetzigen Kenntnisstand leicht ein interessantes Theorem zu dem T1TRS beweisen.

Theorem 2.3.7. *Sei* $k_n \stackrel{\text{def}}{=} \lfloor n \log n + cn \rfloor$ *für ein fixiertes* $c \in \mathbb{R}$. *Es gilt dann*

$$\|Q_1^{*k_n} - U\| = f(c) + o(1), \quad n \to \infty$$

mit

$$f(c) \stackrel{\text{def}}{=} \frac{1}{2}(1 - e^{-e^{-c}}(1 + e^{-c})), \quad c \geq 0$$

und

$$f(c) \stackrel{\text{def}}{=} 1 - e^{-e^{-c}} \sum_{u=0}^{l^*} e^{-uc} \left(\frac{1}{u!} - \frac{1}{(l^*+1)!} \right) - \frac{1}{(l^*+1)!}, \quad c < 0,$$

wobei

$$l^* = l^*(c) = \left\lfloor \frac{\log(e^{e^{|c|}}(e^{|c|} - 1) + 1)}{|c|} \right\rfloor - 1$$

ist.

Beweis. Nach Bemerkung 2.3.2 gilt in der dortigen Notation

$$P_{k_n}^n(u) \to \mathrm{Poi}(e^{-c})(u), \quad n \to \infty, \quad u \in \mathbb{N}_0 \text{ fixiert.}$$

Wir erinnern daran, dass $P_{k_n}^n(u)$ die Wahrscheinlichkeit ist, bei k_n zufälligen, voneinander unabhängigen Besetzungen einer n Zellenkonstellation genau u Zellen *nicht* zu besetzen. μ_n sei eine Verteilung auf $\{0,\dots,n\}$ und werde durch

$$\mu_n(n-j) \stackrel{\text{def}}{=} P_{k_n}^n(j), \quad j=0,\dots,n$$

definiert. Direkt nach Definition von der Verknüpfung $\sharp$ (siehe Prosa vor Lemma 2.1.3) gilt

$$\mu_n = \overbrace{\delta_1 \sharp \dots \sharp \delta_1}^{k_n\text{-mal}}, \quad \delta_1 \text{ Dirac-Verteilung auf } \{0,\dots,n\},$$

woraus nach Satz 2.1.4

$$Q_1^{*k_n} = Q_{\delta_1}^{*k_n} = Q_{\delta_1 \sharp \dots \sharp \delta_1} = Q_{\mu_n}$$

folgt. Mit $\lambda \stackrel{\text{def}}{=} e^{-c}$ sind offenbar die Voraussetzungen von Theorem 2.3.4 erfüllt. Dieser Spezialfall liefert mit $f(c) \stackrel{\text{def}}{=} v(e^{-c})$ das, was zu zeigen war. v ist hierbei die Grenzvariation aus Theorem 2.3.4, d.h. v ist durch (2.86) definiert. □

Wir werden im Folgenden den Variationsabstand in (2.68), der sich für $n \to \infty$ ergibt, genauer untersuchen. Von Interesse sind hierbei die Fälle $\lambda \to 0$ und $\lambda \to \infty$, wobei hier mit dem komplizierteren Fall $\lambda \to \infty$ begonnen wird. Hierzu notieren wir zunächst drei technische Hilfsaussagen, die wir im Beweis vom Satz 2.3.9 benötigen. Diese fassen wir im folgenden Lemma zusammen.

Lemma 2.3.8. *(i) Sei $r > 0$ und $\delta : (r,\infty) \to \mathbb{R}$ eine beschränkte Funktion, d.h.*

$$a \stackrel{\text{def}}{=} \sup_{\gamma \in (r,\infty)} |\delta(\gamma)| < \infty.$$

Dann gilt

$$\left(1 + \frac{\delta(\gamma)}{\gamma}\right)^\gamma - e^{\delta(\gamma)} \to 0, \quad \gamma \to \infty.$$

(ii) Seien f_i, g_i, $i = 1,2$ nichtnegative Funktionen auf $(0,\infty)$ mit

$$g_i(\lambda) = (1 + o_i(1)) f_i(\lambda), \quad \lambda \to \infty, \quad i = 1,2.$$

Dann existiert eine Funktion φ auf $(0,\infty)$ mit $\varphi(\lambda) = o(1)$, $\lambda \to \infty$ und

$$g_1(\lambda) + g_2(\lambda) = (1 + \varphi(\lambda))(f_1(\lambda) + f_2(\lambda)), \quad \lambda > 0. \tag{2.83}$$

(iii) Es gilt

$$\log\left(e^\lambda(\lambda - 1) + 1\right) = \lambda + \log(\lambda) - (1 + o(1))\frac{1}{\lambda}, \quad \lambda \to \infty.$$

Beweis. (i) Es kann o.E. $\delta(\gamma) \neq 0$, $\gamma > r$ angenommen werden, da $\left(1 + \frac{0}{\gamma}\right)^\gamma - e^0 = 0$ gilt, und es daher ausreicht, obige Aussage mit

$$\widetilde{\delta}(\gamma) \stackrel{\text{def}}{=} \begin{cases} \delta(\gamma) & \text{falls } \delta(\gamma) \neq 0, \\ 1 & \text{sonst} \end{cases}, \quad \gamma > r$$

anstelle von δ zu beweisen. Wir haben $\frac{\delta(\gamma)}{\gamma} \to 0$, $\gamma \to \infty$, so dass aus $\frac{\mathrm{d}}{\mathrm{d}x} \log x\big|_{x=1} = 1$

$$\frac{\gamma}{\delta(\gamma)} \log\left(1 + \frac{\delta(\gamma)}{\gamma}\right) = \frac{\log\left(1 + \frac{\delta(\gamma)}{\gamma}\right) - \log 1}{\frac{\delta(\gamma)}{\gamma}} \to 1, \quad \gamma \to \infty$$

folgt. Sei $r_0 > r$ so groß, dass $\frac{\delta(\gamma)}{\gamma} > -1$, $\gamma > r_0$ gilt. Durch Anwendung der Exponentialfunktion erhalten wir dann

$$\left(1 + \frac{\delta(\gamma)}{\gamma}\right)^{\frac{\gamma}{\delta(\gamma)}} = e + \epsilon(\gamma) > 0, \quad \gamma > r_0$$

mit einer Funktion $\epsilon : (r_0, \infty) \to \mathbb{R}$, $\epsilon(\gamma) \to 0$, $\gamma \to \infty$. Dies ergibt

$$\left(1 + \frac{\delta(\gamma)}{\gamma}\right)^{\gamma} = (e + \epsilon(\gamma))^{\delta(\gamma)}, \quad \gamma > r_0. \tag{2.84}$$

Da

$$f(x, y) \stackrel{\text{def}}{=} (e + y)^x = e^{x \log(e+y)}, \quad (x, y) \in \mathbb{R} \times (-e, \infty)$$

als stetige Funktion auf jeder kompakten Teilmenge von $\mathbb{R} \times (-e, \infty)$, d.h. insbesondere auf $[-a, a] \times [-1, 1]$, gleichmäßig stetig ist, folgt aus

$$\|(\delta(\gamma), \epsilon(\gamma)) - (\delta(\gamma), 0)\|_2 = |\epsilon(\gamma)| \to 0, \quad \gamma \to \infty$$

die Konvergenz

$$(e + \epsilon(\gamma))^{\delta(\gamma)} - e^{\delta(\gamma)} \to 0, \quad \gamma \to \infty,$$

wobei mit $\|\cdot\|_2$ die euklidische Norm auf $\mathbb{R}^2$ gemeint ist. Mit (2.84) erhalten wir schließlich

$$\left(1 + \frac{\delta(\gamma)}{\gamma}\right)^{\gamma} - e^{\delta(\gamma)} = (e + \epsilon(\gamma))^{\delta(\gamma)} - e^{\delta(\gamma)} \to 0, \quad \gamma \to \infty.$$

(ii) Falls $f_1(\lambda) = f_2(\lambda) = 0$, wähle $\varphi(\lambda) = 0$. Im Falle $f_1(\lambda) + f_2(\lambda) > 0$ lässt sich (2.83) eindeutig nach $\varphi(\lambda)$ auflösen, und es gilt

$$\begin{aligned} \varphi(\lambda) &= \frac{g_1(\lambda) + g_2(\lambda)}{f_1(\lambda) + f_2(\lambda)} - 1 \\ &= \frac{o_1(1) f_1(\lambda) + o_2(1) f_2(\lambda)}{f_1(\lambda) + f_2(\lambda)}. \end{aligned}$$

Wegen

$$\begin{aligned} |o_1(1) f_1(\lambda) + o_2(1) f_2(\lambda)| &\le |o_1(1)| f_1(\lambda) + |o_2(1)| f_2(\lambda) \\ &\le (|o_1(1)| + |o_2(1)|)(f_1(\lambda) + f_2(\lambda)) \end{aligned}$$

folgt hieraus $|\varphi(\lambda)| \le |o_1(1)| + |o_2(1)|$, $\lambda > 0$ und daraus schließlich $\varphi(\lambda) = o(1)$.

(iii) Für $z > 0$ gilt

$$z(\log(z + 1) - \log z) = z \int_z^{z+1} \frac{dx}{x} \stackrel{y = \frac{x}{z}}{=} z \int_1^{1+\frac{1}{z}} \frac{dy}{y}.$$

Wegen

$$1 \overset{z\to\infty}{\longleftarrow} z \cdot \frac{1}{1+\frac{1}{z}} \cdot \frac{1}{z} \leq z \int_1^{1+\frac{1}{z}} \frac{dy}{y} \leq z \cdot \frac{1}{z} = 1$$

folgt

$$\log(z+1) - \log z = \frac{1}{z}(1+o(1)), \quad z \to \infty. \tag{2.85}$$

Mit $z \overset{\text{def}}{=} e^\lambda(\lambda-1)$ ergibt sich

$$\log\left(e^\lambda(\lambda-1)+1\right) = \log(e^\lambda(\lambda-1)) + \frac{1}{e^\lambda(\lambda-1)}(1+o(1)), \quad \lambda \to \infty,$$

was unter erneuter Anwendung von (2.85) mit $z \overset{\text{def}}{=} \lambda - 1$

$$\begin{aligned} \log\left(e^\lambda(\lambda-1)+1\right) &= \lambda + \log\lambda - \frac{1}{\lambda-1}(1+o(1)) + \frac{1}{e^\lambda(\lambda-1)}(1+o(1)) \\ &= \lambda + \log\lambda - (1+o(1))\frac{1}{\lambda}, \quad \lambda \to \infty \end{aligned}$$

nach sich zieht. □

Satz 2.3.9. *Sei*

$$v(\lambda) \overset{\text{def}}{=} 1 - e^{-\lambda} \sum_{u=0}^{l^*} \lambda^u \left(\frac{1}{u!} - \frac{1}{(l^*+1)!}\right) - \frac{1}{(l^*+1)!}, \quad \lambda > 0, \tag{2.86}$$

mit l^* wie in (2.69). Mit den Abkürzungen $L\lambda = \log\lambda$ und $\Delta(\lambda) = \frac{\lambda}{L\lambda} - \left\lfloor \frac{\lambda}{L\lambda} \right\rfloor$ gilt dann

$$\begin{aligned} & 1 - v(\lambda) \qquad (2.87) \\ = \; & (1+o(1))\frac{1}{\sqrt{2\pi}} e^{-\lambda + \frac{\lambda LL\lambda}{L\lambda} + \frac{\lambda}{L\lambda} - \frac{1}{2}(L\lambda - LL\lambda)} \cdot \left(\frac{1}{(L\lambda)^{\Delta(\lambda)}} + e^{-(1-\Delta(\lambda))(L\lambda - LL\lambda)}\right), \end{aligned}$$

für $\lambda \to \infty$.

Beweis. Der Beweis erfolgt in zwei Schritten. Wir zeigen im Schritt 1

$$\begin{aligned} & 1 - v(\lambda) \qquad (2.88) \\ = \; & (1+o(1)) \overbrace{\frac{1}{\sqrt{2\pi}} e^{-\lambda + \frac{\lambda LL\lambda}{L\lambda} + \frac{\lambda}{L\lambda} - \frac{1}{2}(L\lambda - LL\lambda)}}^{A} \cdot \left(\overbrace{\frac{1}{(L\lambda)^{1-\delta(\lambda)}}}^{B} + \overbrace{e^{-\delta(\lambda)(L\lambda - LL\lambda)}}^{C} \right), \end{aligned}$$

für $\lambda \to \infty$ mit

$$\begin{aligned} \delta(\lambda) &\overset{\text{def}}{=} l^*(\lambda) + 1 - \frac{\lambda}{L\lambda}, \quad \lambda > 1 \\ &= \left\lfloor \frac{L(e^\lambda(\lambda-1)+1)}{L\lambda} \right\rfloor - \frac{\lambda}{L\lambda}, \end{aligned} \tag{2.89}$$

siehe (2.69). Dies entspricht (2.87), wobei jedoch $\Delta(\lambda)$ durch das etwas undurchsichtigere $1-\delta(\lambda)$ ersetzt wurde. In Schritt 2 werden wir sehen, dass $1-\delta(\lambda)$ in (2.88) durch $\Delta(\lambda)$ ersetzt werden darf, womit der Satz bewiesen ist.

SCHRITT 1. Zunächst werden wir $v(\lambda)$ in eine andere Form bringen. Sei hierzu $p(u) \stackrel{\text{def}}{=} \frac{e^{-\lambda}\lambda^u}{u!}$ und $F(l) \stackrel{\text{def}}{=} \sum_{u=0}^{l} p(u)$. Dann gilt

$$v(\lambda) = 1 - F(l^*) - \frac{1}{(l^*+1)!} + e^{-\lambda} \sum_{u=0}^{l^*} \frac{\lambda^u}{(l^*+1)!}.$$

Nun ist

$$e^{-\lambda} \sum_{u=0}^{l^*} \frac{\lambda^u}{(l^*+1)!} = F(l^*)o(1),$$

denn

$$\frac{\sum_{u=0}^{l^*} \frac{\lambda^u}{(l^*+1)!} \cdot \frac{u!}{u!}}{\sum_{u=0}^{l^*} \frac{\lambda^u}{u!}} \leq \frac{1}{1+l^*} \longrightarrow 0, \quad \lambda \to \infty,$$

das bedeutet

$$v(\lambda) = 1 - (1-o(1))F(l^*) - \frac{1}{(l^*+1)!}, \quad \lambda \to \infty.$$

Weiter folgt

$$1 \leq \frac{F(l^*)}{p(l^*)} = \sum_{u=0}^{l^*} \frac{l^*(l^*-1)\cdot\ldots\cdot(u+1)}{\lambda^{l^*-u}} \leq \sum_{u=0}^{l^*} \left(\frac{l^*}{\lambda}\right)^{l^*-u} = \frac{1-\left(\frac{l^*}{\lambda}\right)^{l^*+1}}{1-\frac{l^*}{\lambda}} = 1+o(1),$$

da aus (2.69) offenbar sofort $\frac{l^*}{\lambda} = o(1)$ folgt. Dieses benutzend, gelangen wir zu

$$\begin{aligned} v(\lambda) &= 1 - (1-o(1))F(l^*) - \frac{1}{(l^*+1)!} \\ &= 1 - (1+o(1))p(l^*) - \frac{1}{(l^*+1)!} \\ &= 1 - \frac{1}{(l^*+1)!}\left((1+o(1))e^{-\lambda}\lambda^{l^*}l^* + 1\right), \end{aligned} \tag{2.90}$$

wobei in der letzten Zeile $\frac{l^*+1}{l^*} = (1+o(1))$, $\lambda \to \infty$ ausgenutzt wurde.

Als Nächstes wird $\frac{1}{(l^*+1)!}$ mit Hilfe der *Stirling Formel* [11, S.212] abgeschätzt. Für $z > 0$ sei $s(z) \stackrel{\text{def}}{=} (2\pi z)^{\left(\frac{1}{2}\right)} \left(\frac{z}{e}\right)^z$. Dann gilt mit $\gamma \stackrel{\text{def}}{=} \frac{\lambda}{L\lambda}$

$$\frac{1}{(l^*+1)!} = \frac{1+o(1)}{s(l^*+1)} = \frac{1+o(1)}{s(\gamma+\delta)} = \frac{1+o(1)}{s(\gamma)} \cdot \frac{s(\gamma)}{s(\gamma+\delta)}.$$

Der zweite Faktor des letzten Terms wird nun genauer betrachtet: Wir notieren zuerst, dass $|\delta(\lambda)| \leq 1$, $\lambda \geq 2$ aus (2.89) und

$$-1 \leq \left\lfloor \frac{\lambda}{L\lambda} \right\rfloor - \frac{\lambda}{L\lambda} \leq \left\lfloor \frac{L(e^\lambda(\lambda-1)+1)}{L\lambda} \right\rfloor - \frac{\lambda}{L\lambda} \leq \left\lfloor \frac{L(e^\lambda \cdot \lambda)}{L\lambda} \right\rfloor - \frac{\lambda}{L\lambda} \leq 1, \quad \lambda \geq 2$$

folgt. Da $\lambda \mapsto \gamma(\lambda)$ auf (e, ∞) streng monoton wachsend ist, kann δ auf diesem Intervall mittels $\delta(\lambda) = \delta(\lambda(\gamma))$ auch als Funktion von γ aufgefasst werden. Mit Lemma 2.3.8 (i) erhalten wir dann

$$\left(1+\frac{\delta}{\gamma}\right)^\gamma - e^\delta = o(1), \quad \lambda \to \infty \,(\Rightarrow \gamma \to \infty), \tag{2.91}$$

woraus unter Berücksichtigung von $e^{\delta} \geq e^{-1}$, $\lambda \geq 2$

$$\left(1+\frac{\delta}{\gamma}\right)^{\gamma} = o(1) + e^{\delta} = \left(1+\frac{o(1)}{e^{\delta}}\right) e^{\delta} = (1+o(1))e^{\delta}, \quad \lambda \to \infty$$

folgt, was

$$\left(\frac{\gamma}{\gamma+\delta}\right)^{\gamma} = \left(1+\frac{\delta}{\gamma}\right)^{-\gamma} = (1+o(1))e^{-\delta}, \quad \lambda \to \infty$$

begründet. Wir können somit

$$\begin{aligned}
\frac{s(\gamma)}{s(\gamma+\delta)} &= \frac{(2\pi)^{\frac{1}{2}}\gamma^{\frac{1}{2}}\left(\frac{\gamma}{e}\right)^{\gamma}}{(2\pi)^{\frac{1}{2}}(\gamma+\delta)^{\frac{1}{2}}\left(\frac{\gamma+\delta}{e}\right)^{\gamma+\delta}} \\
&= (1+o(1))e^{-\delta}\left(\frac{e}{\gamma+\delta}\right)^{\delta} \\
&= (1+o(1))\overbrace{\left(\frac{\gamma}{\gamma+\delta}\right)^{\delta}}^{\to 1,\, \lambda\to\infty}\gamma^{-\delta} \\
&= (1+o(1))\gamma^{-\delta}, \quad \lambda \to \infty
\end{aligned}$$

schreiben. Insgesamt haben wir folglich

$$\frac{1}{(l^*+1)!} = \frac{1+o(1)}{s(\gamma)}\gamma^{-\delta}, \quad \lambda \to \infty. \tag{2.92}$$

Als Nächstes wird A in (2.88) in eine uns ansprechendere Darstellung umgeformt.

$$\begin{aligned}
e^{-\lambda+\frac{\lambda LL\lambda}{L\lambda}+\frac{\lambda}{L\lambda}-\frac{1}{2}(L\lambda-LL\lambda)} &= e^{-\lambda}\cdot\left(e^{LL\lambda}\right)^{\gamma}\cdot e^{\gamma}\cdot\left(\lambda^{-\frac{1}{2}}\cdot e^{\frac{1}{2}LL\lambda}\right) \\
&= e^{-\lambda}\cdot\left(\gamma^{-\gamma}e^{\lambda}\right)\cdot e^{\gamma}\cdot\gamma^{-\frac{1}{2}} \\
&= \left(\frac{e}{\gamma}\right)^{\gamma}\gamma^{-\frac{1}{2}}.
\end{aligned} \tag{2.93}$$

C in (2.88) ergibt umgeformt

$$e^{-\delta(L\lambda-LL\lambda)} = \gamma^{-\delta}. \tag{2.94}$$

(2.92) lautet ausgeschrieben

$$\frac{1}{(l^*+1)!} = (1+o(1))\frac{1}{\sqrt{2\pi}}\gamma^{-\frac{1}{2}}\left(\frac{e}{\gamma}\right)^{\gamma}\gamma^{-\delta}. \tag{2.95}$$

Wir setzen (2.93) und (2.94) in (2.95) ein und erhalten

$$\frac{1}{(l^*+1)!} = (1+o(1))\frac{1}{\sqrt{2\pi}}e^{-\lambda+\frac{\lambda LL\lambda}{L\lambda}+\frac{\lambda}{L\lambda}-\frac{1}{2}(L\lambda-LL\lambda)}\cdot e^{-\delta(L\lambda-LL\lambda)},$$

womit unter Verwendung von (2.90) die „erste Hälfte" von (2.88) bewiesen ist.

Wir betrachten nun $A \cdot B$ in (2.88) und zeigen als Nächstes die unter Berücksichtigung von (2.90) noch fehlende Gleichung

$$(1+o(1))\frac{1}{\sqrt{2\pi}}e^{-\lambda+\frac{\lambda LL\lambda}{L\lambda}+\frac{\lambda}{L\lambda}-\frac{1}{2}(L\lambda-LL\lambda)}\cdot\frac{1}{(L\lambda)^{1-\delta}}=\frac{e^{-\lambda}\lambda^{l^*}l^*}{(l^*+1)!},\quad \lambda\to\infty.$$

Einerseits gilt

$$\begin{aligned} e^{-\lambda}\lambda^{l^*}l^* &= e^{-\lambda}\lambda^{\gamma+\delta-1}(\gamma+\delta-1)\\ &= \lambda^{\delta-1}(\gamma+\delta-1), \end{aligned}$$

andererseits

$$\frac{1}{(L\lambda)^{1-\delta}}=\gamma^{1-\delta}\lambda^{\delta-1}.$$

Unter Beachtung von (2.93) und (2.95) haben wir daher folgende Gleichung nachzuweisen:

$$(1+o(1))\gamma^{1-\delta}\lambda^{\delta-1}\frac{1}{\sqrt{2\pi}}\left(\frac{e}{\gamma}\right)^{\gamma}\gamma^{-\frac{1}{2}}=\lambda^{\delta-1}(\gamma+\delta-1)\frac{1}{\sqrt{2\pi}}\gamma^{-\frac{1}{2}}\left(\frac{e}{\gamma}\right)^{\gamma}\gamma^{-\delta}.$$

Nach Kürzen bleibt also

$$(1+o(1))\gamma=(\gamma+\delta-1).$$

Dies ist äquivalent zu

$$1+o(1)=\frac{\gamma+\delta-1}{\gamma}=1+(\delta-1)\frac{1}{\gamma},\quad \lambda\to\infty\,(\Rightarrow\gamma\to\infty),$$

was offenbar richtig ist, da δ beschränkt ist. Insgesamt haben wir daher

$$1-v(\lambda)=(1+o_1(1))\cdot A\cdot B+(1+o_2(1))\cdot A\cdot C \tag{2.96}$$

nachgewiesen. Nach Lemma 2.3.8 (ii) lassen sich die Funktionen, welche in (2.96) durch Landausymbole dargestellt werden, auf einen Vorfaktor $1+o(1)$ zusammenziehen. Damit ist Beweisschritt 1 abgeschlossen.

Schritt 2. Wir werden

$$\frac{\frac{1}{(L\lambda)^{1-\delta(\lambda)}}+e^{-\delta(\lambda)(L\lambda-LL\lambda)}}{\frac{1}{(L\lambda)^{\Delta(\lambda)}}+e^{(\Delta(\lambda)-1)(L\lambda-LL\lambda)}}\longrightarrow 1,\quad \lambda\to\infty \tag{2.97}$$

zeigen, womit der Beweis vollbracht ist. (2.69) zusammen mit Lemma 2.3.8 (iii) ergibt

$$l^*(\lambda)=\left\lfloor\frac{\lambda}{L\lambda}-\frac{1+o(1)}{\lambda L\lambda}\right\rfloor,\quad \lambda\to\infty. \tag{2.98}$$

Sei $\Lambda>1$ so groß, dass

$$0<\frac{1+o(1)}{\lambda L\lambda}<1,\quad \lambda\geq\Lambda \tag{2.99}$$

gilt. Ferner bezeichnen $(\lambda^{(l)})_{l\geq 3}$ die eindeutigen Lösungen der Gleichungen

$$\frac{\lambda}{L\lambda}=l,\quad l\geq 3$$

auf $\lambda\in[e,\infty)$. Die Existenz folgt wegen $\frac{e}{L(e)}=e<3$ und $\frac{\lambda}{L\lambda}\to\infty$, $\lambda\to\infty$

aus dem Zwischenwertsatz, da $\lambda \mapsto \frac{\lambda}{L\lambda}$ stetig ist. Weil $\frac{\lambda}{L\lambda}$ auf $[e,\infty)$ streng monoton wachsend ist, erhalten wir die Eindeutigkeit der Lösungen und auch $\lambda^{(l)} \uparrow \infty$. Ferner sei $N \in \mathbb{N}$, $N \geq 3$ so groß, dass $\lambda^{(N)} > \Lambda$ ist. Für $l \geq N$ gilt dann wegen (2.98) und (2.99)

$$l^*(\lambda^{(l)}) = \left\lfloor \frac{\lambda^{(l)}}{L\lambda^{(l)}} - \frac{1+o(1)}{\lambda^{(l)} L\lambda^{(l)}} \right\rfloor = \left\lfloor l - \frac{1+o(1)}{\lambda^{(l)} L\lambda^{(l)}} \right\rfloor = l-1.$$

Nach (2.79) folgt hieraus

$$\lambda_{l-1} \leq \lambda^{(l)} < \lambda_l, \quad l \geq N.$$

Wir erkennen deshalb mit (2.79)

$$l^*(\lambda) = \begin{cases} l-1 & \text{falls } \lambda \in [\lambda^{(l)}, \lambda_l), \\ l & \text{falls } \lambda \in [\lambda_l, \lambda^{(l+1)}) \end{cases}, \quad l \geq N. \tag{2.100}$$

Sei

$$g(\lambda) \stackrel{\text{def}}{=} \left\lfloor \frac{\lambda}{L\lambda} \right\rfloor - \frac{\lambda}{L\lambda} + 1 = 1 - \Delta(\lambda).$$

Aus (2.100) folgt dann

$$g(\lambda) - \delta(\lambda) = \left\lfloor \frac{\lambda}{L\lambda} \right\rfloor - l^*(\lambda) = \begin{cases} 1 & \text{falls } \lambda \in [\lambda^{(l)}, \lambda_l), \\ 0 & \text{falls } \lambda \in [\lambda_l, \lambda^{(l+1)}) \end{cases}, \quad l \geq N. \tag{2.101}$$

Beachte, dass der Nenner in (2.97) seinem Zähler entspricht, wenn wir in dem Zähler $\delta(\lambda)$ durch $g(\lambda)$ ersetzen. Falls $\lambda \in [\lambda_l, \lambda^{(l+1)})$, $l \geq N$, ist (2.97) wegen $g(\lambda) = \delta(\lambda)$ gleich Eins, bedarf also keiner weiteren Betrachtung.

Wir untersuchen folglich die Situation $\lambda \in [\lambda^{(l)}, \lambda_l)$ und werden als Nächstes

$$|\delta(\lambda)| \leq \frac{C}{\lambda}, \quad \lambda \in [\lambda^{(l)}, \lambda_l), \quad l \geq N \tag{2.102}$$

zeigen, wobei die Konstante C *nicht* von l abhängt. Es gilt für $\lambda \in [\lambda^{(l)}, \lambda_l)$, $l \geq N$

$$\begin{aligned} \delta(\lambda) = l^*(\lambda) - \frac{\lambda}{L\lambda} + 1 &= l - 1 - \frac{\lambda}{L\lambda} + 1 - l + \frac{\lambda^{(l)}}{L\lambda^{(l)}} \\ &= \frac{\lambda^{(l)}}{L\lambda^{(l)}} - \frac{\lambda}{L\lambda}. \end{aligned}$$

Hieraus folgt

$$|\delta(\lambda)| = \frac{\lambda}{L\lambda} - \frac{\lambda^{(l)}}{L\lambda^{(l)}} \leq \frac{\lambda - \lambda^{(l)}}{L\lambda^{(l)}} \leq \lambda - \lambda^{(l)} \leq \lambda_l - \lambda^{(l)}. \tag{2.103}$$

Aus einer Anwendung des Mittelwertsatzes auf die Funktion $\varphi(x) \stackrel{\text{def}}{=} \frac{x}{Lx}$, $x > 1$ wird im Folgenden (2.102) gewonnen: Aus diesem folgt

$$\varphi(x_2) - \varphi(x_1) \geq m(x_2 - x_1), \quad 1 < x_1 \leq x_2, \tag{2.104}$$

falls $m \leq \inf_{x_1 \leq x \leq x_2} \varphi'(x)$ gilt. Routineumformungen ergeben

$$\varphi'(x) = \frac{1}{Lx} - \frac{1}{(Lx)^2} \geq \frac{1}{2Lx}, \quad x \geq e^2.$$

Durch evtl. Vergrößerung von N kann o.E. $\lambda^{(N)} \geq e^2$ angenommen werden. Für $\lambda \in [\lambda^{(l)}, \lambda_l]$, $l \geq N$ folgt speziell

$$\varphi'(\lambda) \geq \frac{1}{2L\lambda} \geq \frac{1}{2L\lambda_l}$$

und daraus wegen (2.104) mit $m \stackrel{\text{def}}{=} \frac{1}{2L\lambda_l}$, $x_1 \stackrel{\text{def}}{=} \lambda^{(l)}$, $x_2 \stackrel{\text{def}}{=} \lambda_l$

$$\frac{1}{2L\lambda_l}(\lambda_l - \lambda^{(l)}) \leq \frac{\lambda_l}{L\lambda_l} - \frac{\lambda^{(l)}}{L\lambda^{(l)}} = \frac{\lambda_l}{L\lambda_l} - l^*(\lambda_l) = \frac{1+o(1)}{\lambda_l L\lambda_l},$$

wobei die letzte Gleichheit richtig ist, da λ_l eine Sprungstelle von l^* ist und daher wegen der Stetigkeit der Funktion

$$\lambda \mapsto \frac{L(e^\lambda(\lambda-1)+1)}{L\lambda} = \frac{\lambda}{L\lambda} - \frac{1+o(1)}{\lambda L\lambda} + 1$$

schon im Innern der Gaußklammer von l^* notwendig

$$\frac{\lambda_l}{L\lambda_l} - \frac{1+o(1)}{\lambda_l L\lambda_l} + 1 \in \mathbb{Z}$$

gelten muss. Dies ergibt

$$\lambda_l - \lambda^{(l)} \leq \frac{2(1+o(1))}{\lambda_l} \leq \frac{C}{\lambda}, \quad \lambda \in [\lambda^{(l)}, \lambda_l), \quad l \geq N$$

für eine von l unabhängige Konstante $C > 0$. Zusammen mit (2.103) folgt daraus (2.102).

Nun lässt sich (2.97) leicht zeigen. Sei hierzu $(x_n)_{n\in\mathbb{N}}$ eine Folge mit $x_n \to \infty$ für $n \to \infty$ und

$$x_n \in \bigcup_{l\geq N} [\lambda^{(l)}, \lambda_l), \quad n \in \mathbb{N}.$$

Nach der Prosa unter (2.101) reicht es zu zeigen, dass der Ausdruck in (2.97) für $n \to \infty$ gegen Eins konvergiert, wobei λ durch x_n ersetzt wird. Genauer werden wir nachweisen, dass Zähler und Nenner gegen Eins konvergieren.

Es wird mit dem Zähler begonnen: Wir haben mit (2.102) für hinreichend große $n \in \mathbb{N}$

$$\begin{aligned} \frac{1}{(Lx_n)^{1-\delta(x_n)}} &= (Lx_n)^{\delta(x_n)-1} \leq (Lx_n)^{-\frac{1}{2}}, \\ |-\delta(x_n)(Lx_n - LLx_n)| &= |\delta(x_n)|(Lx_n - LLx_n) \\ &\leq \frac{C}{x_n}(Lx_n - LLx_n), \end{aligned}$$

woraus sofort die behauptete Konvergenz im Zähler folgt.

Für den Nenner notieren wir zunächst wegen (2.101)

$$\begin{aligned} \Delta(x_n) &= \frac{x_n}{Lx_n} - \left\lfloor \frac{x_n}{Lx_n} \right\rfloor = 1 - g(x_n) = 1 - (\delta(x_n)+1) = -\delta(x_n) \\ \Rightarrow \Delta(x_n) - 1 &= -(1+\delta(x_n)), \quad n \in \mathbb{N}. \end{aligned}$$

Wir notieren weiter für alle hinreichend großen $n \in \mathbb{N}$ unter erneuter Beachtung von

(2.102)

$$(\Delta(x_n)-1)(Lx_n-LLx_n) = -(1+\delta(x_n))(Lx_n-LLx_n) \le -\frac{1}{2}(Lx_n-LLx_n),$$
$$\frac{1}{(Lx_n)^{\Delta(x_n)}} = \frac{1}{(Lx_n)^{-\delta(x_n)}} = e^{\delta(x_n)LLx_n}, \quad |\delta(x_n)LLx_n| \le \frac{C}{x_n}LLx_n.$$

Hieraus folgt unmittelbar die Konvergenz des Nenners gegen Eins. Damit ist (2.97) und somit der gesamte Satz bewiesen. □

Mit (2.87) haben wir eine Annäherung an $1-v(\lambda)$, $\lambda\to\infty$ in dem Sinne gefunden, dass der relative Fehler für $\lambda\to\infty$ gegen Null konvergiert. Genauer folgen, falls

$$a(\lambda) \stackrel{\text{def}}{=} \frac{1-v(\lambda)}{1+o(1)}$$

die Approximation in (2.87) bezeichnet, die Umformungen

$$\begin{aligned}\frac{(1-v(\lambda))-a(\lambda)}{1-v(\lambda)} &= \frac{1-v(\lambda)-\left(\frac{1-v(\lambda)}{1+o(1)}\right)}{1-v(\lambda)} = \frac{(1-v(\lambda))\left(1-\frac{1}{1+o(1)}\right)}{1-v(\lambda)}\\ &= 1-\frac{1}{1+o(1)} \to 0, \quad \lambda\to\infty.\end{aligned}$$

Es besteht ebenso die Konvergenz

$$\frac{(1-v(\lambda))-a(\lambda)}{a(\lambda)} \to 0, \quad \lambda\to\infty.$$

In z.B. der Stirling Formel konvergiert ebenfalls der relative Approximationsfehler gegen Null. Wir sind nun an einer einfacheren und weniger subtilen Abschätzung als (2.87) interessiert und notieren hierzu ein Korollar.

Korollar 2.3.10. *Es gilt mit v wie in (2.86)*

$$1-v(\lambda) = e^{-(1+o(1))\lambda}, \quad \lambda\to\infty.$$

Beweis. Wir definieren eine Funktion f auf (e,∞) durch

$$f(\lambda) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}}\left(\frac{1}{(L\lambda)^{\Delta(\lambda)}} + e^{-(1-\Delta(\lambda))(L\lambda-LL\lambda)}\right)$$

und notieren für $\lambda\in(e,\infty)$ die Abschätzungen

$$\frac{1}{\lambda} = e^{-L\lambda} \le e^{-LL\lambda} \le e^{-\Delta(\lambda)LL\lambda} = \frac{1}{(L\lambda)^{\Delta(\lambda)}} \le 1,$$
$$\frac{1}{\lambda} = e^{-L\lambda} \le e^{-(L\lambda-LL\lambda)} \le e^{-(1-\Delta(\lambda))(L\lambda-LL\lambda)} \le 1,$$

woraus $\frac{1}{\sqrt{2\pi}}\frac{2}{\lambda} \le f(\lambda) \le \frac{1}{\sqrt{2\pi}}\cdot 2$ und daraus

$$\frac{L\left(\frac{2}{\sqrt{2\pi}}\right)-L\lambda}{\lambda} = \frac{L\left(\frac{1}{\sqrt{2\pi}}\frac{2}{\lambda}\right)}{\lambda} \le \frac{Lf(\lambda)}{\lambda} \le \frac{L\left(\frac{2}{\sqrt{2\pi}}\right)}{\lambda}$$

folgt. Dies ergibt insbesondere

$$\frac{Lf(\lambda)}{\lambda} = o(1), \quad \lambda \to \infty.$$

Sei weiter eine Funktion g auf (e, ∞) durch

$$g(\lambda) \stackrel{\text{def}}{=} -\lambda + \frac{\lambda LL\lambda}{L\lambda} + \frac{\lambda}{L\lambda} - \frac{1}{2}(L\lambda - LL\lambda) = -\lambda(1 + o(1)), \quad \lambda \to \infty$$

definiert. Nach Satz 2.3.9 gilt

$$\begin{aligned} 1 - v(\lambda) &= (1 + o(1))f(\lambda)e^{g(\lambda)} \\ &= e^{L(1+o(1))+Lf(\lambda)+g(\lambda)} \\ &= e^{\left(\frac{L(1+o(1))}{\lambda} + \frac{Lf(\lambda)}{\lambda} + \frac{g(\lambda)}{\lambda}\right)\lambda} \\ &= e^{-(1+o(1))\lambda}, \quad \lambda \to \infty, \end{aligned}$$

womit der Beweis erbracht ist. □

Bemerkung 2.3.11. Nach $k_n \stackrel{\text{def}}{=} \lfloor n \log n + nc \rfloor$ Mischvorgängen des T1TRS entfernt sich die Verteilung des Kartendecks wegen $\lambda = e^{-c}$ (siehe Theorem 2.3.7) für $c \to -\infty$ sogar in *doppelt* exponentieller Geschwindigkeit von der Gleichverteilung auf $\mathfrak{S}_n$ für großes n. Da hier zwei Größen unendlich werden (nämlich n und c), muss der Grenzübergang noch präzisiert werden: In einem ersten Schritt wähle man ein $c < 0$ und fixiere dieses c. Je kleiner dieses c ist, desto kleiner wird $1 - v(e^{-c})$ (siehe Korollar 2.3.21). Die Annäherung an Null geschieht hierbei, wie bereits bemerkt, in *doppelt* exponentieller Geschwindigkeit. In einem zweiten Schritt betrachte man den Grenzübergang $n \to \infty$ (c ist fixiert). Nach Theorem 2.3.7 und Korollar 2.3.10 gilt dann

$$\|Q_1^{*k_n} - U\| \stackrel{n\to\infty}{\longrightarrow} f(c) = v(e^{-c}) = 1 - e^{-(1+o(1))e^{-c}}, \tag{2.105}$$

wobei mit $o(1)$ hier eine Funktion gemeint ist mit

$$o(1) = o(1)(\lambda) = o(1)(e^{-c}) \to 0, \text{ falls } c \to -\infty.$$

Naheliegenderweise sollte im Sinne der Anschauung das fixierte $c < 0$ so klein gewählt werden, dass (2.105) nahezu Eins ergibt. Es reichen dann $\lfloor k_n = n \log n + nc \rfloor$ Mischvorgänge für großes n nicht mehr aus, um das Kartendeck zu durchmischen.

Das nächste, weitaus einfachere Vorgehen besteht in der asymptotischen Untersuchung von $v(\lambda)$ für $\lambda \to 0$.

Korollar 2.3.12. *Es gilt* $v(\lambda) = \frac{1}{4}\lambda^2(1 + O(\lambda)) = o(1), \quad \lambda \to 0.$

Beweis. Für $\lambda \leq 1$ ist $v(\lambda) = \frac{1}{2}(1 - e^{-\lambda}(1 + \lambda))$, da $l^* = 1$ für $\lambda \leq 1$ in (2.69). Einige elementare Umformungen ergeben unter Benutzung der Reihendarstellung von $e^{-\lambda}$

$$v(\lambda) = \frac{1}{2}(1 - e^{-\lambda}(1 + \lambda)) = \frac{1}{2}\sum_{j=2}^{\infty}(-1)^j(j-1)\frac{\lambda^j}{j!},$$

woraus die Behauptung sofort folgt. □

Bemerkung 2.3.13. Für den T1TRS bedeutet dies insbesondere

$$v(\lambda) = v(e^{-c}) = \frac{1}{4}e^{-2c}(1 + O(e^{-c})) = \frac{1}{4}(1 + o(1))e^{-2c}, \quad c \to \infty.$$

$v(e^{-c})$ nähert sich folglich in *einfacher* exponentieller Geschwindigkeit der Null an. Wir sehen also insgesamt, dass es bei $k_n = \lfloor n \log n + cn \rfloor$ Mischvorgängen auf die Wahl von $c \in \mathbb{R}$ ankommt, um zu entscheiden, ob k_n Mischvorgänge für großes n ausreichen oder nicht. Konkreter bedeutet dies, dass die Fälle $v(e^{-c}) \widetilde{=} 0$ bzw. $v(e^{-c}) \widetilde{=} 1$ für $n \to \infty$ bei fixiertem $c \in \mathbb{R}$ eintreffen, je nachdem ob $c > 0$ hinreichend groß bzw. $c < 0$ hinreichend klein gewählt ist.

Zur Anschauung wollen wir schließlich noch $\lambda \mapsto v(\lambda)$ in einer von *Maple* berechneten Abbildung darstellen.

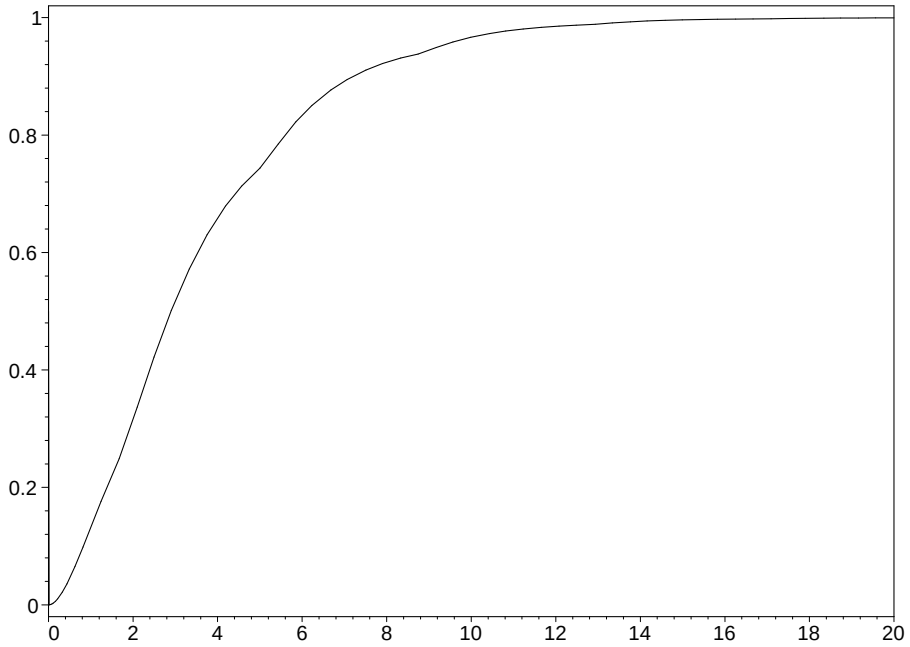

Abbildung 2.1: Darstellung der Funktion $\lambda \mapsto v(\lambda)$

Betrachten wir obige Abbildung genauer, so fallen uns an den Stellen λ_l leichte „Dellen" auf. Genauer sind diese zumindest noch für $l = 2, 3, 4$ in vorliegender Auflösung zu erkennen, siehe Tabelle 1.1 in Lemma 2.3.5. Mit Blick auf (2.69) ist klar, dass diese von der Gaußklammer stammen. Wir erkennen ferner mittels (2.86) und der Rechtsstetigkeit von $l^*(\lambda)$ (siehe Lemma 2.3.5) sofort, dass $v(\lambda)$ rechtsstetig ist. Eine interessante Frage ist allerdings, ob $v(\lambda)$ auch stetig ist. Der Computerplot deutet zumindest darauf hin. Außerdem spräche es mit der Interpretation von $v(\lambda)$ im Sinne von Theorem 2.3.7 auch eher gegen unsere Intuition, wäre $v(\lambda)$ tatsächlich nicht stetig. Dennoch erkennt man diese nicht unmittelbar an (2.86). Wir werden die Stetigkeit von $v(\lambda)$ nun beweisen.

Lemma 2.3.14. *$\lambda \mapsto v(\lambda)$ ist auf $(0, \infty)$ stetig.*

Beweis. Offenbar muss nur die Linksstetigkeit in den Punkten $(\lambda_l)_{l \geq 2}$ nachgewiesen werden. Seien hierzu $l \geq 2$ und $\epsilon > 0$ gegeben, so dass $\lambda_{l-1} < \lambda_l - \epsilon$ gilt. Wir haben dann

$$\begin{aligned} & v(\lambda_l) - v(\lambda_l - \epsilon) \\ = \; & e^{-(\lambda_l - \epsilon)} \sum_{u=0}^{l-1} (\lambda_l - \epsilon)^u \left(\frac{1}{u!} - \frac{1}{l!} \right) + \frac{1}{l!} - e^{-\lambda_l} \sum_{u=0}^{l} \lambda_l^u \left(\frac{1}{u!} - \frac{1}{(l+1)!} \right) - \frac{1}{(l+1)!}. \end{aligned}$$

Im Grenzübergang $\epsilon \downarrow 0$, $\epsilon > 0$, $\lambda_{l-1} < \lambda_l - \epsilon$ erhalten wir folglich

$$\begin{aligned} & e^{-\lambda_l} \sum_{u=0}^{l-1} \lambda_l^u \left(\frac{1}{u!} - \frac{1}{l!} \right) + \frac{1}{l!} - e^{-\lambda_l} \sum_{u=0}^{l} \lambda_l^u \left(\frac{1}{u!} - \frac{1}{(l+1)!} \right) - \frac{1}{(l+1)!} \\ = \quad & e^{-\lambda_l} \left(\frac{1}{(l+1)!} - \frac{1}{l!} \right) \sum_{u=0}^{l} \lambda_l^u + \left(\frac{1}{l!} - \frac{1}{(l+1)!} \right) = 0, \end{aligned}$$

wobei letzte Gleichheit wegen $\sum_{u=0}^{l} \lambda_l^u = e^{\lambda_l}$, $l \geq 2$ folgt, siehe Lemma 2.3.5. □

$v(\lambda)$ ist in den Punkten $(\lambda_l)_{l \geq 2}$ allerdings *nicht* differenzierbar. Genauer ist die rechtsseitige Ableitung in diesen Punkten echt größer als die linksseitige. Da jedoch solche Betrachtungen für unsere Zwecke keinen greifbaren Nutzen ergeben, werden wir auf eine genauere Kurvendiskussion verzichten.

Der Leser stelle sich ein vorsortiertes n-Kartendeck vor, das $k = n \log n + nc$ - mal im Sinne eines T1TRS gemischt werde, wobei n eine große, fixierte natürliche Zahl sei. Die Anzahl der Mischschritte k variiere nun, d.h. c ist in diesem Gedankenexperiment unsere Variable. Etwas salopp formuliert haben wir durch die Analyse von $\lambda \mapsto v(\lambda)$ den Phasenübergang zwischen den drei Zuständen „ungemischt", „teilweise gemischt" und „gemischt" analysiert. Unter Beachtung von $\lambda = e^{-c}$ veranschaulicht obige Abbildung, dass der Zustand „teilweise gemischt" in einem kleinen Intervall um $c = 0$ anzusiedeln ist. Entscheidend ist aber, dass außerhalb dieses Zustandes die beiden anderen Zustände exponentiell schnell erreicht werden ($c \to \pm\infty$). Es gibt also keinen langsamen, kontinuierlichen Wechsel zwischen ungemischt und gemischt. Dieser geschieht, lassen wir c gedanklich von $-\infty$ nach $+\infty$ laufen, nach einem „kurzen" Aufenthalt in dem Zustand „teilweise gemischt" quasi sprunghaft. In diesem Sinne ist der zuvor erwähnte *Cutoff-Effekt* zu verstehen. Es macht daher durchaus Sinn zu sagen, ein Kartendeck benötige $n \log n$ T1TRS Schritte, um durchmischt zu werden. Streng genommen werden allerdings „etwas mehr" als $n \log n$ Schritte benötigt, da wir schließlich nicht in dem „teilweise gemischt" Zustand schon von einem durchmischten Kartendeck sprechen möchten. Um dies anzudeuten und auch die „richtige" Relation zwischen k und n für $n \to \infty$ noch einmal zu unterstreichen, wird oft geschrieben, dass $n \log n + nc$ T1TRS vonnöten sind, um das Deck zu durchmischen. Die Wahl von $c > 0$ hängt natürlich davon ab, wie dicht man sich an der Gleichverteilung befinden möchte. Für quantitative Aussagen müssten einige Werte von $v(\lambda)$ evtl. numerisch berechnet werden. Solche Betrachtungen sind allerdings für qualitative Betrachtungen, worum es in dieser Arbeit geht, eher uninteressant.

Der Leser beachte an dieser Stelle auch die Abbildung 1 aus der Einleitung. Hier haben wir mit $n = 52$ Karten

$$k \mapsto \|Q_1^{*k} - U\|$$

für $k \in \{20m + 5,\ 0 \leq m \leq 20\}$ exakt berechnet und

$$k \mapsto v(\lambda(k))$$

mit

$$\lambda(k) \stackrel{\text{def}}{=} e^{-c} = e^{-\frac{k - n \log n}{n}} = n e^{-\frac{k}{n}} = 52 e^{-\frac{k}{52}}, \quad k \in [5, 405]$$

geplottet. Zum einen erkennen wir die Approximation von $v(\lambda)$, die in Theorem 2.3.7 mit $f(c) = v(e^{-c})$ nachgewiesen wird, zum anderen wird auch ein *Cutoff-Effekt* sichtbar ($52 \log 52 \widetilde{=} 205.46$).

Wir werden als Nächstes einen *Cutoff-Effekt* bei einer ganzen Klasse von Wahrscheinlichkeitsverteilungen nachweisen. Zu dieser gehört auch der TmTRS, $m \geq 1$. Ein Spezialfall des folgenden Lemmas besagt, dass es $\frac{n}{m}(\log n + c)$ TmTRS bedarf, um ein Kartendeck mit n Karten zu durchmischen. Die Ordnungsklasse $o(n)$ in (2.106) besagt u.a., dass eine eventuelle Rundung von $\frac{n}{m}(\log n + c)$ keine Auswirkungen auf das Ergebnis hat, d.h. $\lfloor \frac{n}{m}(\log n + c) \rfloor$ und $\lceil \frac{n}{m}(\log n + c) \rceil$ sind in diesem Sinne gleichwertig.

Satz 2.3.15. *Sei $b \in \mathbb{N}$ fixiert und μ eine Verteilung auf $\{0, 1, \ldots, b\}$ mit dem Erwartungswert $\overline{\mu} \stackrel{\text{def}}{=} \sum_{i=1}^{b} i\mu(i) > 0$. Wir werden im Folgenden μ bei gegebenen $n \geq b$ als Verteilung auf $\{0, \ldots, n\}$ mit $\mu(\{b+1, \ldots, n\}) = 0$ auffassen. Q_μ sei wieder wie in (2.5) und v wie in (2.86) definiert. Für fixiertes $c \in \mathbb{R}$ sei*

$$k_n \stackrel{\text{def}}{=} \frac{n}{\overline{\mu}}(\log n + c) + o(n) \in \mathbb{N}, \quad \forall n \in \mathbb{N}. \tag{2.106}$$

Dann gilt

$$\|Q_\mu^{*k_n} - U\| = v(e^{-c}) + o(1), \quad n \to \infty. \tag{2.107}$$

Beweis. Nach Satz 2.1.4 haben wir

$$Q_\mu^{*k_n} = Q_{\underbrace{\mu \sharp \ldots \sharp \mu}_{k_n-\text{mal}}},$$

so dass es nach Theorem 2.3.4 ausreicht zu zeigen, dass für jedes fixierte $u \in \mathbb{N}_0$ die Konvergenz

$$\overbrace{\mu \sharp \ldots \sharp \mu}^{k_n-\text{mal}}(n-u) \to \text{Poi}(e^{-c})(u), \quad n \to \infty$$

besteht. Um dieses nachzuweisen, werden wir als Nächstes Zufallsvariablen $(Z_n)_{n \geq 1}$ konstruieren, die der Bedingung

$$Z_n \sim \overbrace{\mu \sharp \ldots \sharp \mu}^{k_n-\text{mal}}, \quad n \geq 1$$

und der Verteilungskonvergenz

$$n - Z_n \to \text{Poi}(e^{-c}), \quad n \to \infty \tag{2.108}$$

genügen:

Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum mit einer unabhängigen Familie $\mathcal{F}$ von Zufallsvariablen $X_1, \ldots, X_{k_n}, Y_1, Y_2, \ldots$, wobei

$$P^{X_i} = \mu,\ i = 1, \ldots, k_n, \quad P^{Y_t} = \text{Gleichverteilung auf } \{1, 2, \ldots, n\},\ t \geq 1$$

gilt. Nach dem Satz von Andersen/Jessen [3, S.98] ist die Existenz eines solchen Wahrscheinlichkeitsraumes gesichert. Genau genommen müssten wir wegen der n-Abhängigkeit $(\Omega_n, \mathfrak{A}_n, P_n)$ anstelle von $(\Omega, \mathfrak{A}, P)$ schreiben und ebenso jede Zufallsvariable hierauf mit n indizieren, wovon wir aus Gründen der Übersichtlichkeit aber absehen werden. Der Leser halte sich jedoch diese Tatsache innerhalb des gesamten

Beweises immer vor Augen. Wir definieren weiter

$$
\begin{aligned}
&N_s^t \stackrel{\text{def}}{=} \sum_{r=1}^{n} \left(1 - \prod_{j=s}^{t} (1 - \mathbb{1}_{\{r\}}(Y_j)) \right), \quad 1 \le s \le t,\ s,t \in \mathbb{N},\\
&N(0) \stackrel{\text{def}}{=} 0, \quad N(t) \stackrel{\text{def}}{=} N_1^t, \quad t \in \mathbb{N},\\
&W_0 \stackrel{\text{def}}{=} 0,\\
&W_i \stackrel{\text{def}}{=} \begin{cases} \inf\{t \ge \sum_{j=1}^{i-1} W_j + 1 \,:\, N^t_{\sum_{j=1}^{i-1} W_j + 1} = X_i\} - \sum_{j=1}^{i-1} W_j & \text{für } X_i > 0,\\ 0 & \text{für } X_i = 0, \end{cases}\\
&i = 1, \ldots, k_n.
\end{aligned}
$$

Die etwas kompliziert aussehenden Definitionen haben für folgendes Experiment eine einfache Bedeutung: Es werde in eine n Zellenanordnung sukzessive und unabhängig voneinander jeweils eine Kugel rein zufällig platziert. $(N(t))_{t\ge 0}$ beschreibt dann die Anzahl der besetzten Zellen nach t solchen Platzierungen, wobei die t-te Kugel ($t \ge 1$) in Zelle Y_t gelegt werde. $N(t)$ ist offenbar eine Markov-Kette auf dem Zustandsraum $\{0, 1, \ldots, n\}$ mit Übergangsmatrix $\mathbb{P}(i, i+1) = 1 - \frac{i}{n}$, $\mathbb{P}(i,i) = \frac{i}{n}$, gestartet in $N(0) = 0$. W_1 beschreibt die Anzahl der Schritte, um X_1 disjunkte Zellen mit jeweils mindestens einer Kugel auszufüllen. W_2 ist die Anzahl weiterer Schritte, deren es bedarf, um X_2 disjunkte Zellen mit jeweils einer Kugel zu füllen. Die X_2 Zellen im zweiten Schritt können eventuell mit einigen der X_1 Zellen aus dem ersten Schritt überlappen. Die Disjunktheit im zweiten Schritt bezieht sich also nur auf die folgenden X_2 Zellen, die ersten X_1 Zellen werden hierbei ignoriert. Auf diese Art und Weise erhalten wir auch sukzessive die Bedeutung von $W_3, \ldots, W_{k_n}$. Entscheidend ist, dass offenbar

$$
N(W_1 + \ldots + W_{k_n}) \sim \overbrace{\mu \sharp \ldots \sharp \mu}^{k_n - \text{mal}}, \quad n \ge 1
$$

gilt. Um dies einzusehen, betrachte der Leser noch einmal die Definition von der Verknüpfung $\sharp$ vor Lemma 2.1.3. Mit

$$
Z_n \stackrel{\text{def}}{=} N(W_1 + \ldots + W_{k_n}), \quad n \ge 1
$$

bleibt folglich nur noch (2.108) zu zeigen:

Es erweist sich die spezielle Konstruktion der $(Z_n)_{n\ge 1}$ als hilfreich. Nach der Interpretation der $(W_i)_{i=1,\ldots,k_n}$ im obigen Experiment und der Unabhängigkeit der Familie $\mathcal{F}$ ist klar, dass $(W_i)_{i=1,\ldots,k_n}$ unabhängig und identisch verteilt sind. Sei $W \stackrel{\text{def}}{=} W_1$, $X \stackrel{\text{def}}{=} X_1$, und seien $(T_i)_{i=1,\ldots,n}$ unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\widetilde{\Omega}, \widetilde{\mathfrak{A}}, \widetilde{P})$. Ferner gelte $T_i \sim \text{Geo}\left(\frac{n-i+1}{n}\right)$, $i = 1, \ldots, n$ (geometrische Verteilung), wobei hierbei der Zeitpunkt des ersten Erfolges und *nicht* der Zeitpunkt des letzten Misserfolges betrachtet wird. Offenbar haben wir

$$
P^{W|X=x} = \widetilde{P}^{\sum_{i=1}^{x} T_i}, \quad x = 0, \ldots, n,
$$

und erinnern uns: Falls $Y \sim \text{Geo}(\theta)$, $\theta \in (0,1]$, gilt

$$
E(Y) = \frac{1}{\theta} \text{ und } \text{Var}(Y) = \frac{1-\theta}{\theta^2}, \quad \text{siehe [3, S.127]}.
$$

Hieraus folgt

$$E(W|X=x) = \sum_{i=0}^{x-1} \frac{n}{n-i} = x + \frac{1}{n} \sum_{i=0}^{x-1} \frac{ni}{n-i} = x + O_x\left(\frac{1}{n}\right), \quad \text{für } n \to \infty \ (x \text{ fixiert})$$

und damit

$$\begin{aligned} E(W) &= \int E(W|X=x)\, P^X(dx) \\ &= \sum_{x=0}^{b} E(W|X=x)\mu(x) \\ &= \sum_{x=0}^{b} \left(x + O_x\left(\frac{1}{n}\right)\right)\mu(x) \\ &= \overline{\mu} + O\left(\frac{1}{n}\right). \end{aligned}$$

Mit $\mathrm{Var}(W|X=x)$ bezeichnen wir im Folgenden die Varianz von W bedingt auf $\{X=x\}$, d.h. präziser von der Verteilung $P^{W|X=x}$. Wir definieren daher

$$\begin{aligned} \mathrm{Var}(W|X=x) &\stackrel{\text{def}}{=} \int \left(w - \int w P^{W|X=x}(dw)\right)^2 P^{W|X=x}(dw), \\ \mathrm{Var}(W|X) &\stackrel{\text{def}}{=} \mathrm{Var}(W|X=\cdot) \circ X. \end{aligned}$$

Es gilt

$$\begin{aligned} \mathrm{Var}(W|X=x) &= \sum_{i=0}^{x-1} \frac{1-\theta_i}{\theta_i^2}, \quad \theta_i \stackrel{\text{def}}{=} \frac{n-i}{n},\ i = 0, \ldots, x-1 \\ &= \sum_{i=0}^{x-1} \frac{ni}{(n-i)^2} \\ &= o_x(1), \quad n \to \infty \ (x \text{ fixiert}). \end{aligned}$$

Wegen

$$\mathrm{Var}(W|X) = E(W^2|X) - (E(W|X))^2$$

folgt

$$E(\mathrm{Var}(W|X)) = E(W^2) - E((E(W|X))^2). \tag{2.109}$$

Andererseits gilt

$$\begin{aligned} \mathrm{Var}(E(W|X)) &= E((E(W|X))^2) - (E(E(W|X)))^2 \\ &= E((E(W|X))^2) - (E(W))^2, \end{aligned} \tag{2.110}$$

woraus durch Addition von (2.109) und (2.110)

$$\mathrm{Var}\, W = E(W^2) - (E(W))^2 = E(\mathrm{Var}(W|X)) + \mathrm{Var}(E(W|X))$$

folgt. Insgesamt haben wir damit

$$\mathrm{Var}\, W = E(o_X(1)) + \mathrm{Var}\left(X + O_X\left(\frac{1}{n}\right)\right)$$

$$\begin{aligned}
&= \sum_{x=0}^{b} o_x(1)\mu(x) + \sum_{x=0}^{b}\left(x + o_x(1) - \sum_{x'=0}^{b}(x' + o_{x'}(1))\mu(x')\right)^2 \mu(x)\\
&\to \sum_{x=0}^{b}\left(x - \sum_{x'=0}^{b} x'\mu(x')\right)^2 \mu(x), \quad n \to \infty\\
&= \operatorname{Var} X.
\end{aligned}$$

Sei ϵ_n, $n \geq 1$ definiert durch die Gleichung

$$\sum_{i=1}^{k_n} W_i = n\log n + (c + \epsilon_n)n. \tag{2.111}$$

Wir werden als Nächstes zeigen, dass für alle fixierten $\alpha > 0$ die Konvergenz

$$P(|\epsilon_n| \geq \alpha) \to 0, \quad n \to \infty$$

besteht. Es sei daran erinnert, dass $(\Omega, \mathfrak{A}, P)$ von $n \in \mathbb{N}$ abhängt, weshalb hier nicht von Konvergenz in Wahrscheinlichkeit gesprochen werden sollte. Der folgende Beweis hierzu basiert auf der Chebychev Ungleichung. Es gilt

$$\begin{aligned}
E\,\epsilon_n &= E\left(\frac{1}{n}\sum_{i=1}^{k_n} W_i - \log n - c\right)\\
&= \frac{k_n}{n}\left(\overline{\mu} + O\left(\frac{1}{n}\right)\right) - \log n - c\\
&= \left(\frac{\log n + c}{\overline{\mu}} + \frac{o(n)}{n}\right)\left(\overline{\mu} + O\left(\frac{1}{n}\right)\right) - \log n - c\\
&= \frac{\log n + c}{\overline{\mu}} O\left(\frac{1}{n}\right) + \frac{o(n)}{n}\left(\overline{\mu} + O\left(\frac{1}{n}\right)\right)\\
&= o(1).
\end{aligned}$$

Sei $t > 0$ fixiert. Wir haben dann für alle hinreichend großen n

$$P(|\epsilon_n| \geq 2t) \leq P(|\epsilon_n| \geq t + |o(1)|) \leq P(|\epsilon_n - o(1)| \geq t) = P(|\epsilon_n - E\epsilon_n| \geq t).$$

Nach Chebychev gilt

$$\begin{aligned}
P(|\epsilon_n - E\epsilon_n| \geq t) &\leq t^{-2}\operatorname{Var}\epsilon_n\\
&= t^{-2}\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{k_n} W_i - \log n - c\right)\\
&= t^{-2}\frac{k_n}{n^2}\operatorname{Var} W \to 0, \quad n \to \infty,
\end{aligned}$$

da $\frac{k_n}{n^2} \to 0$, $\operatorname{Var} W \to \operatorname{Var} X \leq E(X^2) \leq b^2 < \infty$, $n \to \infty$, womit (2.111) bewiesen ist.

Nach Bemerkung 2.3.2 gilt für jedes fixierte $u \in \mathbb{N}_0$ und $a \in \mathbb{R}$

$$\begin{aligned}
P(n - N(\lfloor n\log n + an\rfloor = u) = P^n_{\lfloor n\log n + an\rfloor}(u) &\to \operatorname{Poi}(e^{-a})(u),\\
P(n - N(\lceil n\log n + an\rceil = u) = P^n_{\lceil n\log n + an\rceil}(u) &\to \operatorname{Poi}(e^{-a})(u)
\end{aligned}$$

jeweils für $n \to \infty$. Ferner ist $t \mapsto n - N(t)$ bei fixiertem $n \in \mathbb{N}$ und $\omega \in \Omega$ offenbar monoton fallend. Bevor wir fortfahren, notieren wir noch eine triviale Tatsache:

Für jedes $n \in \mathbb{N}$ seien zwei Ereignisse $A_n, B_n \in \mathfrak{A}\,(= \mathfrak{A}_n)$ gegeben. Ferner gelte $\lim P(B_n) = 1$. Es folgt dann

$$\limsup P(A_n \cap B_n) = \limsup(P(A_n) + P(B_n) - P(A_n \cup B_n)) = \limsup P(A_n).$$

Eine analoge Aussage besteht für den *Limes inferior*.

Sei $\delta > 0$ fixiert. Wir erhalten unter Benutzung des obigen Faktums und der Monotonieeigenschaft von $n - N(t)$

$$\begin{aligned} & \limsup P(n - N(n\log n + (c+\epsilon_n)n) \le x) \\ = \quad & \limsup P(n - N(n\log n + (c+\epsilon_n)n) \le x, |\epsilon_n| \le \delta) \\ \le \quad & \limsup P(n - N(\lceil n\log n + (c+\delta)n\rceil) \le x, |\epsilon_n| \le \delta) \\ = \quad & \mathrm{Poi}\left(e^{-(c+\delta)}\right)(\{0,\dots,x\}). \end{aligned}$$

Die Eigenschaft $P(|\epsilon_n| \le \delta) \to 1$, $n \to \infty$ ($\forall \delta > 0, \delta$ fixiert) haben wir unter (2.111) nachgewiesen. Mit $\delta \downarrow 0$ folgt aus der Stetigkeit von $\delta \mapsto \mathrm{Poi}\left(e^{-(c+\delta)}\right)(\{0,\dots,x\})$ schließlich obige Ungleichung auch für $\delta = 0$. Andererseits gilt

$$\begin{aligned} & \liminf P(n - N(n\log n + (c+\epsilon_n)n) \le x) \\ = \quad & \liminf P(n - N(n\log n + (c+\epsilon_n)n) \le x, |\epsilon_n| \le \delta) \\ \ge \quad & \liminf P(n - N(\lfloor n\log n + (c-\delta)n\rfloor) \le x, |\epsilon_n| \le \delta) \\ = \quad & \mathrm{Poi}(e^{-(c-\delta)})(\{0,\dots,x\}). \end{aligned}$$

Betrachten wir hierbei auch wieder $\delta \downarrow 0$, so gilt insgesamt für alle $x \in \mathbb{N}_0$, x fixiert

$$\lim P(n - N(n\log n + (c+\epsilon_n)n) \le x) \overset{n\to\infty}{\longrightarrow} \mathrm{Poi}\left(e^{-c}\right)(\{0,\dots,x\}).$$

Hieraus folgt durch Subtraktion sofort

$$\lim P(n - N(n\log n + (c+\epsilon_n)n) = u) \overset{n\to\infty}{\longrightarrow} \mathrm{Poi}\left(e^{-c}\right)(u), \quad \forall u \in \mathbb{N}_0,$$

also

$$n - \overbrace{N(W_1 + \ldots + W_{k_n})}^{Z_n} \to \mathrm{Poi}(e^{-c}), \quad n \to \infty$$

in Verteilung. Hiermit ist die noch fehlende Konvergenz von (2.108) gezeigt, womit der Satz bewiesen ist. □

Im Sinne der Asymptotik von Satz 2.3.15 scheint es also keinen Unterschied zwischen $k \cdot m$ sukzessive ausgeführten T1TRS und k ausgeführten TmTRS zu geben. Genauer folgen mit $k_n \overset{\text{def}}{=} \left\lfloor \frac{n}{m}(\log n + c)\right\rfloor$ die Konvergenzen

$$\|Q_{\delta_m}^{*k_n} - U\| \overset{n\to\infty}{\longrightarrow} v(e^{-c}), \quad \|Q_{\delta_1}^{*(k_n\cdot m)} - U\| \overset{n\to\infty}{\longrightarrow} v(e^{-c}),$$

wobei $m \in \mathbb{N}$, $c \in \mathbb{R}$ fixiert sind. Dennoch ist die m-fache Faltung eines T1TRS nicht dasselbe wie ein TmTRS, was im Extremfall $m = n$ sofort klar wird. Ein TmTRS scheint sich der Gleichverteilung schneller anzunähern als m T1TRS. Diese Aussage wollen wir im Folgenden präzisieren und verallgemeinern. Der Leser betrachte zur Motivation des Folgenden schon an dieser Stelle die Aussage von Korollar 2.3.20.

Definition 2.3.16. Wir schreiben $\mu \leq \nu$ für zwei Verteilungen μ, ν auf $\{0, 1, \ldots, n\}$, falls

$$\sum_{i=0}^{j} \mu(i) \leq \sum_{i=0}^{j} \nu(i), \quad 0 \leq j \leq n$$

gilt, m.a.W. die Verteilungsfunktion von ν diejenige von μ dominiert. ν heißt dann *stochastisch kleiner* als μ.

Es wird sich herausstellen: Je größer eine Verteilung μ auf $\{0, 1, \ldots, n\}$ im Sinne der Relation „$\leq$" aus Definition 2.3.16 ist, desto weiter ist Q_μ von der Gleichverteilung U entfernt. Anschaulich gesprochen liegen wir umso näher an der Gleichverteilung, je mehr Karten abgehoben und wieder zwischengeschoben werden. Ein intuitiv klares Ergebnis, das natürlich noch präzisiert werden muss. Die Vorgehensweise im Zusammenhang mit Definition 2.3.16 wurde übrigens durch ein statistisches Problem in Alsmeyer [4] motiviert. Hier werden im Kontext der nichtparametrischen Statistik Verfahren entwickelt, darauf zu testen, ob eine Verteilung eine andere stochastisch dominiert oder nicht. An dieser Stelle notieren wir zunächst ein recht nützliches Lemma.

Lemma 2.3.17. *Es gilt $\mu \leq \nu$ genau dann, falls die Ungleichung*

$$\sum_{i=0}^{n} g(i)\mu(i) \geq \sum_{i=0}^{n} g(i)\nu(i)$$

für alle monoton wachsenden Folgen $(g(i))_{i=0,\ldots,n}$ besteht. Ferner ist

$$\mu(i) > 0,\ 0 \leq i \leq n \text{ und } i \mapsto \frac{\nu(i)}{\mu(i)} \text{ monoton fallend}$$

hierfür eine hinreichende Bedingung.

Beweis. Wir zeigen zuerst die Äquivalenz und nehmen $\mu \leq \nu$ an. Sei g eine monoton wachsende Folge, d.h.

$$g(j+1) - g(j) \geq 0, \quad j = 0, \ldots, n-1.$$

Durch Multiplikation folgt dann

$$(g(j+1) - g(j)) \sum_{i=0}^{j} \mu(i) \leq (g(j+1) - g(j)) \sum_{i=0}^{j} \nu(i), \quad j = 0, \ldots, n-1. \quad (2.112)$$

Wir summieren den linken Term über $j = 0, \ldots, n-1$ und erhalten nach Umsummierung

$$\begin{aligned}
\sum_{j=0}^{n-1} (g(j+1) - g(j)) \sum_{i=0}^{j} \mu(i) &= \sum_{i=0}^{n-1} \left(\sum_{j=i}^{n-1} (g(j+1) - g(j)) \right) \mu(i) \\
&= \sum_{i=0}^{n-1} (g(n) - g(i))\mu(i) \\
&= \sum_{i=0}^{n} (g(n) - g(i))\mu(i)
\end{aligned}$$

$$= \quad g(n) - \sum_{i=0}^{n} g(i)\mu(i).$$

Der rechte Term von (2.112) lässt sich ebenso umformen, so dass aus (2.112)

$$\sum_{i=0}^{n} g(i)\mu(i) \geq \sum_{i=0}^{n} g(i)\nu(i) \tag{2.113}$$

folgt.

Sei umgekehrt (2.113) gegeben. Wir setzen dann für ein $0 \leq j \leq n$

$$g(i) \stackrel{\text{def}}{=} \mathbb{1}_{\{j+1,\ldots,n\}}(i)$$

und erhalten daraus mit (2.113)

$$\sum_{i=j+1}^{n} \mu(i) \geq \sum_{i=j+1}^{n} \nu(i), \quad j = 0, \ldots, n,$$

was wegen

$$\sum_{i=0}^{j} \mu(i) = 1 - \sum_{j+1}^{n} \mu(i) \leq 1 - \sum_{i=j+1}^{n} \nu(i) = \sum_{i=0}^{j} \nu(i), \quad j = 0, \ldots, n$$

$\mu \leq \nu$ liefert.

Es bleibt noch die im Lemma behauptete hinreichende Bedingung zu zeigen: Es kann o.E. $\mu \neq \nu$ angenommen werden, woraus

$$\frac{\nu(i_0)}{\mu(i_0)} \neq 1 \text{ für ein } 0 \leq i_0 \leq n$$

folgt. Wegen der Normiertheit von μ und ν muss $\left(\frac{\nu(i)}{\mu(i)} - 1\right)_{i=0,\ldots,n}$ einen Vorzeichenwechsel aufweisen, woraus wegen der vorausgesetzten Monotonie

$$\frac{\nu(n)}{\mu(n)} < 1 < \frac{\nu(0)}{\mu(0)}$$

folgt. Daher existiert unter erneuter Nutzung der Monotonie ein $a \in \mathbb{N}$, $1 \leq a \leq n$, so dass

$$\nu(k) \begin{matrix} > \\ \leq \end{matrix} \mu(k), \quad \text{falls } k \begin{matrix} < \\ \geq \end{matrix} a$$

gilt. Gäbe es ein $0 \leq x \leq n$ mit $\sum_{k=0}^{x} \mu(k) > \sum_{k=0}^{x} \nu(k)$, so folgte $x \geq a$, was

$$\sum_{k=0}^{n} \mu(k) > \sum_{k=0}^{n} \nu(k) = 1$$

nach sich zöge. Das ist ein Widerspruch. Es folgt daher $\mu \leq \nu$. □

Schließlich erhalten wir hiermit

Satz 2.3.18. *Seien μ und ν zwei Verteilungen auf $\{0, 1, \ldots, n\}$ mit $\mu \leq \nu$. Dann gilt*

$$\|Q_\mu - U\| \leq \|Q_\nu - U\|.$$

Beweis. Mit Lemma 2.1.1 und der Abkürzung $c_l \stackrel{\text{def}}{=} |\{L = l\}|$, $l = 1, \ldots, n$ gilt

$$\begin{aligned}
&\sum_{\pi\in\mathfrak{S}_n} |Q_\nu(\pi) - U(\pi)| \\
&= \sum_{l=1}^{n} c_l \left| \sum_{m=n-l}^{n} \nu(m) \frac{(n-m)!}{n!} - \frac{1}{n!} \right| \\
&= \frac{1}{n!} \sum_{l=1}^{n} c_l \left| \sum_{m=n-l}^{n} (n-m)!\,\nu(m) - 1 \right|, \quad k \stackrel{\text{def}}{=} n-m \\
&= \frac{1}{n!} \sum_{l=1}^{n} c_l \left| \sum_{k=0}^{l} k!\,\nu(n-k) - 1 \right|. \qquad (2.114)
\end{aligned}$$

Sei $l_\rho \stackrel{\text{def}}{=} \min\{1 \le s \le n : \sum_{k=0}^{s} k!\,\rho(n-k) \ge 1\}$, wobei ρ eine Verteilung auf $\{0, \ldots, n\}$ ist. Mit dieser Bezeichnung formen wir (2.114) weiter um zu

$$\begin{aligned}
&\frac{1}{n!} \left(\sum_{l=l_\nu}^{n} c_l \sum_{k=0}^{l} k!\,\nu(n-k) - \sum_{l=l_\nu}^{n} c_l + \sum_{l=1}^{l_\nu-1} c_l - \sum_{l=1}^{l_\nu-1} c_l \sum_{k=0}^{l} k!\,\nu(n-k) \right) \\
&\ge \frac{1}{n!} \left(\sum_{l=l_\mu}^{n} c_l \sum_{k=0}^{l} k!\,\nu(n-k) - \sum_{l=l_\mu}^{n} c_l + \sum_{l=1}^{l_\mu-1} c_l - \sum_{l=1}^{l_\mu-1} c_l \sum_{k=0}^{l} k!\,\nu(n-k) \right), \qquad (2.115)
\end{aligned}$$

wobei letzte Ungleichung besteht, da für $l_\mu \neq l_\nu$ in (2.114) einige der Summanden

$$\left| \sum_{k=0}^{l} k!\,\nu(n-k) - 1 \right|_{l=1,\ldots,n}$$

eventuell mit negativem Vorzeichen aufsummiert werden, sich sonst aber nichts ändert. Im Folgenden wird der Term innerhalb der Klammern von (2.115) in die Form

$$\sum_{k=0}^{n} a(k)\nu(n-k) - \sum_{l=l_\mu}^{n} c_l + \sum_{l=1}^{l_\mu-1} c_l$$

gebracht, wobei eine Wahl der $a(k)$ nun bestimmt wird. Es gilt einerseits

$$\begin{aligned}
\sum_{l=l_\mu}^{n} c_l \sum_{k=0}^{l} k!\,\nu(n-k) &= \sum_{\substack{l\ge l_\mu \\ k\le l}} c_l k!\,\nu(n-k) \\
&= \sum_{k=0}^{l_\mu-1} \left(\sum_{l=l_\mu}^{n} c_l k! \right) \nu(n-k) + \sum_{k=l_\mu}^{n} \left(\sum_{l=k}^{n} c_l k! \right) \nu(n-k)
\end{aligned}$$

und andererseits

$$\sum_{l=1}^{l_\mu-1} c_l \sum_{k=0}^{l} k!\,\nu(n-k) = \sum_{\substack{l\le l_\mu-1 \\ k\le l}} c_l k!\,\nu(n-k)$$

$$= \sum_{k=0}^{l_\mu-1} \left(\sum_{l=k\vee 1}^{l_\mu-1} c_l k! \right) \nu(n-k).$$

Hieraus folgt für $k = 0, \ldots, l_\mu - 1$ als naheliegende Wahl

$$\begin{aligned} a(k) &= \sum_{l=l_\mu}^{n} c_l k! - \sum_{l=k\vee 1}^{l_\mu-1} c_l k! \\ &= k!(|\{L \geq l_\mu\}| - (|\{L \geq k\}| - |\{L \geq l_\mu\}|)) \\ &= k!(2|\{L \geq l_\mu\}| - |\{L \geq k\}|) \\ &= \frac{2n!}{l_\mu!} k! - n!. \end{aligned}$$

Für $k = l_\mu, \ldots, n$ erhalten wir

$$a(k) = \sum_{l=k}^{n} c_l k! = k!|\{L \geq k\}| = n!.$$

Wegen

$$a(l_\mu - 1) = \frac{2n!}{l_\mu!}(l_\mu - 1)! - n! \leq n! = a(l_\mu)$$

ist $k \mapsto a(k)$ monoton wachend. Ferner folgt aus $\mu \leq \nu$ leicht $\nu(n - \cdot) \leq \mu(n - \cdot)$, so dass nach Lemma 2.3.17

$$\begin{aligned} & \frac{1}{n!} \left(\sum_{k=0}^{n} a(k)\mu(n-k) - \sum_{l=l_\mu}^{n} c_l + \sum_{l=1}^{l_\mu-1} c_l \right) \\ = & \frac{1}{n!} \left(\sum_{l=l_\mu}^{n} c_l \sum_{k=0}^{l} k!\, \mu(n-k) - \sum_{l=l_\mu}^{n} c_l + \sum_{l=1}^{l_\mu-1} c_l - \sum_{l=1}^{l_\mu-1} c_l \sum_{k=0}^{l} k!\, \mu(n-k) \right) \\ = & \sum_{\pi \in \mathfrak{S}_n} |Q_\mu(\pi) - U(\pi)| \end{aligned}$$

eine untere Schranke von (2.115) ist, womit der Satz bewiesen ist. □

Bemerkung 2.3.19. Aus (2.114) folgt mit $\nu \stackrel{\text{def}}{=} \delta_m$, $0 \leq m \leq n$ leicht

$$\begin{aligned} \|Q_{\delta_m} - U\| &= \frac{1}{n!} \sum_{l=1}^{n} c_l \left| \sum_{k=0}^{l} k! \delta_m(n-k) - 1 \right| \\ &= \frac{1}{n!} \sum_{l=n-m}^{n} c_l((n-m)! - 1) \\ &= \frac{|\{L \geq n-m\}|}{n!}((n-m)! - 1) \\ &= 1 - \frac{1}{(n-m)!}. \end{aligned}$$

Wir bemerken weiter, dass Satz 2.3.18 auch für den Beweis von Satz 2.5.18 benötigt wird.

Korollar 2.3.20. *Es gilt*

$$\|Q_{\delta_m}^{*k} - U\| \leq \|Q_{\delta_1}^{*k\cdot m} - U\|, \quad m,k \geq 1. \tag{2.116}$$

Beweis. Nach den Sätzen 2.1.4 und 2.3.18 bleibt noch

$$\overbrace{\delta_m \sharp \ldots \sharp \delta_m}^{k-\text{mal}} \leq \overbrace{\delta_1 \sharp \ldots \sharp \delta_1}^{k\cdot m-\text{mal}}$$

zu zeigen. Das ist aber klar, da für jedes $0 \leq l \leq n$ die Wahrscheinlichkeit, mit $k \cdot m$ sukzessiven, voneinander unabhängigen, rein zufälligen Bestückungen einer n-Zellenkonstellation *nicht* mehr als l Zellen zu besetzen, größer ist als die Wahrscheinlichkeit, dieses mit k sukzessiven, voneinander unabhängigen Bestückungen mit jeweils m Kugeln zu erreichen, wobei diese m Kugeln in jedem Schritt in m *verschiedene* Zellen gelegt werden müssen, die Besetzungen ansonsten aber willkürlich sind. Als formales Argument verweisen wir den Leser auf den Anfang des Beweises von Satz 2.3.15. Mit $\mu \stackrel{\text{def}}{=} \delta_m$ erhalten wir in der dortigen Notation

$$N(k\cdot m) \sim \overbrace{\delta_1 \sharp \ldots \sharp \delta_1}^{k\cdot m-\text{mal}}, \quad N(W_1 + \ldots + W_k) \sim \overbrace{\delta_m \sharp \ldots \sharp \delta_m}^{k-\text{mal}}.$$

Wegen

$$W_i \geq m, \quad 1 \leq i \leq k \implies W_1 + \ldots + W_k \geq k \cdot m$$

und da $t \mapsto N(t)$ monoton wachsend ist, erhalten wir für $0 \leq l \leq n$

$$\begin{aligned} \overbrace{\delta_m \sharp \ldots \sharp \delta_m(\{0,\ldots,l\})}^{k-\text{mal}} &= P(N(W_1 + \ldots + W_k) \leq l) \\ &\leq P(N(k\cdot m) \leq l) \\ &= \underbrace{\delta_1 \sharp \ldots \sharp \delta_1}_{k\cdot m-\text{mal}}(\{0,\ldots,l\}). \end{aligned}$$

Dies war zu zeigen. □

Wir werden als weiteres Korollar zu Satz 2.3.18 noch beweisen, dass $\lambda \mapsto v(\lambda)$ eine auf $(0,\infty)$ monoton wachsende Funktion ist. Selbstverständlich lässt sich dieses auch direkt nachrechnen. Hierzu wird (2.86) auf den Intervallen $(\lambda_l, \lambda_{l+1})$, $l \geq 1$ nach λ abgeleitet und gezeigt, dass die Ableitung nichtnegativ ist. Unter Benutzung der in Lemma 2.3.14 gezeigten Stetigkeit in $(\lambda_l)_{l\geq 2}$ folgt dann sofort die Monotonie. Folgendes ist daher eher als ein interessantes Spiel mit unserer Theorie zu sehen. Dies ziehen wir den Standardrechnungen vor.

Korollar 2.3.21. *Die Funktion $\lambda \mapsto v(\lambda)$, $\lambda \in (0,\infty)$ ist monoton wachsend. v ist hierbei wie in (2.86) definiert.*

Beweis. Wir definieren Verteilungen μ_n^λ für $n \in \mathbb{N}$, $\lambda \in (0,\infty)$ durch

$$\mu_n^\lambda(n-k) \stackrel{\text{def}}{=} c(n,\lambda)\frac{\lambda^k}{k!}, \quad k = 0,\ldots,n$$

auf $\{0,\ldots,n\}$ mit

$$c(n,\lambda) \stackrel{\text{def}}{=} \left(\sum_{k=0}^{n} \frac{\lambda^k}{k!}\right)^{-1}.$$

Es gilt offenbar für fixierte $k \in \mathbb{N}_0$, $\lambda \in (0, \infty)$

$$\mu_n^\lambda(n-k) \to e^{-\lambda}\frac{\lambda^k}{k!}, \quad n \to \infty.$$

Nach Theorem 2.3.4 haben wir folglich

$$v(\lambda) = \lim_{n\to\infty} \|Q_{\mu_n^\lambda} - U\|, \quad \lambda \in (0, \infty).$$

Sei $0 < \lambda_1 < \lambda_2$. Da $\mu_n^{\lambda_2}(n-k) > 0$, $k = 0, \ldots, n$ und

$$k \mapsto \frac{\mu_n^{\lambda_1}(n-k)}{\mu_n^{\lambda_2}(n-k)} = \frac{\sum_{i=0}^n \frac{\lambda_2^i}{i!}}{\sum_{i=0}^n \frac{\lambda_1^i}{i!}} \left(\frac{\lambda_1}{\lambda_2}\right)^k$$

monoton fallend ist, folgt aus Lemma 2.3.17 $\mu_n^{\lambda_2}(n-\cdot) \le \mu_n^{\lambda_1}(n-\cdot)$, d.h. $\mu_n^{\lambda_1} \le \mu_n^{\lambda_2}$. Satz 2.3.18 besagt dann

$$\|Q_{\mu_n^{\lambda_1}} - U\| \le \|Q_{\mu_n^{\lambda_2}} - U\|,$$

woraus durch Grenzübergang die Behauptung folgt. □

Wir werden schließlich noch einen *Cutoff-Effekt* des T1TRS mit Hilfe des Prinzips der *stark stationären Zeiten* (siehe Satz 1.2.13) nachweisen. Die hier konstruierte *stark stationäre Zeit* hat eine sehr anschauliche Bedeutung, und alles Notwendige lässt sich relativ leicht berechnen. Allerdings erhalten wir auch nicht so subtile Aussagen wie bei dem direkten Ansatz, der uns z.B. die Grenzvariation $f(c) = v(e^{-c})$ liefert (siehe Theorem 2.3.7).

Die folgenden Ausführungen basieren auf Aldous, Diaconis [1].

Theorem 2.3.22. *Für den T1TRS gilt mit $d(\cdot)$ wie in (1.11) bzgl. der Gleichverteilung auf $\mathfrak{S}_n$ als stationäres Maß*

(i) $d(n\log n + cn) \le e^{-c}, \quad c \ge 0, n \ge 1,$

(ii) $d(n\log n - c_n n) \to 1, \quad n \to \infty,\ (c_n)_{n\in\mathbb{N}}, c_n \to \infty,$

wobei zusätzlich $c, (c_n)_{n\in\mathbb{N}}$ derart gewählt sei, dass die Argumente in $d(\cdot)$ natürliche Zahlen sind.

Für den Beweis des Theorems benötigen wir zunächst ein Lemma. Sei eine n Zellenanordnung gegeben. In diese werden sukzessive Kugeln gelegt. Genauer wird in jedem Schritt genau eine Kugel unabhängig von allen anderen Schritten in die Zellenanordnung gelegt, wobei jede Zelle mit derselben Wahrscheinlichkeit besetzt werde. Die Zufallsvariable V beschreibe den ersten Schritt, bei dem alle Zellen besetzt sind.

Lemma 2.3.23. *Für $c \ge 0$ und $n \in \mathbb{N}$ mit $n\log n + cn \in \mathbb{N}$ gilt*

$$P(V > n\log n + cn) \le e^{-c}.$$

Beweis. Sei $k \stackrel{\text{def}}{=} n\log n + cn \in \mathbb{N}$ und A_i, $1 \le i \le n$ das Ereignis, dass Zelle i *nicht* innerhalb der ersten k Schritte besetzt wird. Es gilt dann

$$P(V > k) = P\left(\bigcup_{i=1}^n A_i\right) \le \sum_{i=1}^n P(A_i) = n\left(1 - \frac{1}{n}\right)^k \le ne^{-k/n} = e^{-c},$$

wobei die zweite Ungleichheit wegen

$$\left(1-\frac{1}{n}\right)^n \leq e^{-1}, \quad n \in \mathbb{N} \tag{2.117}$$

gilt. Für $n = 1$ ist (2.117) trivial, und für $n \geq 2$ ist (2.117) äquivalent mit

$$n \log\left(1-\frac{1}{n}\right) \leq -1 \iff \log\left(1-\frac{1}{n}\right) \leq -\frac{1}{n}, \quad n \geq 2,$$

was sofort aus der geläufigen Ungleichung $\log(1+x) \leq x$, $x \in (-1, \infty)$ folgt. □

Wir werden nun mittels einiger Prosa eine *stark stationäre Zeit* für den T1TRS konstruieren: Gegeben sei ein vorsortiertes Kartendeck ($\mathrm{id} \in \mathfrak{S}_n$). Es wird im Folgenden bei jedem T1TRS Mischvorgang immer die *unterste* Karte des ursprünglich vorsortierten Kartendecks (Karte n) näher betrachtet. Diese wird hierzu vor Beginn der Mischiterationen markiert. Sei T_1 der erste Zeitpunkt (oder Mischschritt), bei dem innerhalb des T1TRS die oberste Karte unter die markierte Karte gelegt wird. Sei T_2 der erste Zeitpunkt, bei dem sich unter der markierten Karte genau zwei Karten befinden. Offensichtlich ist zu diesem Zeitpunkt jede der beiden Anordnungsmöglichkeiten der unteren beiden Karten gleichwahrscheinlich. Das Verfahren wird nun konsequent fortgesetzt, d.h. $T_i > T_{i-1}$, $3 \leq i \leq n-1$ beschreibe den ersten Zeitpunkt, bei dem sich unter der markierten Karte genau i Karten befinden. Auch hier sind alle $i!$ möglichen Permutationen offenbar gleichwahrscheinlich. Zum Zeitpunkt $T \stackrel{\text{def}}{=} T_{n-1} + 1$, d.h. wenn die ursprünglich unterste Karte (Karte n) oben liegt und danach an eine rein zufällige Stelle untergemischt wird, ist schließlich jede Kartenordnung gleichwahrscheinlich oder in Formeln:

$$P(X_k = \pi | T = k) = \frac{1}{n!}, \quad k \in \mathbb{N}_0, \pi \in \mathfrak{S}_n,$$

wobei $(X_k)_{k\geq 0}$ mit $X_0 \stackrel{\text{def}}{=} \mathrm{id}$ einen *Random Walk auf* $\mathfrak{S}_n$ im Sinne von Abschnitt 1.3 beschreibt und Q in (1.12) gleich der Verteilung Q_1 in (2.3) ($m = 1$) ist. Ferner ist das Ereignis $\{T = k\}$ bis zum Zeitpunkt k entscheidbar, weshalb T eine Stoppzeit ist. Nach Beispiel 1.3.2 ist $(X_k)_{k\geq 0}$ eine *irreduzible*, endliche Markov-Kette. Wegen der Endlichkeit des Zustandraumes ist diese auch rekurrent, so dass jeder Zustand nach endlich vielen Schritten P-f.s. erreicht wird, siehe z.B. [5, S.54]. Dies impliziert die P-f.s. Endlichkeit von T, so dass T im Sinne von Definition 1.2.10 eine eine *stark stationäre Zeit* ist.

Heuristisch können wir schon an dieser Stelle konstatieren, dass es nicht mehr als eine Größenordnung von $n \log n$ Mischschritten bedarf, bis das Kartendeck durchmischt ist, da

$$T = (T - T_{n-1}) + (T_{n-1} - T_{n-2}) + \ldots + T_1 \tag{2.118}$$

gilt und wir mit $T_n \stackrel{\text{def}}{=} T$, $T_0 \stackrel{\text{def}}{=} 0$ wegen $T_i - T_{i-1} \sim \mathrm{Geo}\left(\frac{i}{n}\right)$, $i = 1, \ldots, n$ den Erwartungswert

$$E(T) = \sum_{i=1}^{n} \frac{n}{i} = n \sum_{i=1}^{n} \frac{1}{i} = n(\log n + O(1)), \quad n \to \infty$$

erhalten, wobei wir in der letzten Gleichung auf (2.121) vorgegriffen haben. Wir sind nun in der Lage Theorem 2.3.22 zu beweisen.

Beweis. (i) Nach Lemma 2.3.23 und Satz 1.2.13 zusammen mit (1.9) reicht es zu zeigen, dass T dieselbe Verteilung wie V besitzt. Sei V_i, $1 \leq i \leq n$ der erste Zeitpunkt, bei dem genau i Zellen der n Zellenkonstellation in Lemma 2.3.23 besetzt sind. Es gilt

$$V = (V - V_{n-1}) + (V_{n-1} - V_{n-2}) + \ldots + V_1. \tag{2.119}$$

Wenn genau $1 \leq i \leq n$ Zellen besetzt sind, beträgt die Wahrscheinlichkeit, eine unbesetzte Zelle zu belegen, $\frac{n-i}{n}$, woraus wegen der vorausgesetzten Unabhängigkeit der jeweiligen Belegungen

$$P(V_{i+1} - V_i = j) = \frac{n-i}{n}\left(1 - \frac{n-i}{n}\right)^{j-1}, \quad j \geq 1,\, i = 1, \ldots, n-1$$

folgt. Vergleichen wir dies mit den Ausführungen direkt vor diesem Beweis, so folgt

$$T_{i+1} - T_i \sim V_{n-i} - V_{n-i-1}, \quad i = 1, \ldots, n-2.$$

Da zusätzlich noch

$$T - T_{n-1} = V_1 = 1, \quad T_1 \sim V - V_{n-1}$$

gilt und die Summanden in (2.118) und (2.119) jeweils unabhängig sind, ist somit (i) bewiesen.

(ii) Sei $A_j \stackrel{\text{def}}{=} \{L \geq j\}$, $1 \leq j \leq n$ mit L wie im Lemma 2.1.1. Es gilt dann $U(A_j) = \frac{1}{j!}$. Es reicht mit $k_n \stackrel{\text{def}}{=} n \log n - c_n n$ und $j \geq 1$ fixiert

$$Q^{*k_n}(A_j) \to 1, \quad n \to \infty$$

zu zeigen, da hieraus

$$d(k_n) \geq \max_{j=1,\ldots,n} (Q^{*k_n}(A_j) - U(A_j)) \to 1, \quad n \to \infty$$

folgt. Um dieses nachzuweisen, mache man sich die Ungleichung

$$Q^{*k_n}(A_j) \geq P(T - T_{j-1} > k_n), \quad j \geq 2$$

klar: $T - T_{j-1}$ ist so verteilt wie die Anzahl der Schritte, bis die $(n-j+1)$-te Karte (die j-te Karte von unten) oben auf das Kartendeck gekommen ist und dann wieder eingefügt wird. Falls dies nicht bis zum Zeitpunkt k_n passiert ist, müssen die anfänglich j untersten Karten immer noch in unveränderter relativer Ordnung zueinander stehen. Es bleibt für fixiertes $j \geq 1$ folglich noch

$$P(T - T_j \leq k_n) \to 0, \quad n \to \infty$$

zu zeigen. Wir werden hierzu die Chebychev Ungleichung benutzen. Es ist bekannt, dass

$$T_{i+1} - T_i \sim \text{Geo}\left(\frac{i+1}{n}\right), \quad 1 \leq i \leq n-1$$

gilt, woraus

$$E(T_{i+1} - T_i) = \frac{n}{i+1}, \quad \text{Var}(T_{i+1} - T_i) = \left(\frac{n}{i+1}\right)^2 \left(1 - \frac{i+1}{n}\right), \quad 1 \leq i \leq n-1$$

folgt. Wegen

$$T - T_j = (T_n - T_{n-1}) + (T_{n-1} - T_{n-2}) + \ldots + (T_{j+1} - T_j), \quad 1 \leq j \leq n-1$$

und der Unabhängigkeit der einzelnen Summanden haben wir daher für $1 \leq j \leq n-1$

$$E(T - T_j) = \sum_{i=j}^{n-1} \frac{n}{i+1}, \quad \mathrm{Var}(T - T_j) = \sum_{i=j}^{n-1} \left(\frac{n}{i+1}\right)^2 \left(1 - \frac{i+1}{n}\right). \tag{2.120}$$

Wegen

$$\sum_{i=1}^{n-1} \frac{1}{i+1} \leq \int_1^n \frac{dx}{x} = \log n \leq \sum_{i=1}^{n-1} \frac{1}{i}$$

folgt

$$0 \leq \log n - \sum_{i=1}^{n-1} \frac{1}{i+1} \leq 1 - \frac{1}{n},$$

woraus sich sofort

$$O(1) + \log n = \sum_{i=j}^{n-1} \frac{1}{i+1}, \quad n \to \infty \ (j \text{ fixiert}) \tag{2.121}$$

ergibt. Ferner haben wir

$$\sum_{i=j}^{n-1} \frac{n-i-1}{(i+1)^2 n} \leq \sum_{i=1}^{n-1} \frac{n-i}{i^2 n} \leq \sum_{i=1}^{\infty} \frac{1}{i^2} < \infty,$$

woraus wir mit (2.120) für fixiertes $j \in \mathbb{N}$

$$E(T - T_j) = n \log n + O(n), \quad \mathrm{Var}(T - T_j) = O(n^2), \quad n \to \infty$$

erhalten. Mit Hilfe der Chebychev Ungleichung folgt hieraus

$$\begin{aligned} P(T - T_j \leq k_n) &= P(T_j - T \geq -k_n) \\ &= P(T_j - T - E(T_j - T) \geq E(T - T_j) - k_n) \\ &\leq P(|T_j - T - E(T_j - T)| \geq E(T - T_j) - k_n) \\ &\leq \frac{\mathrm{Var}(T_j - T)}{(E(T_j - T) - k_n)^2} \\ &= \frac{O(n^2)}{(O(n) + c_n n)^2}. \end{aligned}$$

Es bleibt daher noch

$$\frac{O(n^2)}{(O(n) + c_n n)^2} \to 0, \quad n \to \infty \tag{2.122}$$

zu zeigen. Unter Berücksichtigung von $c_n \to \infty$, $n \to \infty$ und der Definition von der Landauordnungsklasse O existieren reelle Zahlen $\xi_1, \xi_2 > 0$ und eine natürliche Zahl N, so dass für alle $n \geq N$

$$|O(n)| \leq \xi_1 n, \quad |O(n^2)| \leq \xi_2 n^2, \quad c_n \geq 2\xi_1 > 0$$

gilt. Hieraus folgt für den Nenner in (2.122) für $n \geq N$

$$\begin{aligned}(O(n)+c_n n)^2 &\geq 2c_n O(n)n + c_n^2 n^2 \\ &\geq -2c_n\xi_1 n^2 + c_n^2 n^2 \\ &= c_n(c_n - 2\xi_1)n^2 \geq 0.\end{aligned}$$

Insgesamt ergibt sich damit

$$\left|\frac{O(n^2)}{(O(n)+c_n n)^2}\right| \overset{(n\geq N)}{\leq} \frac{\xi_2 n^2}{c_n(c_n-2\xi_1)n^2} = \frac{\xi_2}{c_n(c_n-2\xi_1)} \to 0, \quad n\to\infty,$$

womit der Beweis vollbracht ist. □

Zum Abschluss dieses Abschnitts wird noch der *Cutoff-Effekt* bzgl. der Separation nachgewiesen. Dieses erweist sich im Vergleich zum *Cutoff-Effekt* bzgl. des Variationsabstandes als weniger Aufwendig, da Theorem 2.3.4 durch folgendes einfaches Lemma ersetzt werden kann:

Lemma 2.3.24. *Mit denselben Voraussetzungen wie in Theorem 2.3.4 gilt*

$$\mathrm{sep}(Q_{\mu_n}, U) = 1 - e^{-\lambda}(1+\lambda) + o(1), \quad n\to\infty. \tag{2.123}$$

Beweis. Sei

$$\tau \overset{\mathrm{def}}{=} \begin{pmatrix} 1 & 2 & \dots & n \\ n & n-1 & \dots & 1 \end{pmatrix} \in \mathfrak{S}_n\,.$$

Offenbar gilt in der Notation von Lemma 2.1.1 $L(\tau) = 1$. Aus (2.3) erhält man daher ohne weiteren Aufwand sofort

$$Q_\mu(\tau) = \min_{\pi\in\mathfrak{S}_n}\{Q_\mu(\pi)\},$$

woraus

$$\begin{aligned}\mathrm{sep}(Q_{\mu_n}, U) &= 1 - n!\sum_{m=0}^{n}\mu(m)Q_m(\tau) \\ &= 1-\mu(n-1)-\mu(n) \\ &= 1 - e^{-\lambda}(1+\lambda)+o(1)\end{aligned}$$

folgt. □

Nach einer kurzen Betrachtung des Beweises von Satz 2.3.15 erhalten wir daher u.a. Informationen über einen *Cutoff-Effekt* bzgl. der Separation des TmTRS, sofern wir $v(e^{-c})$ in (2.107) durch $1 - e^{-e^{-c}}(1+e^{-c})$ ersetzen. Es zeigt sich, dass ein *Cutoff-Effekt* genau wie bei dem Variationsabstand wieder nach $\frac{n}{m}(\log n + c)$ Mischvorgängen erscheint. Diesbezüglich besteht hier zwischen Variation und Separation folglich *kein* Unterschied. Bei dem *Riffle-Shuffle* ist das z.B. nicht der Fall: vgl. Satz 2.5.13 mit Theorem 2.5.16.

2.4 Spektralanalyse und Algebren

Wir beginnen mit einem Lemma aus der linearen Algebra, das im Beweis von Satz 2.4.3 benötigt wird.

Lemma 2.4.1. *Folgendes gilt:*

(i) Es seien $r \geq 1$ paarweise verschiedene komplexe Zahlen $c_1, \ldots, c_r$ gegeben. Dann besitzt das lineare Gleichungssystem

$$\sum_{j=1}^{r} \alpha_j c_j^i = 0, \quad i = 0, \ldots, r-1 \tag{2.124}$$

mit den Variablen $\alpha_j \in \mathbb{C}$, $j = 1, \ldots, r$ die eindeutige Lösung

$$\alpha_1 = \ldots = \alpha_r = 0.$$

(ii) Gegeben sei eine Matrix $A \in \mathbb{C}^{n\times n}$. Falls $s \geq 1$ Matrizen

$$Z_i \in \mathbb{C}^{n\times n} - \{0\}, \quad i = 1, \ldots, s$$

und s paarweise verschiedene komplexe Zahlen $\eta_1, \ldots, \eta_s$ existieren, so dass die Gleichung

$$A^k = \sum_{i=1}^{s} \eta_i^k Z_i$$

für alle $k \in \mathbb{N}_0$ erfüllt ist, so ist A diagonalisierbar und besitzt genau die Eigenwerte $\eta_1, \ldots, \eta_s$.

Beweis. (i): Sei $C \in \mathbb{C}^{r\times r}$ durch $C_{ij} \stackrel{\text{def}}{=} c_j^{i-1}$, $1 \leq i, j \leq r$ definiert. (2.124) entspricht in vektorieller Schreibweise dann der Gleichung

$$C\alpha = 0 \quad \text{mit} \quad \alpha \stackrel{\text{def}}{=} (\alpha_1, \ldots, \alpha_r)^t. \tag{2.125}$$

Es gilt nach Fischer [10, S.196]

$$\det C = \prod_{1\leq i<j\leq r} (c_j - c_i) \neq 0 \quad \text{(Vandermonde Determinante)}.$$

C ist folglich invertierbar, weshalb (2.125) die *eindeutige* Lösung $\alpha = 0$ besitzt, womit (i) gezeigt ist.

(ii): Als Erstes wird gezeigt, dass $\eta_1, \ldots, \eta_s$ genau die Eigenwerte von A sind. Seien hierzu $\lambda_1, \ldots, \lambda_l$ die paarweise verschiedenen Eigenwerte von A und $v_1, \ldots, v_l$ eine Wahl zugehöriger Eigenvektoren. Für beliebiges, aber fixiertes $1 \leq u \leq l$ gilt dann

$$\lambda_u^k v_u = A^k v_u = \sum_{i=1}^{s} \eta_i^k Z_i v_u, \quad k \in \mathbb{N}_0,$$

woraus

$$\sum_{i=1}^{s} \eta_i^k (Z_i v_u)^{(\iota)} + \lambda_u^k (-v_u)^{(\iota)} = 0, \quad 1 \leq \iota \leq n,\ k = 0, 1, 2, \ldots, s$$

folgt. Hierbei sind mit den hochgestellten $^{(\iota)}$ die jeweiligen Koordinaten gemeint. Falls $\lambda_u \notin \{\eta_1, \ldots, \eta_s\}$, existierte wegen (i) für jedes $1 \leq \iota \leq n$ nur die triviale Lösung. Es ergäbe sich folglich $v_u = 0$, was nicht sein kann, da v_u ein Eigenvektor ist. Hiermit haben wir $\{\lambda_1, \ldots, \lambda_l\} \subseteq \{\eta_1, \ldots, \eta_s\}$ gezeigt. Für die umgekehrte Inklusion wird zuerst die Idempotenz von Z_i, $i = 1, \ldots, s$ nachgewiesen, genauer sogar

$$Z_i Z_j = \delta_{ij} Z_i, \quad 1 \leq i, j \leq s. \tag{2.126}$$

Sei $\nu \in \mathbb{N}_0$ fixiert. Dann gilt für alle $k \in \mathbb{N}_0$

$$A^k A^\nu = \sum_{i=1}^{s} \eta_i^k \eta_i^\nu Z_i = \left(\sum_{i_1=1}^{s} \eta_{i_1}^k Z_{i_1} \right) \left(\sum_{i_2=1}^{s} \eta_{i_2}^\nu Z_{i_2} \right) = \sum_{i_1, i_2} \eta_{i_1}^k \eta_{i_2}^\nu Z_{i_1} Z_{i_2}.$$

Dies impliziert

$$\sum_{i=1}^{s} \eta_i^k \left(\eta_i^\nu Z_i - \sum_{j=1}^{s} \eta_j^\nu Z_i Z_j \right) = 0, \quad k = 0, 1, \ldots, s-1.$$

Nach komponentenweiser Betrachtung obiger Matrizen liefert eine erneute Anwendung von (i)

$$\eta_i^\nu Z_i = \sum_{j=1}^{s} \eta_j^\nu Z_i Z_j, \quad i = 1, \ldots, s.$$

Da $\nu \in \mathbb{N}_0$ beliebig war, folgt hieraus

$$\eta_i^\nu (Z_i Z_i - Z_i) + \sum_{\substack{j=1 \\ j \neq i}}^{s} \eta_j^\nu Z_i Z_j = 0, \quad \nu = 0, \ldots, s-1,\ i = 1, \ldots, s,$$

wobei wir wieder nur die für uns interessanten ν angegeben haben. Unter erneuter Berücksichtigung von (i) erhalten wir schließlich (2.126). Da Z_i idempotent ist, gilt für jeden Eigenvektor $v \neq 0$ zu einem Eigenwert λ

$$Z_i v = \lambda v \Rightarrow \lambda v = Z_i v = Z_i^2 v = \lambda^2 v \Rightarrow \lambda \in \{0, 1\}.$$

Besäße Z_i nur Null als Eigenwert, so folgte aus der Jordanschen Normalform von Z_i sofort die Nilpotenz von Z_i, die einen Widerspruch zur Idempotenz darstellte, da $Z_i \neq 0$ (vgl. auch Fischer [10, S.257]). Sei v also ein Eigenvektor zum Eigenwert Eins von Z_i. Dann gilt

$$Av = AZ_i v = \left(\sum_{j=1}^{s} \eta_j Z_j \right) Z_i v = \left(\sum_{j=1}^{s} \eta_j Z_j Z_i \right) v = \eta_i Z_i v = \eta_i v,$$

weshalb η_i ein Eigenwert ist. Damit ist die umgekehrte Inklusion gezeigt, und wir haben die Darstellung ($l = s$)

$$A^k = \sum_{i=1}^{l} \lambda_i^k Z_i, \quad k \in \mathbb{N}_0.$$

Als Nächstes wird die Diagonalisierbarkeit von A nachgewiesen. Wir zeigen dies-

bezüglich zuerst, dass es ausreicht, die Diagonalisierbarkeit von

$$C \stackrel{\text{def}}{=} \sum_{i=1}^{l} iZ_i$$

nachzuweisen: Wegen (2.126) gilt

$$C^k = \sum_{i=1}^{l} i^k Z_i, \quad k \in \mathbb{N}_0.$$

C besitzt nach dem soeben Gezeigten genau die Eigenwerte $1, \ldots, l$. Wir zeigen, dass für $v \in \mathbb{C}^n$ folgende Äquivalenz besteht:

$$Cv = jv \iff Av = \lambda_j v, \quad j = 1, \ldots, l.$$

$Cv = jv$ für ein beliebig fixiertes $1 \leq j \leq l$ impliziert

$$C^k v = j^k v = \sum_{i=1}^{l} i^k Z_i v, \quad k \in \mathbb{N}_0,$$

woraus

$$\sum_{\substack{i=1 \\ i \neq j}}^{l} i^k Z_i v + j^k (Z_j v - v) = 0, \quad k = 0, \ldots, l-1$$

folgt. Unter Berücksichtigung von (i) ergibt dies

$$Z_i v = \begin{cases} v & \text{falls } i = j, \\ 0 & \text{sonst,} \end{cases}$$

was sofort $Av = \lambda_j v$ nach sich zieht. Die umgekehrte Richtung lässt sich genauso zeigen. A besitzt daher genau dann eine Basis von Eigenvektoren, wenn auch C eine solche Basis besitzt. Es reicht folglich, die Diagonalisierbarkeit von C nachzuweisen, m.a.W. kann wie behauptet o.E. $\lambda_i = i$, $1 \leq i \leq l$ angenommen werden.

Sei

$$J^k = SA^kS^{-1} = \sum_{i=1}^{l} i^k SZ_iS^{-1}, \quad k \in \mathbb{N}_0$$

für passendes $S \in \mathbb{C}^{n \times n}$ die Jordansche Normalform ($k = 1$) zu A. Diese ist bekanntlich bis auf Permutationen der Jordanblöcke eindeutig. Mit $E_i \stackrel{\text{def}}{=} SZ_iS^{-1}$ gilt

$$J^k = \sum_{i=1}^{l} i^k E_i, \quad k \in \mathbb{N}_0. \tag{2.127}$$

Falls A nicht diagonalisierbar ist, so existiert für ein $1 \leq \lambda \leq l$, $\lambda \in \mathbb{N}$ ein Jordanblock der Form

$$B \stackrel{\text{def}}{=} \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{C}^{t \times t}, \quad t \geq 2. \tag{2.128}$$

Mit $B = \lambda I_t + N_t$, wobei I_t die Identität auf $\mathbb{R}^t$ bezeichne und N_t über diese Gleichung definiert werde, gilt

$$B^k = \sum_{u=0}^{k} \binom{k}{u} \lambda^u N_t^{k-u}, \quad k \in \mathbb{N}_0.$$

Es wird nun das Element $\left(B^k\right)_{1,2} = \binom{k}{k-1}\lambda^{k-1} = k\lambda^{k-1}$ betrachtet. Da (2.127) gilt, folgte unter der Tatsache, dass die k-te Potenz von J gebildet wird, indem die k-ten Potenzen der jeweiligen Jordanblöcke gebildet werden

$$k\lambda^{k-1} = \sum_{i=1}^{l} i^k (E_i)_{1,2} = \sum_{i=1}^{l} c_i i^k, \quad c_i \stackrel{\text{def}}{=} (E_i)_{1,2} \in \mathbb{C},\ k \in \mathbb{N}, \tag{2.129}$$

wobei hier o.E. B als der erste Jordanblock von J angenommen wurde. Sei $1 \leq u \leq l$ maximal mit $c_u \neq 0$. Ein solches u existiert immer, da aus $c_1 = \ldots = c_l = 0$ mit $k \stackrel{\text{def}}{=} 1$ in (2.129) der Widerspruch $1 = 0$ folgt. Wir haben also

$$k\lambda^{k-1} = \sum_{i=1}^{u} c_i i^k, \quad c_i \in \mathbb{C},\ c_u \neq 0,\ k \in \mathbb{N}.$$

Hieraus ergibt sich

$$\frac{1}{c_u \lambda} = \frac{1}{k} \sum_{i=1}^{u} \frac{c_i}{c_u} \left(\frac{i}{\lambda}\right)^k = \frac{1}{k} \left(\frac{u}{\lambda}\right)^k \left(1 + \sum_{i=1}^{u-1} \frac{c_i}{c_u} \left(\frac{i}{u}\right)^k\right), \quad k \in \mathbb{N},$$

woraus

$$0 < \frac{1}{|c_u|\lambda} = \frac{1}{k} \left(\frac{u}{\lambda}\right)^k \overbrace{\left|1 + \sum_{i=1}^{u-1} \frac{c_i}{c_u} \left(\frac{i}{u}\right)^k\right|}^{\to 1,\, k \to \infty}, \quad k \in \mathbb{N} \tag{2.130}$$

folgt. Falls $\lambda \geq u$, erhalten wir in (2.130) durch Grenzübergang $k \to \infty$ den Widerspruch $\frac{1}{|c_u|\lambda} = 0$ und im Fall $\lambda < u$ wegen

$$\frac{1}{k} \left(\frac{u}{\lambda}\right)^k = e^{-\log k + k \log\left(\frac{u}{\lambda}\right)} = e^{k\left(\overbrace{\log\left(\frac{u}{\lambda}\right)}^{>0} - \overbrace{\frac{\log k}{k}}^{\to 0}\right)} \to \infty, \quad k \to \infty$$

den Widerspruch $\frac{1}{|c_u|\lambda} = \infty$. Mit (2.127) kann folglich kein B in der Form von (2.128) Bestandteil von J sein, weshalb J in Diagonalform ist. Da J und A ähnlich sind, ist A per Definition diagonalisierbar. □

Bemerkung 2.4.2. Für eine Matrix $A \in \mathbb{C}^{n \times n}$ seien die Voraussetzungen von Lemma 2.4.1 (ii) erfüllt. Wir wollen noch die Gestalt der Z_i in diesem Lemma näher beleuchten: Es ist A diagonalisierbar und die Jordansche Normalform folglich in Diagonalgestalt. Deshalb kann

$$J = SAS^{-1} = \sum_{i=1}^{s} \eta_i \widetilde{E_i} \quad \text{mit } (\widetilde{E_i})_{u,v} = \begin{cases} 1 & \text{falls} \quad u = v \text{ und } \kappa_i^{(1)} \leq u \leq \kappa_i^{(2)} \\ 0 & \text{sonst} \end{cases}$$

geschrieben werden, wobei hier o.E.

$$1 = \kappa_1^{(1)} < \kappa_1^{(2)} + 1 = \kappa_2^{(1)} < \kappa_2^{(2)} + 1 = \kappa_3^{(1)} < \kappa_3^{(2)} + 1 \ldots = \kappa_s^{(1)} < \kappa_s^{(2)} + 1 = n + 1$$

angenommen werden kann. Insbesondere folgt mit $E_i = SZ_iS^{-1}$

$$J^k = \sum_{i=1}^{s} \eta_i^k \widetilde{E_i} = \sum_{i=1}^{s} \eta_i^k E_i, \quad k \in \mathbb{N}_0,$$

was

$$\sum_{i=1}^{s} \eta_i^k (\widetilde{E_i} - E_i) = 0, \quad k = 0, \ldots, s-1$$

ergibt, woraus mit Lemma 2.4.1 (i) nach komponentenweiser Betrachtung $E_i = \widetilde{E_i}$ gefolgert werden kann. Wegen

$$SZ_iS^{-1} = E_i = \widetilde{E_i}$$

folgt hieraus

$$Z_i = S^{-1}\widetilde{E_i}S,$$

was auch unserer Intuition entspricht.

Wir bemerken weiter, dass in Lemma 2.4.1 (ii) ebenfalls die Umkehrrichtung gilt.

Satz 2.4.3. *Sei $\mathbb{P}_j$ die Übergangsmatrix des TjTRS, $0 \le j \le n$ und $\mathbb{P} \stackrel{\text{def}}{=} \mathbb{P}_1$. Dann gilt:*

(a) $\mathbb{P}$ ist diagonalisierbar und besitzt genau die n paarweise verschiedenen Eigenwerte $0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-2}{n}, 1$. Der Eigenwert $\frac{i}{n}$ besitzt ferner dieselbe Multiplizität, wie es Permutationen aus $\mathfrak{S}_n$ gibt, die genau i Fixpunkte besitzen.

(b) $\mathbb{P}$ besitzt die spektrale Zerlegung

$$\mathbb{P}^k = \sum_{i=0}^{n} \lambda_i^k Z_i = \sum_{\substack{i=0 \\ i \neq n-1}}^{n} \lambda_i^k Z_i, \quad k \in \mathbb{N}_0 \tag{2.131}$$

mit

$$\lambda_i \stackrel{\text{def}}{=} \frac{i}{n}, \; Z_i \stackrel{\text{def}}{=} \sum_{j=i}^{n} (-1)^{j-i} \binom{j}{i} \binom{n}{j} \mathbb{P}_j, \quad 0 \le i \le n \tag{2.132}$$

und

$$\{0 \le i \le n \,:\, Z_i = 0\} = \{n-1\}. \tag{2.133}$$

Beweis. Wegen Lemma 2.4.1 folgt die Behauptung (a), bis auf die Multiplizität der Eigenwerte $\frac{i}{n}$, $i \neq n-1$ aus (b).

(b) Vorab notieren wir, dass mit $\mathbb{P}_{n-1} = \mathbb{P}_n$ in (2.132) unmittelbar $Z_{n-1} = 0$ folgt. Es wird zuerst (2.131) gezeigt: Wegen $\mathbb{P}^k(\pi, \sigma) = Q_1^{*k}(\sigma\pi^{-1})$, $\pi, \sigma \in \mathfrak{S}_n$ und Korollar 2.1.8 folgt

$$\mathbb{P}^k(\pi, \sigma) = \frac{1}{n!} \sum_{u=0}^{L(\sigma\pi^{-1})} u! \sum_{\nu=u}^{n} (-1)^{\nu-u} \binom{n}{\nu} \binom{\nu}{u} \left(1 - \frac{\nu}{n}\right)^k, \quad k \ge 1.$$

Die rechte Seite der Gleichung (2.131) ergibt durch Einsetzen von (2.132) unter

Beachtung von (2.3) und $Z_{n-1} = 0$

$$\left(\sum_{i=0}^{n} \lambda_i^k Z_i\right)(\pi,\sigma) = \sum_{i=0}^{n} \left(\frac{i}{n}\right)^k \sum_{\substack{j=n-L(\sigma\pi^{-1}) \\ j\geq i}}^{n} (-1)^{j-i} \binom{j}{i}\binom{n}{j} \frac{(n-j)!}{n!}.$$

Mittels der Indextransformation

$$\nu \stackrel{\text{def}}{=} n-i,\ u \stackrel{\text{def}}{=} n-j \quad \Longleftrightarrow \quad i = n-\nu,\ j = n-u$$

folgt hieraus weiter

$$\begin{aligned}
\left(\sum_{i=0}^{n} \lambda_i^k Z_i\right)(\pi,\sigma) &= \sum_{\nu=0}^{n} \left(1-\frac{\nu}{n}\right)^k \sum_{u=0}^{L(\sigma\pi^{-1})\wedge\nu} (-1)^{\nu-u} \binom{n-u}{n-\nu}\binom{n}{n-u}\frac{u!}{n!} \\
&= \frac{1}{n!} \sum_{u=0}^{L(\sigma\pi^{-1})} u! \sum_{\nu=u}^{n} (-1)^{\nu-u} \binom{n-u}{n-\nu}\binom{n}{n-u}\left(1-\frac{\nu}{n}\right)^k \\
&= \mathbb{P}^k(\pi,\sigma),
\end{aligned}$$

da eine einfache Rechnung $\binom{n-u}{n-\nu}\binom{n}{n-u} = \binom{n}{\nu}\binom{\nu}{u}$ ergibt, womit wir (2.131) für alle $k \geq 1$ gezeigt haben. Den Fall $k = 0$ rechnet man unter Benutzung von $\mathbb{P}_0 = I_n$ leicht direkt nach. Hiermit ist (2.131) bewiesen. Als Nächstes wird $\mathrm{Sp}(Z_i) \neq 0$, $i \neq n-1$ gezeigt, womit (2.133) bewiesen ist. Wir notieren hierzu

$$\begin{aligned}
\mathrm{Sp}(Z_i) &= \sum_{j=i}^{n} (-1)^{j-i} \binom{j}{i}\binom{n}{j} \mathrm{Sp}(\mathbb{P}_j) \\
&= \sum_{j=i}^{n} (-1)^{j-i} \binom{j}{i}\binom{n}{j} (n-j)! \\
&= \frac{n!}{i!} \sum_{u=0}^{n-i} \frac{(-1)^u}{u!}, \quad u \stackrel{\text{def}}{=} j-i \\
&= \binom{n}{i} D_{n-i}, \qquad (2.134)
\end{aligned}$$

wobei $D_0 \stackrel{\text{def}}{=} 1$ und D_t die Anzahl der Permutationen *ohne* Fixpunkte in $\mathfrak{S}_t$, $t \geq 1$ bezeichne. Das zweite Gleichheitszeichen gilt, da $L(\mathrm{id}) = n$ und $|\,\mathfrak{S}_n\,| = n!$ ist. Das letzte Gleichheitszeichen folgt direkt aus dem bekannten *Rencontre Problem*, das durch Anwendung der Siebformel gelöst wird (vgl. Henze [12, S.77]). Es gilt offensichtlich $D_{n-i} \neq 0$, $i \neq n-1$.

(a) Wir haben nur noch die Behauptung über die Multiplizität der Eigenwerte nachzuweisen. Diese folgt sofort aus (2.134), da einerseits eine triviale kombinatorische Überlegung

$$\binom{n}{i} D_{n-i} = |\{\pi \in \mathfrak{S}_n : \sum_{j=1}^{n} \mathbb{1}_{\{\pi(j)\}}(j) = i\}|$$

liefert und andererseits nach Bemerkung 2.4.2 die Multiplizität von $\frac{i}{n}$, $i \neq n-1$

$$\mathrm{Sp}(\widetilde{E_i}) = \mathrm{Sp}(SZ_iS^{-1}) = \mathrm{Sp}(Z_i)$$

beträgt. □

Mit Hilfe von Satz 2.4.3 lassen sich Strukturaussagen gewisser Algebren nachweisen. Der Rest dieses Abschnitts ist dieser Thematik gewidmet. Zunächst geben wir zur Wiederholung die Definition einer Algebra an:

Definition 2.4.4. Sei R ein kommutativer Ring mit 1. Eine *R-Algebra* ist dann ein linker R-Modul A zusammen mit einer bilinearen Abbildung $A \times A \to A$ (Notation $(x,y) \mapsto xy$), die assoziativ ($x(yz) = (xy)z$, $\forall\, x,y,z \in A$) ist und ein Einselement $1_A \in A$ besitzt, d.h. $1_A x = x1_A = x$, $\forall\, x \in A$. Wir werden zukünftig mit einem R-Modul immer einen linken R-Modul meinen.

Obige Struktur wird z.B. in Pierce [16] durchleuchtet. In einem einführenden Kontext wird hierauf auch in Lang [15] eingegangen.

Genauer sind die Algebren, die *hier* von Interesse sind, eine Verallgemeinerung des wohlbekannten Polynomrings $K[X]$, wobei K ein zunächst beliebiger Körper sei. $K[X]$ wird daher unseren motivierenden Ausgangspunkt bilden. Es gilt bekanntlich

$$K[X] \stackrel{\text{def}}{=} K^{(\mathbb{N}_0)} \stackrel{\text{def}}{=} \{(x_0, x_1, \ldots) \in K^{\mathbb{N}_0} \,:\, x_i = 0 \text{ für fast alle } i \in \mathbb{N}_0\},$$

wobei Addition und Multiplikation auf $K^{(\mathbb{N}_0)}$ wie gewohnt definiert sind und wir mit $X \stackrel{\text{def}}{=} (0,1,0,0,\ldots)$ die gewohnte Polynomdarstellung erhalten. $K[X]$ wird im Sinne von Definition 2.4.4 zu einer K-Algebra, falls $K[X]$ über die kanonische Inklusion $K \hookrightarrow K[X]$ als K-Modul betrachtet wird. $\mathbb{N}_0$ ist offenbar ein kommutatives Monoid. Sei M im Folgenden ein *nicht* notwendig kommutatives Monoid. Wir definieren dann

$$K[M] \stackrel{\text{def}}{=} K^{(M)} \stackrel{\text{def}}{=} \{(x_i)_{i \in M} \in K^M \,:\, x_i = 0 \text{ für fast alle } i \in M\}.$$

Die Addition wird wieder komponentenweise definiert und für die Multiplikation gelte

$$(fg)_i = \sum_{\mu,\nu\,:\,\mu\nu = i} f_\mu g_\nu, \quad f,g \in K[M],$$

wobei wegen der fehlenden Kommutativitätsvoraussetzung an M die Monoidverknüpfung nach der üblichen Konvention multiplikativ geschrieben wird. Offenbar ist $K[M]$ genau dann als Ring kommutativ, falls M kommutativ ist. Dies sehen wir sofort an der kanonischen Monoideinbettung

$$\begin{aligned} \iota :\, M &\hookrightarrow K[M] \\ \mu &\mapsto (\delta_{\nu\mu})_{\nu \in M}, \end{aligned}$$

wobei δ das Kronecker Symbol bezeichne und $K[M]$ unter der Ringmultiplikation als Monoid aufgefasst werde. Anders gesagt gilt i.Allg.

$$X^\mu X^\nu = X^{\mu\nu} \neq X^{\nu\mu} = X^\nu X^\mu$$

mit $X^\kappa \stackrel{\text{def}}{=} (\delta_{\nu\kappa})_{\nu \in M}$, $\kappa \in M$. Im Sinne dieser Verallgemeinerung kann für $K[X]$ auch $K[\mathbb{N}_0]$ geschrieben werden, was aber eher unüblich ist.

Wir werden uns im Folgenden mit $\mathbb{Q}[\mathfrak{S}_n]$ befassen, d.h. $K = \mathbb{Q}$, $M = \mathfrak{S}_n$, wobei $\mathfrak{S}_n$ als Monoid unter der *inversen* Komposition (siehe Beispiel 2.1.2) betrachtet wird. $\mathbb{Q}[\mathfrak{S}_n]$ ist für $n \geq 3$ als Ring nicht kommutativ. Bezüglich der Notation sei noch angemerkt, dass wir Elemente $(c_\pi)_{\pi\in\mathfrak{S}_n} \in \mathbb{Q}[\mathfrak{S}_n]$ zukünftig auch in der Symbolik $\sum_{\pi\in\mathfrak{S}_n} c_\pi \pi$ schreiben werden. Es wird gezeigt werden, dass gewisse Unteralgebren von $\mathbb{Q}[\mathfrak{S}_n]$, d.h. Teilmengen von $\mathbb{Q}[\mathfrak{S}_n]$, die unter denselben Verknüpfungen wie $\mathbb{Q}[\mathfrak{S}_n]$ wieder eine $\mathbb{Q}$-Algebra bilden, als Ring kommutativ sind. Es wird ebenfalls zu jeder solchen Unteralgebra eine in unserem Kontext interessante Basis angegeben werden. Mit Basis ist hier eine Vektorraumbasis gemeint, da in Definition 2.4.4 mit $R = \mathbb{Q}$ der R-Modul A eine Vektorraumstruktur trägt.

Es werden zunächst zwei zu $\mathbb{Q}[\mathfrak{S}_n]$ kanonisch isomorphe $\mathbb{Q}$-Algebren angegeben.

$$L(\mathfrak{S}_n) \stackrel{\text{def}}{=} \{f : \mathfrak{S}_n \to \mathbb{Q}\}, \tag{2.135}$$

die Menge aller Abbildungen von $\mathfrak{S}_n$ nach $\mathbb{Q}$, ist unter bekannten Verknüpfungen ein $\mathbb{Q}$-Vektorraum. Durch

$$f * g(\pi) \stackrel{\text{def}}{=} \sum_{\tau\in\mathfrak{S}_n} f(\pi\tau^{-1})g(\tau), \quad f, g \in L(\mathfrak{S}_n)\,, \pi \in \mathfrak{S}_n$$

wird $L(\mathfrak{S}_n)$ offenbar zu einer $\mathbb{Q}$-Algebra.

$$\begin{aligned} \varphi :\, L(\mathfrak{S}_n) &\to \mathbb{Q}[\mathfrak{S}_n] \\ f &\mapsto \sum_{\pi\in\mathfrak{S}_n} f(\pi)\pi \end{aligned}$$

ist offensichtlich ein $\mathbb{Q}$-Algebrenisomorphismus. Falls $f \geq 0$, $\sum_{\pi\in\mathfrak{S}_n} f(\pi) = 1$, so kann f als Verteilung auf $\mathfrak{S}_n$ angesehen werden. Auf diese Weise können die Bilder solcher f unter φ in $\mathbb{Q}[\mathfrak{S}_n]$ ebenso als Verteilungen auf $\mathfrak{S}_n$ angesehen werden. Hiermit kommt die Stochastik ins Spiel, und damit kommen in diesem Kontext insbesondere die Kartenmischsysteme ins Spiel. Im Hinblick auf Satz 2.4.3 müssen wir noch eine Verbindung zu Übergangsmatrizen herstellen. Hierzu sei

$$\begin{aligned} \psi :\, L(\mathfrak{S}_n) &\to \mathbb{Q}^{\mathfrak{S}_n \times \mathfrak{S}_n} \\ f &\mapsto M :\, M(\pi, \sigma) = f(\sigma\pi^{-1}). \end{aligned}$$

ψ wird zu einer Surjektion, falls wir die Zielmenge auf $\psi(L(\mathfrak{S}_n))$ einschränken. Sei hierzu

$$\mathbb{M} \stackrel{\text{def}}{=} \psi(L(\mathfrak{S}_n)) \tag{2.136}$$

die Menge der sogenannten *gruppenzirkulären* Matrizen. Mit ψ meinen wir zukünftig die surjektive Abbildung $\psi : L(\mathfrak{S}_n) \to \mathbb{M}$, ohne hierfür ein neues Symbol zu verwenden. $\mathbb{M}$ ist eine Unteralgebra der $\mathbb{Q}$-Algebra $\mathbb{Q}^{\mathfrak{S}_n \times \mathfrak{S}_n}$. Letztere ist eine $\mathbb{Q}$-Algebra, da diese Menge mit $\mathbb{Q} \hookrightarrow \mathbb{Q}^{\mathfrak{S}_n \times \mathfrak{S}_n}, \quad q \mapsto q\,\mathrm{id}_{\mathfrak{S}_n}$ ein $\mathbb{Q}$-Vektorraum ist und unter Hinzuziehung der Matrizenmultiplikation alle Axiome einer $\mathbb{Q}$-Algebra erfüllt. Für unsere Zwecke ist es jedoch von Vorteil, $\mathbb{Q}^{\mathfrak{S}_n \times \mathfrak{S}_n}$ als $\mathbb{Q}$-Algebra bzgl. *inverser* Matrizenmultiplikation zu definieren, genauer

$$AB \stackrel{\text{def}}{=} B \cdot A,$$

wobei auf der linken Seite die zu definierende Ringverknüpfung und auf der rechten Seite die gewöhnliche Matrizenmultiplikation steht. Grob gesprochen ist dies von

Vorteil, da bei einer Hintereinanderschaltung von Übergangsmatrizen die erste immer links steht. In unserer bisherigen Konvention steht allerdings das, was zuerst passiert, immer rechts. Daher invertieren wir die gewöhnliche Matrizenmultiplikation. Oder anders gesagt kann nur so erreicht werden, dass folgendes Lemma gilt:

Lemma 2.4.5. *ψ ist ein $\mathbb{Q}$-Algebrenisomorphismus.*

Beweis. Es wird nur die Multiplikativität gezeigt. Der Rest ist trivial. Es gilt für alle $f, g \in L(\mathfrak{S}_n)$ und $\pi, \sigma \in \mathfrak{S}_n$

$$\begin{aligned}
\psi(f*g)(\pi,\sigma) &= (f*g)(\sigma\pi^{-1}) \\
&= \sum_{\tau\in\mathfrak{S}_n} f(\sigma\pi^{-1}\tau^{-1})g(\tau), \quad \gamma \stackrel{\text{def}}{=} \tau\pi \iff \tau = \gamma\pi^{-1} \\
&= \sum_{\gamma\in\mathfrak{S}_n} f(\sigma\gamma^{-1})g(\gamma\pi^{-1}) \\
&= \sum_{\gamma\in\mathfrak{S}_n} \psi(f)(\gamma,\sigma)\psi(g)(\pi,\gamma) \\
&= (\psi(g)\cdot\psi(f))(\pi,\sigma) \\
&= (\psi(f)\psi(g))(\pi,\sigma),
\end{aligned}$$

also

$$\psi(f*g) = \psi(f)\psi(g).$$

□

Somit sind $L(\mathfrak{S}_n)$, $\mathbb{M}$ und $\mathbb{Q}[\mathfrak{S}_n]$ kanonisch isomorph und werden daher künftig miteinander identifiziert.

Satz 2.4.6. *Definiere*

$$A_i \stackrel{\text{def}}{=} \sum_{\pi: L(\pi)=n-i} \pi, \quad i = 0, 1, \ldots, n-1, \tag{2.137}$$

$$B_j \stackrel{\text{def}}{=} \sum_{i=0}^{j} A_i = \sum_{\pi: L(\pi)\geq n-j} \pi, \quad j = 0, \ldots, n-1,\ B_n \stackrel{\text{def}}{=} B_{n-1}. \tag{2.138}$$

(a) B_1 erzeugt eine n dimensionale, kommutative Unteralgebra $\mathcal{B}$ von $\mathbb{Q}[\mathfrak{S}_n]$.

(b) Es gilt

$$Z_i = \sum_{j=i}^{n} (-1)^{j-i} \binom{j}{i} \frac{B_j}{j!}, \tag{2.139}$$

mit Z_i wie in (2.132), außerdem ist

$$B_1^k = \sum_{\substack{i=0 \\ i\neq n-1}}^{n} i^k Z_i, \quad k \geq 0. \tag{2.140}$$

(c) Die B_j aus (2.138) bilden eine Basis von $\mathcal{B}$, ferner gilt

$$B_i B_j = \sum_{k=0}^{n} b_{ij}^k B_k, \quad i, j = 0, \ldots, n, \text{ wobei} \tag{2.141}$$

$$b_{ij}^k \stackrel{\text{def}}{=} \begin{cases} \frac{i!j!}{(k-i)!(k-j)!(i+j-k)!} & \max(i,j) \le k \le i+j, \\ 0 & \text{sonst.} \end{cases} \tag{2.142}$$

(d) Die A_i aus (2.137) sind ebenso eine Basis von $\mathcal{B}$.

Beweis. (a) Im Sinne obiger Identifizierung $\mathbb{M} = L(\mathfrak{S}_n) = \mathbb{Q}[\mathfrak{S}_n]$ gilt offenbar

$$\mathbb{P}_j = \frac{(n-j)!}{n!} \sum_{\pi: L(\pi) \ge n-j} \pi = \frac{(n-j)!}{n!} B_j, \tag{2.143}$$

wobei $\mathbb{P}_j$ in gewohnter Notation die Übergangsmatrix des TjTRS bezeichnet. $\mathcal{B}$ ist also ebenso die von $\mathbb{P}_1$ erzeugte $\mathbb{Q}$-Algebra. Ferner sind $(Z_i)_{i=0,\dots,n,\, i\neq n-1}$ als Vektoren des $\mathbb{Q}$-Vektorraums $\mathbb{M} = L(\mathfrak{S}_n) = \mathbb{Q}[\mathfrak{S}_n]$ linear unabhängig, denn

$$\sum_{\substack{j=0 \\ j\neq n-1}}^{n} \alpha_j Z_j = 0, \quad \alpha_j \in \mathbb{Q}$$

impliziert durch beidseitige Multiplikation von Z_i, $i \in \{0, 1, \dots, n-2, n\}$ wegen (2.126) die Gleichungen $\alpha_i Z_i = 0$, woraus nach (2.133) $\alpha_i = 0$ folgt. Nach Satz 2.4.3 (b) gilt

$$\mathbb{P}_1^k = \sum_{\substack{i=0 \\ i\neq n-1}}^{n} \lambda_i^k Z_i, \quad k \ge 0. \tag{2.144}$$

Wegen der Unabhängigkeit der Z_i und des bekannten Sachverhalts über die *Vandermonde Determinante* folgt

$$\dim_{\mathbb{Q}} \mathcal{B} \ge n. \tag{2.145}$$

Andererseits ist

$$Z_i = \sum_{j=i}^{n} (-1)^{j-i} \binom{j}{i} \binom{n}{j} \mathbb{P}_j, \tag{2.146}$$

woraus mit (2.144) und (2.143)

$$\mathbb{P}_1^k = \sum_{\substack{i=0 \\ i\neq n-1}}^{n} \lambda_i^k Z_i \in \langle B_0, \dots, B_{n-1} \rangle_{\mathbb{Q}}, \quad k \ge 0 \tag{2.147}$$

folgt. Mit rechter Symbolik ist hierbei der kleinste $\mathbb{Q}$-Untervektorraum von $\mathbb{M}$ gemeint, der $B_0, \dots, B_{n-1}$ enthält. Aus Satz 2.1.4 und (2.143) folgt sofort, dass die rechte Seite von (2.147) unter Multiplikation abgeschlossen ist und daher eine $\mathbb{Q}$-Algebra ist. Wir haben also

$$\mathcal{B} \subseteq \langle B_0, \dots, B_{n-1} \rangle_{\mathbb{Q}}. \tag{2.148}$$

Da an dieser Stelle besonders schön die Brücke zwischen dem algebraischen Formalismus und den Kartenmischsystemen zu erkennen ist, werden wir obiges Argument noch etwas genauer ausführen. Seien $i, j \in \{0, \dots, n\}$. Es ist dann

$$\underbrace{B_i}_{\in \mathbb{Q}[\mathfrak{S}_n]} = \underbrace{\frac{n!}{(n-i)!} \mathbb{P}_i}_{\in \mathbb{M}} = \underbrace{\frac{n!}{(n-i)!} Q_{\delta_i}}_{\in L(\mathfrak{S}_n)}, \quad B_j \text{ analog.}$$

Es folgt daher

$$B_iB_j = \left(\frac{n!}{(n-i)!}Q_{\delta_i}\right)\left(\frac{n!}{(n-j)!}Q_{\delta_j}\right) = \frac{n!}{(n-i)!}\frac{n!}{(n-j)!}Q_{\delta_i}Q_{\delta_j}. \tag{2.149}$$

Die Multiplikation in $L(\mathfrak{S}_n)$ ist aber gerade wie eine Faltung definiert, weshalb

$$Q_{\delta_i}Q_{\delta_j} = Q_{\delta_i} * Q_{\delta_j} = Q_{\delta_i \sharp \delta_j}$$

nach Satz 2.1.4 folgt. Andererseits gilt aber

$$Q_{\delta_i\sharp\delta_j} = \sum_{u=0}^{n}(\delta_i\sharp\delta_j)(u)Q_{\delta_u} = \sum_{u=0}^{n}(\delta_i\sharp\delta_j)(u)\frac{(n-u)!}{u!}B_u \in \langle B_0,\dots,B_{n-1}\rangle_{\mathbb{Q}},$$

was die behauptete Abgeschlossenheit zeigt. Insgesamt muss wegen (2.145) und (2.148) aus Dimensionsgründen daher $\mathcal{B} = \langle B_0,\dots,B_{n-1}\rangle_{\mathbb{Q}}$ gelten. Die Kommutativität von $\mathcal{B}$ ist wegen $\mathcal{B} = \{f(\mathbb{P}_1) : f \in \mathbb{Q}[X]\}$ klar.

(b) (2.139) ist eine unmittelbare Konsequenz aus (2.143). Für (2.140) betrachten wir die Gleichungen

$$B_1^k = \left(\frac{n!}{(n-1)!}\mathbb{P}_1\right)^k = n^k \sum_{\substack{i=0\\ i\neq n-1}}^{n} \lambda_i^k Z_i = \sum_{\substack{i=0\\ i\neq n-1}}^{n} i^k Z_i, \quad k \geq 0.$$

(c) Es bleibt noch (2.142) zu zeigen. Der Rest wurde im Beweis von (a) schon gezeigt. (2.141) ist nach (2.149) äquivalent zu

$$\frac{n!}{(n-i)!}\frac{n!}{(n-j)!}Q_{\delta_i\sharp\delta_j} = \sum_{k=0}^{n} b_{ij}^k \frac{n!}{(n-k)!}Q_{\delta_k}.$$

Andererseits gilt nach (2.7)

$$\begin{aligned}(\delta_i\sharp\delta_j)(k) &= \binom{n-j}{k-j}\binom{j}{i-(k-j)}\Big/\binom{n}{i}\\ &= \frac{(n-j)!}{(n-k)!(k-j)!}\frac{j!}{(i-k+j)!(k-i)!}\frac{i!(n-i)!}{n!},\end{aligned}$$

woraus wegen $Q_{\delta_i\sharp\delta_j} = \sum_{k=0}^{n}(\delta_i\sharp\delta_j)(k)Q_{\delta_k}$ die Gleichung (2.141) mit

$$b_{ij}^k \frac{n!}{(n-k)!}\frac{(n-i)!}{n!}\frac{(n-j)!}{n!} = \frac{(n-j)!j!i!(n-i)!}{(n-k)!(k-j)!(i-k+j)!(k-i)!n!}$$

erfüllt ist. Das ist allerdings genau (2.142).

(d) Es gilt $\langle A_0,\dots,A_{n-1}\rangle_{\mathbb{Q}} \supseteq \mathcal{B}$, da $B_j = \sum_{i=0}^{j} A_i$, $j = 0,\dots,n-1$. Andererseits ist $\dim_{\mathbb{Q}}\langle A_0,\dots,A_{n-1}\rangle_{\mathbb{Q}} \leq n$, woraus die Behauptung folgt. □

Schließlich können wir noch die strukturelle Aussage beweisen, dass $\mathcal{B}$ eine sogenannte *halbeinfache* Algebra ist.

Definition 2.4.7. Ein R-Modul A heißt *einfach*, falls $A \neq 0$ ist und dieser nur die trivialen R-Untermoduln 0 und A besitzt. Dementsprechend wird er *halbeinfach* genannt, falls er sich als direkte Summe einfacher R-Moduln schreiben lässt. Eine *R-Algebra A* wird *halbeinfach* genannt, falls sie als A-Modul (d.h. der Ring A wird

als Modul über sich selbst aufgefasst) halbeinfach ist.

Hier haben wir konkret $A = \mathcal{B}$, $R = \mathbb{Q}$.

Lemma 2.4.8. *$\mathcal{B}$ ist eine halbeinfache Algebra.*

Beweis. Satz 2.4.6 liefert zusammen mit der Unabhängigkeit (siehe Beweisanfang von (a)) der $(Z_i)_{i=0,1,\ldots,n-2,n}$

$$\mathcal{B} = \langle Z_0, \ldots, Z_{n-2}, Z_n \rangle_{\mathbb{Q}} .$$

Genauer gilt

$$\mathcal{B} = \bigotimes_{\substack{i=0 \\ i\neq n-1}}^{n} \langle Z_i \rangle_{\mathbb{Q}} ,$$

wobei $\langle Z_i \rangle_{\mathbb{Q}}$ als $\mathcal{B}$-Modul aufgefasst werde. Die Abgeschlossenheit der Multiplikation ist wegen (vgl. (2.126))

$$Z_i Z_j = \delta_{ij} Z_i, \quad i, j \in \{0, \ldots, n-2, n\}$$

gegeben. Die Summe ist direkt, da $(Z_i)_{i=0,\ldots,n-2,n}$ als Vektoren des $\mathbb{Q}$-Vektorraums $\mathcal{B}$ unabhängig sind. Ferner ist $\langle Z_i \rangle_{\mathbb{Q}}$ einfach, da $\dim_{\mathbb{Q}} \langle Z_i \rangle_{\mathbb{Q}} = 1$ und jeder $\mathcal{B}$-Untermodul vermöge der Einbettung

$$\begin{array}{ccc} \mathbb{Q} & \hookrightarrow & \mathcal{B} \\ q & \mapsto & qZ_i \end{array}$$

insbesondere auch ein $\mathbb{Q}$-Untervektorraum ist. □

Durch weitere Analogiebetrachtungen zweier TmTRS Varianten kann die Isomorphie zwischen $\mathcal{B}$ und zwei noch zu definierenden interessanten $\mathbb{Q}$-Algebren nachgewiesen werden. Diese sind damit ebenfalls semisimpel, kommutativ und besitzen als $\mathbb{Q}$-Vektorraum die Dimension n. Hierzu zuerst eine Definition.

Definition 2.4.9. Sei $F(\pi)$ der *erste Abstieg* und $G(\pi)$ der *letzte Abstieg* von π, genauer

$$F(\pi) \stackrel{\text{def}}{=} \min\{i \in \{1, \ldots, n-1\} : \pi(i) > \pi(i+1)\}, \pi \in \mathfrak{S}_n - \{\mathrm{id}\}, \quad (2.150)$$

$$G(\pi) \stackrel{\text{def}}{=} \max\{i \in \{1, \ldots, n-1\} : \pi(i) > \pi(i+1)\}, \pi \in \mathfrak{S}_n - \{\mathrm{id}\}, \quad (2.151)$$

wobei weiter $F(\mathrm{id}) \stackrel{\text{def}}{=} n$ und $G(\mathrm{id}) \stackrel{\text{def}}{=} 0$ gesetzt werde.

Korollar 2.4.10. *Sei*

$$A_j^{\mathrm{INV}} \stackrel{\text{def}}{=} \sum_{\pi: G(\pi)=j} \pi, \quad \widetilde{A_k^{\mathrm{INV}}} \stackrel{\text{def}}{=} \sum_{\pi: F(\pi)=k} \pi, \quad j = 0, 1, \ldots, n-1,\ k = 1, \ldots, n.$$

Es gelten dann die beiden Aussagen

(a) $\left\langle A_j^{\mathrm{INV}} : j = 0, \ldots, n-1 \right\rangle_{\mathbb{Q}}$ *ist eine zu $\mathcal{B}$ isomorphe $\mathbb{Q}$-Algebra.*

(b) $\left\langle \widetilde{A_k^{\mathrm{INV}}} : k = 1, \ldots, n \right\rangle_{\mathbb{Q}}$ *ist eine zu $\mathcal{B}$ isomorphe $\mathbb{Q}$-Algebra.*

Beweis. (a) $B_m^{\mathrm{INV}} \stackrel{\mathrm{def}}{=} \sum_{j=0}^{m} A_j^{\mathrm{INV}} \in \mathbb{Q}[\mathfrak{S}_n]$ entspricht bis auf einen Vorfaktor dem *inversen* TmTRS (ITmTRS). Letzterer bezeichnet ein Mischverfahren, bei dem m Karten aus einem n Kartendeck zufällig ausgewählt werden, dann aus diesem Deck herausgezogen werden und schließlich in einer zufälligen Reihenfolge wieder auf die verbleibenden $n-m$ Karten des Kartendecks oben heraufgelegt werden. Offenbar gilt mit einer zu Lemma 2.1.1 analogen Begründung

$$Q_m^{\mathrm{INV}}(\pi) \stackrel{\mathrm{def}}{=} \begin{cases} \frac{(n-m)!}{n!} & \text{falls } G(\pi) \leq m, \\ 0 & \text{falls } G(\pi) > m. \end{cases} \tag{2.152}$$

Es gilt daher

$$Q_m^{\mathrm{INV}} = \frac{(n-m)!}{n!} \sum_{j=0}^{m} A_j^{\mathrm{INV}}, \quad m = 0, 1, \ldots, n-1.$$

Falls wir $A_n^{\mathrm{INV}} \stackrel{\mathrm{def}}{=} \sum_{\pi: G(\pi)=n} \pi \stackrel{\mathrm{def}}{=} 0$ vereinbaren, ist letztere Gleichung auch für $m = n$ richtig. Aus der Anschauung folgt sofort

$$Q_m^{\mathrm{INV}}(\pi) = Q_m(\pi^{-1}), \quad \forall \pi \in \mathfrak{S}_n,\ 0 \leq m \leq n. \tag{2.153}$$

Als formales Argument wird noch die direkt aus den Definitionen von L und G folgende Gleichung

$$G(\pi) = n - L(\pi^{-1}), \quad \forall \pi \in \mathfrak{S}_n$$

angegeben. So kann (2.153) direkt an (2.3) und (2.152) abgelesen werden.

$$\begin{array}{rccccccccc} \text{Sei} & \gamma: & L(\mathfrak{S}_n) & \longrightarrow & L(\mathfrak{S}_n), & \text{wobei} & \iota: & \mathfrak{S}_n & \longrightarrow & \mathfrak{S}_n \\ & & f & \mapsto & f \circ \iota & & & \pi & \mapsto & \pi^{-1} \end{array}.$$

Es gilt dann $\gamma(Q_m^{\mathrm{INV}}) = Q_m$ und $\gamma(Q_m) = Q_m^{\mathrm{INV}}$, $m = 0, \ldots, n$, wie oben gezeigt. Sei $\widetilde{\mathcal{B}} \stackrel{\mathrm{def}}{=} \gamma(\mathcal{B})$ und $\eta: \mathcal{B} \to \widetilde{\mathcal{B}}$, $f \mapsto \gamma(f)$. Es ist dann $\widetilde{\mathcal{B}} \subseteq \mathbb{Q}[\mathfrak{S}_n]$ eine $\mathbb{Q}$-Unteralgebra und η ein $\mathbb{Q}$-Algebrenisomorphismus, denn:

η ist per Definition surjektiv und wegen $\gamma^2 = \mathrm{id}$ injektiv. Ferner ist η ein $\mathbb{Q}$-Vektorraumhomomorphismus, da γ offensichtlich ein solcher ist. Insbesondere ist $\widetilde{\mathcal{B}}$ als Bild von $\mathcal{B}$ unter γ ein $\mathbb{Q}$-Untervektorraum von $\mathbb{Q}[\mathfrak{S}_n]$. Aus der schon nachgewiesenen Kommutativität von $\mathcal{B}$ ergibt sich nun für $f, g \in \mathcal{B}$, $\pi \in \mathfrak{S}_n$

$$\eta(f * g)(\pi) = \gamma(g * f)(\pi) = (g * f)(\pi^{-1}) = \sum_{\tau \in \mathfrak{S}_n} g(\pi^{-1}\tau^{-1}) f(\tau).$$

Andererseits gilt

$$\begin{aligned} (\eta(f) * \eta(g))(\pi) &= \sum_{\kappa \in \mathfrak{S}_n} (f \circ \iota)(\pi\kappa^{-1})(g \circ \iota)(\kappa) = \sum_{\kappa \in \mathfrak{S}_n} f(\kappa\pi^{-1}) g(\kappa^{-1}) \\ &= \sum_{\tau \in \mathfrak{S}_n} g(\pi^{-1}\tau^{-1}) f(\tau), \quad \text{mit } \tau \stackrel{\mathrm{def}}{=} \kappa\pi^{-1}. \end{aligned}$$

Der Vergleich ergibt

$$\eta(f * g) = \eta(f) * \eta(g), \quad \forall f, g \in \mathcal{B}.$$

η ist daher ein $\mathbb{Q}$-Algebrenisomorphismus und $\widetilde{\mathcal{B}}$ damit eine zu $\mathcal{B}$ isomorphe $\mathbb{Q}$-

Algebra. Nach Satz 2.4.6 ist $\mathcal{B} = \langle Q_0, \ldots, Q_{n-1}\rangle_{\mathbb{Q}}$. Das ergibt

$$\begin{aligned}\widetilde{\mathcal{B}} &= \gamma(\langle Q_0, \ldots, Q_{n-1}\rangle_{\mathbb{Q}}) = \langle \gamma(Q_0), \ldots, \gamma(Q_{n-1})\rangle_{\mathbb{Q}} \\ &= \left\langle Q_0^{\mathrm{INV}}, \ldots, Q_{n-1}^{\mathrm{INV}}\right\rangle_{\mathbb{Q}} = \left\langle B_0^{\mathrm{INV}}, \ldots, B_{n-1}^{\mathrm{INV}}\right\rangle_{\mathbb{Q}} \\ &= \left\langle A_0^{\mathrm{INV}}, \ldots, A_{n-1}^{\mathrm{INV}}\right\rangle_{\mathbb{Q}}.\end{aligned}$$

(b) $\widetilde{B_m^{\mathrm{INV}}} \stackrel{\mathrm{def}}{=} \sum_{j=n-m}^{n} \widetilde{A_j^{\mathrm{INV}}} \in \mathbb{Q}[\mathfrak{S}_n]$ entspricht bis auf einen Vorfaktor dem *inversen Bottom-m-to-Random-Shuffle*(IBmTRS). Es handelt sich hierbei um einen ITmTRS, nur dass im letzten Schritt die Karten nicht *auf* die verbleibenden $n-m$ Karten, sondern *unter* diese gelegt werden. Das Vorgehen ist bis auf einige Modifikationen im Wesentlichen wie im Beweisteil (a). Wir werden zur Demonstration dennoch ähnlich ausführlich argumentieren. In der gewohnten Notation gilt

$$\widetilde{Q_m^{\mathrm{INV}}}(\pi) \stackrel{\mathrm{def}}{=} \begin{cases} 0 & \text{falls } F(\pi) < n-m, \\ \frac{(n-m)!}{n!} & \text{falls } F(\pi) \geq n-m. \end{cases} \tag{2.154}$$

Wir haben daher

$$\widetilde{Q_m^{\mathrm{INV}}} = \frac{(n-m)!}{n!} \sum_{j=n-m}^{n} \widetilde{A_j^{\mathrm{INV}}}, \quad m = 0, \ldots, n-1.$$

Falls $\widetilde{A_0^{\mathrm{INV}}} \stackrel{\mathrm{def}}{=} \sum_{\pi\,:\,F(\pi)=0} \pi \stackrel{\mathrm{def}}{=} 0$ vereinbart wird, ist letztere Gleichung auch für $m = n$ richtig. Mit den Definitionen

$$\widetilde{\gamma}: \begin{array}{ccc} L(\mathfrak{S}_n) & \longrightarrow & L(\mathfrak{S}_n), \\ f & \mapsto & f \circ \widetilde{\iota} \end{array} \qquad \widetilde{\iota}: \begin{array}{ccc} \mathfrak{S}_n & \longrightarrow & \mathfrak{S}_n, \\ \pi & \mapsto & \sigma\pi\sigma \end{array}, \quad \sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix}$$

gilt

$$\widetilde{\gamma}(\widetilde{Q_m^{\mathrm{INV}}}) = Q_m^{\mathrm{INV}},\ m = 0, 1, \ldots, n. \tag{2.155}$$

Anschaulich gesprochen invertiert $\widetilde{\gamma}$ die Kartennummerierung und dreht das Kartendeck gerade so um, dass aus „oben" „unten" wird. Das formale Argument an dieser Stelle ist $G(\pi) = n - F(\sigma\pi\sigma)$, $\forall \pi \in \mathfrak{S}_n$. So kann (2.155) direkt an (2.152) und (2.154) abgelesen werden.

Sei $\widehat{\mathcal{B}} \stackrel{\mathrm{def}}{=} \widetilde{\gamma}(\widetilde{\mathcal{B}})$ und $\widetilde{\eta}: \widetilde{\mathcal{B}} \to \widehat{\mathcal{B}}$, $f \mapsto \widetilde{\gamma}(f)$. Es stellt sich mit derselben Vorgehensweise wie in (a) heraus, dass $\widetilde{\eta}$ ein $\mathbb{Q}$-Algebrenisomorphismus ist. Wir weisen wieder die Multiplizität nach: Für $f, g \in \widetilde{\mathcal{B}}$, $\pi \in \mathfrak{S}_n$ gilt

$$\begin{aligned}\widetilde{\eta}(f * g)(\pi) &= (f * g)(\sigma\pi\sigma) \\ &= \sum_{\kappa \in \mathfrak{S}_n} f(\sigma\pi\sigma\kappa^{-1}) g(\kappa), \quad \tau \stackrel{\mathrm{def}}{=} \sigma\kappa\sigma \\ &= \sum_{\tau \in \mathfrak{S}_n} f(\sigma\pi \overbrace{\sigma\sigma}^{=\mathrm{id}} \tau^{-1}\sigma) g(\sigma\tau\sigma) \\ &= \sum_{\tau \in \mathfrak{S}_n} \widetilde{\eta}(f)(\pi\tau^{-1}) \widetilde{\eta}(g)(\tau) \\ &= (\widetilde{\eta}(f) * \widetilde{\eta}(g))(\pi).\end{aligned}$$

$\widehat{\mathcal{B}}$ ist folglich eine zu $\widetilde{\mathcal{B}}$ und damit nach Teil (a) zu $\mathcal{B}$ isomorphe $\mathbb{Q}$-Algebra. Wir

haben

$$\begin{aligned}\widehat{\mathcal{B}} &= \widetilde{\gamma}\left(\left\langle Q_0^{\mathrm{INV}},\dots,Q_{n-1}^{\mathrm{INV}}\right\rangle_{\mathbb{Q}}\right) = \left\langle \widetilde{\gamma}(Q_0^{\mathrm{INV}}),\dots,\widetilde{\gamma}(Q_{n-1}^{\mathrm{INV}})\right\rangle_{\mathbb{Q}} \\ &= \left\langle \widetilde{Q_0^{\mathrm{INV}}},\dots,\widetilde{Q_{n-1}^{\mathrm{INV}}}\right\rangle_{\mathbb{Q}} \qquad = \left\langle \widetilde{B_0^{\mathrm{INV}}},\dots,\widetilde{B_{n-1}^{\mathrm{INV}}}\right\rangle_{\mathbb{Q}} \\ &= \left\langle \widetilde{A_1^{\mathrm{INV}}},\dots,\widetilde{A_n^{\mathrm{INV}}}\right\rangle_{\mathbb{Q}}.\end{aligned}$$

Damit ist auch Aussage (b) bewiesen. □

Bemerkung 2.4.11. Bisher wurde die symmetrische Gruppe $\mathfrak{S}_n$ mit der *inversen* Verknüpfung

$$\sigma\pi = \pi\circ\sigma, \quad \pi,\sigma\in\mathfrak{S}_n$$

betrachtet. $\widehat{\mathfrak{S}_n}$ bezeichne $\mathfrak{S}_n$, versehen mit der gewöhnlichen Komposition „$\circ$" von Abbildungen. Wir werden als Nächstes zeigen, dass die in diesem Abschnitt gemachten Aussagen sich auch auf $\mathbb{Q}[\widehat{\mathfrak{S}_n}]$ übertragen: Dieses folgt aus der Kommutativität entsprechender Unteralgebren, da das Produkt in beiden Ringen dann gleich definiert ist: „$\cdot$" bezeichne die Multiplikation in $\mathbb{Q}[\mathfrak{S}_n]$ und „$\bullet$" die Multiplikation in $\mathbb{Q}[\widehat{\mathfrak{S}_n}]$. Sei $A\subseteq\mathbb{Q}[\mathfrak{S}_n]$ eine kommutative $\mathbb{Q}$-Unteralgebra. Dann gelten für alle

$$f,g\in A, \quad f=\sum_{\pi\in\mathfrak{S}_n} a_\pi\pi,\ g=\sum_{\pi\in\mathfrak{S}_n} b_\pi\pi$$

die Gleichungen

$$\begin{aligned}f\cdot g &= g\cdot f \\ &= \sum_{\pi\in\mathfrak{S}_n}\left(\sum_{\sigma\tau=\pi} b_\sigma a_\tau\right)\pi \\ &= \sum_{\pi\in\mathfrak{S}_n}\left(\sum_{\tau\circ\sigma=\pi} a_\tau b_\sigma\right)\pi \\ &= f\bullet g.\end{aligned}$$

2.5 Riffle- und Top-to-Random-Shuffles

Gegeben sei ein Kartendeck mit n Karten. Von diesen werden $0\le m\le n$ Karten abgehoben und an zufälligen Positionen wieder in das verbleibende Deck zurückgesteckt. Allerdings soll hierbei die relative Ordnung der Karten erhalten bleiben. Es handelt sich also *nicht* um einen TmTRS. Die inverse Variante dieses Mischverfahrens besteht darin, aus dem n Kartendeck per Zufall m Karten auszuwählen (d.h. jede m-elementige Teilmenge wird mit der gleichen Wahrscheinlichkeit $1/\binom{n}{m}$ gewählt) und diese dann in derselben Reihenfolge wieder auf das Deck von oben zurückzulegen. Dieses Mischverfahren wird ein $(m,n-m)$ Shuffle genannt.

Ein Spezialfall des folgenden Lemmas besagt, dass das Konvergenzverhalten der *nicht* inversen Variante äquivalent zur inversen Variante ist, da sich diese offenbar durch den Übergang von Q zu $\widetilde{Q}$ (Notation Lemma 2.5.1) manifestiert, wobei mit Q die Verteilung der nicht inversen Variante gemeint ist. Es reicht daher, die inverse Variante zu studieren, falls es uns ausschließlich auf den Variations- bzw. Separationsabstand ankommt, was hier der Fall ist.

Lemma 2.5.1. *Seien Q, P Verteilungen auf $\mathfrak{S}_n$ und $\iota : \pi \mapsto \pi^{-1}$ die Inversion auf $\mathfrak{S}_n$. Wir schreiben $\widetilde{W} \stackrel{\text{def}}{=} W^{\iota}$, d.h. $\widetilde{W}(\pi) = W(\pi^{-1})$, $\pi \in \mathfrak{S}_n$ für Verteilungen W auf $\mathfrak{S}_n$. Es gilt dann*

$$\widetilde{Q} * \widetilde{P}(\pi) = \widetilde{P * Q}(\pi), \quad \pi \in \mathfrak{S}_n,$$

woraus insbesondere

$$\begin{aligned} \|\big(\widetilde{Q}\big)^{*k} - U\| &= \|Q^{*k} - U\|, \\ \operatorname{sep}\big(\big(\widetilde{Q}\big)^{*k}, U\big) &= \operatorname{sep}\big(Q^{*k}, U\big), \quad k \in \mathbb{N}. \end{aligned}$$

folgt.

Beweis. Es gilt

$$\begin{aligned} \widetilde{Q} * \widetilde{P}(\pi) &= \sum_{\tau \in \mathfrak{S}_n} \widetilde{Q}(\pi\tau^{-1})\widetilde{P}(\tau) = \sum_{\tau \in \mathfrak{S}_n} Q(\tau\pi^{-1})P(\tau^{-1}), \quad \gamma \stackrel{\text{def}}{=} \tau\pi^{-1} \\ &= \sum_{\gamma \in \mathfrak{S}_n} P(\pi^{-1}\gamma^{-1})Q(\gamma) = P * Q(\pi^{-1}) \\ &= \widetilde{P * Q}(\pi), \quad \pi \in \mathfrak{S}_n. \end{aligned}$$

Daraus folgt induktiv $\big(\widetilde{Q}\big)^{*k} = \widetilde{Q^{*k}}$, $\forall k \in \mathbb{N}$. Wegen $U(\pi) = U(\pi^{-1}) = \frac{1}{n!}$, $\forall \pi \in \mathfrak{S}_n$ erhalten wir schließlich

$$\begin{aligned} \|Q^{*k} - U\| &= \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} |Q^{*k}(\pi) - U(\pi)| = \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} |Q^{*k}(\pi^{-1}) - U(\pi^{-1})| \\ &= \frac{1}{2} \sum_{\pi \in \mathfrak{S}_n} |\big(\widetilde{Q}\big)^{*k}(\pi) - U(\pi)| = \|\big(\widetilde{Q}\big)^{*k} - U\|, \quad k \in \mathbb{N} \end{aligned}$$

und ebenso

$$\begin{aligned} \operatorname{sep}\big(Q^{*k}, U\big) &= \max_{\pi \in \mathfrak{S}_n} \left(1 - \frac{Q^{*k}(\pi)}{U(\pi)}\right) = \max_{\pi \in \mathfrak{S}_n} \left(1 - \frac{Q^{*k}(\pi^{-1})}{U(\pi^{-1})}\right) \\ &= \max_{\pi \in \mathfrak{S}_n} \left(1 - \frac{\big(\widetilde{Q}\big)^{*k}(\pi)}{U(\pi)}\right) = \operatorname{sep}\big(\big(\widetilde{Q}\big)^{*k}, U\big), \quad k \in \mathbb{N}. \end{aligned}$$

□

Definition 2.5.2. Zur Verallgemeinerung des $(m, n-m)$ Shuffles vereinbaren wir:

(1) $\nu = (\nu_1, \ldots, \nu_r) \in \mathbb{N}^r$ heißt *Komposition von n*, falls $\sum_{i=1}^r \nu_i = n$, $1 \le \nu_i \le n$ gilt. $\mathfrak{Z}$ bezeichne die Menge aller Kompositionen von n.

(2) $D(\pi) \stackrel{\text{def}}{=} \{i \in \{1, \ldots, n-1\} : \pi(i+1) < \pi(i)\}$ bezeichne die *Menge der Abstiege* von π, $\pi \in \mathfrak{S}_n$.

(3) Mit einem *ν-Shuffle*, wobei $\nu = (\nu_1, \ldots, \nu_r)$ eine Komposition von n sei, ist folgender Mischvorgang gemeint: Gegeben sei ein n Kartendeck. Es werden nun sukzessive r Teilmengen des Kartendecks der Größe $\nu_1, \ldots, \nu_r$ unterschiedlich markiert (z.B. mit $1, 2, \ldots, r$). Die Wahl dieser Teilmengen erfolgt *rein zufällig*, d.h. jede Mengenkombination besitzt die gleiche Wahrscheinlichkeit $1/\binom{n}{\nu_1, \ldots, \nu_r}$. Die Karten mit der Markierung 1 werden nach oben gelegt, die

Karten mit der Markierung 2 darunter u.s.w. Hierbei soll die *relative Ordnung* der Karten in jeder der r Teilmengen *beibehalten* werden.

$\mathcal{P}(\{1,2,\ldots,n-1\})$ bezeichne die Potenzmenge von $\{1,2,\ldots,n-1\}$. Dann ist

$$\begin{array}{rccl} \varphi: & \mathcal{P}(\{1,2,\ldots,n-1\}) & \longrightarrow & \mathfrak{Z} \\ & \{d_1 < d_2 < \ldots < d_k\} & \mapsto & (d_1 - d_0, d_2 - d_1, \ldots, d_{k+1} - d_k) \end{array}$$

mit $d_0 \stackrel{\text{def}}{=} 0$, $d_{k+1} \stackrel{\text{def}}{=} n$ und $\varphi(\emptyset) \stackrel{\text{def}}{=} (n)$ offenbar eine Bijektion. Wir können nun die Verteilung eines ν-Shuffles in prägnanter Form angeben.

Lemma 2.5.3. *Sei $\nu = (\nu_1, \ldots, \nu_r)$ eine Komposition von n. Die dem ν-Shuffle zugehörige Verteilung lautet dann*

$$Q_\nu(\pi) = \begin{cases} \frac{1}{\binom{n}{\nu_1,\ldots,\nu_r}} & \textit{falls } D(\pi) \subseteq \varphi^{-1}(\nu), \\ 0 & \textit{sonst.} \end{cases} \tag{2.156}$$

Beweis. Offenbar teilt sich $\mathfrak{S}_n$ wieder in zwei disjunkte Mengen auf: Gegeben sei ein vorsortiertes Kartendeck, d.h. $\mathrm{id} \in \mathfrak{S}_n$, kann die eine Teilmenge von $\mathfrak{S}_n$ durch einen ν-Shuffle Schritt erreicht werden, die andere jedoch nicht. Es folgt unmittelbar aus der Definition des ν-Shuffles, dass jede erreichbare Permutation die gleiche Wahrscheinlichkeit besitzt. Hiervon gibt es gerade $\binom{n}{\nu_1,\ldots,\nu_r}$ Permutationen. So ist der Multibinomialkoeffizient gerade definiert. Jede erreichbare Permutation ist offensichtlich dadurch gekennzeichnet, dass eventuelle Abstiege nur an den Positionen $\varphi^{-1}(\nu)$ liegen können. Man mache sich hierzu noch einmal klar, dass die relative Ordnung der r Kartenteilmengen bestehen bleibt, wir also, anschaulich formuliert, r Ketten der Länge $\nu_1, \nu_2, \ldots, \nu_r$ haben, die zusammengelegt nur an ihren Verbindungsstellen einen Abstieg haben können (aber nicht müssen). So erklärt sich die Charakterisierung der erreichbaren Permutationen durch die Bedingung $D(\pi) \subseteq \varphi^{-1}(\nu)$. Die anderen, nicht erreichbaren Permutationen besitzen die Wahrscheinlichkeit Null. □

Bemerkung 2.5.4. Sei $1^m \stackrel{\text{def}}{=} (1,1,\ldots,1) \in \mathbb{N}^m$, $0 \leq m \leq n$. Dann beschreibt $(1^m, n-m)$ offenbar den ITmTRS. Es gilt $\varphi^{-1}((1^m, n-m)) = \{1,2,\ldots,m\}$, d.h. Abstiege sind nur innerhalb der ersten m Positionen möglich. Ein plausibles Ergebnis, da bei einem ITmTRS genau an diesen Positionen Willkür herrscht. Ferner gilt $\binom{n}{1^m,n-m} = \frac{n!}{(n-m)!}$, was ebenfalls zu erwarten war, siehe (2.152).

Sei $m \in \mathbb{N}$ und $n \stackrel{\text{def}}{=} 2m$. $\nu = (m,m)$ beschreibt dann einen *Riffle-Shuffle*, jedoch wird das Kartendeck vor dem Zusammenfächern *genau* halbiert. Es handelt sich folglich um eine Vereinfachung des ursprünglichen *Riffle-Shuffles*, bei dem die Anzahl der abgehobenen Karten variieren kann. Wie genau diese Anzahl variieren soll und welche Verteilung daraus resultiert, wird später diskutiert.

Im Folgenden wird die Verteilung untersucht, welche sich bei sukzessiven, unabhängig voneinander ausgeführten ν-Shuffles ergibt. ν kann hierbei zwischen den einzelnen Shuffles variieren. Die Art der Variation liegt vor Beginn der Mischvorgänge aber schon deterministisch fest. $Q_{\nu^k} * \ldots * Q_{\nu^1}$ mit $\nu^1, \ldots, \nu^k \in \mathfrak{Z}$ soll folglich studiert werden. Wir veranschaulichen diese k Mischvorgänge durch ein geschickt gewähltes Markierungsschema. Sei hierzu $A \in \mathbb{N}^{n \times k}$ eine Matrix, deren Spalten unabhängige Zufallsvektoren sind. Genauer soll bei einer konkreten Realisierung von A die i-te Spalte an genau ν^i_j zufälligen Stellen mit j belegt werden, $1 \leq j \leq r_i$, falls $\nu^i \in \mathbb{N}^{r_i}$. Bei unserem vorsortierten n Kartendeck ($\mathrm{id} \in \mathfrak{S}_n$) schreiben wir nun

die l-te Zeile der soeben konstruierten Matrix auf die l-te Karte. Die k Shuffleiterationen werden in k Schritten durchgeführt. Im ersten Schritt werden alle Karten, auf deren Markierung (die Zeilen der Matrix) sich an erster Stelle eine Eins befindet, nach oben gelegt. Direkt im Anschluss werden die Karten, die an erster Stelle ihrer Markierung eine Zwei aufweisen, unter erstere Karten gelegt u.s.w. Die relative Ordnung der einzelnen Kartenteilmengen, die bewegt werden, bleibt hierbei erhalten. Im zweiten Schritt wiederholen wir dasselbe Prozedere, betrachten jedoch immer die zweite Stelle der Kartenmarkierungen. Der erste Schritt entspricht folglich einem ν^1-Shuffle und der zweite einem vom ersten Schritt unabhängigen ν^2-Shuffle. Das wird sukzessive bis zum k-ten Schritt fortgesetzt. Folgende Skizze veranschaulicht obige Prosa. Es werden hier mit $n = 6$ drei sukzessive $(2, 4)$-Shuffles durchgeführt. Hierbei wurde natürlich nur *eine* mögliche Realisierung von A ausgewählt.

$$\begin{array}{cccc} & & A & \\ 1 & 1 & 2 & 2 \\ 2 & 2 & 2 & 1 \\ 3 & 2 & 1 & 2 \\ 4 & 1 & 1 & 1 \\ 5 & 2 & 2 & 2 \\ 6 & 2 & 2 & 2 \end{array} \quad \overset{\text{shuffle } 1}{\longrightarrow} \quad \begin{array}{ccc} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{array} \quad \overset{\text{shuffle } 2}{\longrightarrow} \quad \begin{array}{ccc} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 1 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{array} \quad \overset{\text{shuffle } 3}{\longrightarrow} \quad \begin{array}{ccc} 1 & 1 & 1 \\ 2 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{array} \tag{2.157}$$

Beachte, dass die letzte Anordnung in (2.157) lexikographisch von rechts nach links geordnet ist. Sei $\nu(A)$ die Komposition von n, die zu den sich wiederholenden Reihen der lexikographisch von rechts nach links sortierten Matrix A korrespondiert. Im obigen Beispiel also $\nu(A) = (1, 1, 1, 1, 2)$. Ähnlich wie Satz 2.1.4 gilt dann

Satz 2.5.5. *Seien $\nu^1, \nu^2, \ldots, \nu^k \in \mathfrak{Z}$. Dann gilt*

$$Q_{\nu^k} * \ldots * Q_{\nu^1} = \sum_{\nu \in \mathfrak{Z}} c(\nu; \nu^1, \ldots, \nu^k) Q_\nu,$$

wobei $c(\nu; \nu^1, \ldots, \nu^k)$ die Wahrscheinlichkeit des Ereignisses $\{\nu(A) = \nu\}$ ist.

Beweis. Die Summe resultiert aus der disjunkten Zerlegung des zugrunde liegenden Wahrscheinlichkeitsraumes in die Ereignisse $\{\nu(A) = \nu\}$, $\nu \in \mathfrak{Z}$. Es besitzt

$$(A_{ij})_{i=1,\ldots,n,\, j=1,\ldots,k}$$

offenbar die gleiche Verteilung wie

$$(A_{\pi(i),j})_{i=1,\ldots,n,\, j=1,\ldots,k}, \quad \pi \in \mathfrak{S}_\mathrm{n}\,.$$

Das folgt direkt aus der Definition von A. Wir schreiben $\nu = (\nu_1, \ldots, \nu_l)$, $1 \leq l \leq n$. Bedingt auf $\{\nu(A) = \nu\}$ ist folglich nach dem Mischvorgang jede Teilmenge von ν_1 Karten gleichwahrscheinlich oben auf dem Kartendeck. Ebenso ist jede Teilmenge von ν_2 Karten unter den verbleibenden $n - \nu_1$ Karten gleichwahrscheinlich direkt unter den zuerst genannten ν_1 Karten u.s.w. Das ist aber genau die Definition eines ν-Shuffles. □

Als Nächstes werden wir obiges Markierungsschema, d.h. die Matrix A (siehe (2.157)), nutzen, um den Variationsabstand zwischen k nacheinander ausgeführten ν-Shuffles und der Gleichverteilung U abzuschätzen. Hierbei wird das Prinzip der

stark stationären Zeiten zum Einsatz kommen. Wir wollen noch darauf aufmerksam machen, dass im folgenden Satz $\nu \in \mathbb{N}^r$ zwischen den einzelnen Schritten *nicht* variieren darf. Nur so kann die Homogenität der resultierenden Markov-Kette garantiert werden, und wir können unsere Theorie in Kapitel 1 anwenden. Allerdings ist es nicht aufwendig, Satz 2.5.6 auf Shuffels, zwischen denen ν variieren darf, zu verallgemeinern. Das ist für unsere Zwecke allerdings unnötig.

Satz 2.5.6. *Mit $\nu = (\nu_1, \ldots, \nu_r) \in \mathfrak{Z}$ gilt*

$$\|Q_\nu^{*k} - U\| \leq \operatorname{sep}(Q_\nu^{*k}, U) \leq \binom{n}{2} \left(\sum_{j=1}^{r} \frac{\nu_j(\nu_j - 1)}{n(n-1)} \right)^k, \quad k \in \mathbb{N}. \tag{2.158}$$

Beweis. Der Fall $\nu = (n)$ ergibt auf der rechten Seite in (2.158) die triviale Schranke $\binom{n}{2}$, weshalb o.E. $\nu \neq (n)$ angenommen wird. A^k bezeichne die $\mathbb{N}^{n\times k}$ Matrix, welche zu k sukzessiven, unabhängig voneinander ausgeführten ν-Shuffles gehört. Genauer sind hierbei A^k, $k \geq 1$ Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$. A^k wird dabei so generiert, dass A^{k+1} eine Fortsetzung von A^k ist, d.h.

$$\left(A_{ij}^{k+1}\right)_{1\leq i\leq n, 1\leq j\leq k} = \left(A_{ij}^{k}\right)_{1\leq i\leq n, 1\leq j\leq k}, \quad k \geq 1.$$

Wir definieren wie weiter oben beschrieben $(X_k)_{k\geq 0}$ auf $(\Omega, \mathfrak{A}, P)$ als Funktion von $(A^k)_{k\geq 1}$, so dass $X_0 \stackrel{\text{def}}{=} \mathrm{id}$ und X_k das Ergebnis der mittels A^k ausgeführten Mischschritte ist. Nach Konstruktion ist daher $(X_k)_{k\in\mathbb{N}_0}$ ein *Random Walk* auf $\mathfrak{S}_\mathrm{n}$, wobei die Übergangsmatrix im Sinne von Abschnitt 1.3 von Q_ν generiert wird. Seien

$$\begin{aligned} \mathcal{F}_0 &\stackrel{\text{def}}{=} \{\emptyset, \Omega\}, \\ \mathcal{F}_k &\stackrel{\text{def}}{=} \sigma(A^m,\, 1 \leq m \leq k), \quad k \geq 1. \end{aligned}$$

Offenbar ist $(X_k)_{k\in\mathbb{N}_0}$ dann bezüglich der Filtration $(\mathcal{F}_k)_{k\in\mathbb{N}_0}$ eine DMK. Wir definieren eine $(\mathcal{F}_k)_{k\in\mathbb{N}_0}$ Stoppzeit T durch

$$T \stackrel{\text{def}}{=} \inf\{k \geq 1 \,:\, A_i^k \neq A_j^k, \quad \forall i \neq j\},$$

wobei mit A_i^k, $i = 1, \ldots, n$ die i-te Zeile von A^k gemeint ist. T ist P-f.s. endlich, da wegen $\nu \neq (n)$ zwei beliebig fixierte Zeilen nach endlich vielen Schritten P-f.s. verschieden sind und die Anzahl n der Zeilen fixiert ist. Entscheidend ist das offensichtliche Faktum

$$P(X_k = \pi | T = k) = \frac{1}{n!} = U(\pi), \quad \pi \in \mathfrak{S}_\mathrm{n}\,.$$

T ist daher eine *stark stationäre Zeit*. Nach Satz 1.2.13 und wegen $P^{X_k} = Q_\nu^{*k}$ gilt folglich

$$\|Q_\nu^{*k} - U\| \leq \operatorname{sep}(Q_\nu^{*k}, U) \leq P(T > k).$$

Wir müssen folglich noch $P(T > k)$ abschätzen. Seien hierzu $X_{ij}^{(k)}$, $1 \leq i < j \leq n$ Indikatorzufallsvariablen mit

$$X_{ij}^{(k)} = \begin{cases} 1 & \text{falls die } i\text{-te und } j\text{-te Zeile von } A^k \text{ übereinstimmen,} \\ 0 & \text{sonst.} \end{cases}$$

Es gilt $\{T > k\} = \bigcup_{i<j}\{X_{ij}^{(k)} = 1\}$, $k \in \mathbb{N}$. Daraus folgt

$$P(T > k) \leq \sum_{1\leq i<j\leq n} P(X_{ij}^{(k)} = 1) = \binom{n}{2} P(X_{12}^{(k)} = 1), \tag{2.159}$$

wobei letztere Gleichheit gültig ist, da offenbar

$$\begin{aligned} P(X_{ij}^{(k)} = 1) &= P(X_{12}^{(k)} = 1), \quad 1 \leq i < j \leq n, \\ |\{\{i,j\} \in \{1,2,\ldots,n\} \,:\, i<j\}| &= |\{M \subseteq \{1,2,\ldots,n\} \,:\, |M| = 2\}| = \binom{n}{2} \end{aligned}$$

gilt. Es bleibt daher nur noch

$$P(X_{12}^{(k)} = 1) = \left(\sum_{j=1}^{r} \frac{\nu_j(\nu_j - 1)}{n(n-1)}\right)^k \tag{2.160}$$

zu zeigen. Die ersten beiden Zeilen von A^k stimmen genau dann überein, falls alle k Einträge übereinstimmen. Da die k Spalten von A^k per Voraussetzung unabhängig und identisch verteilt sind, folgt hieraus die k-te Potenz in (2.160). Für jede Spalte gibt es r Möglichkeiten der Übereinstimmung, nämlich die Markierungen $1, 2, \ldots, r$. Diese Ereignisse sind paarweise disjunkt, woraus die Summe in (2.160) resultiert. Die Wahrscheinlichkeit, dass Symbol j, $1 \leq j \leq r$ in Zeile eins, Spalte i steht, ist wegen der vorausgesetzten Gleichverteilung der Bestückung der Matrix ν_j/n. Bedingt auf dieses Ereignis sind die Zeilen eins und zwei an der Position i genau dann gleich, falls sich das Symbol j ebenso an der Stelle i der zweiten Zeile befindet. Die Wahrscheinlichkeit hierfür beträgt $(\nu_j - 1)/(n-1)$. Durch Multiplikation der bedingten Wahrscheinlichkeit mit dem, worauf bedingt wurde, ergibt sich der Schnitt der Ereignisse, was hier bedeutet, dass in beiden Zeilen an Position i das Markierungssymbol j steht. Hiermit ist (2.160) und damit der Satz bewiesen. □

Bemerkung 2.5.7. $\nu(A) \stackrel{\text{def}}{=} (\nu(A^k))_{k\geq 0}$ mit $\nu(A^0) \stackrel{\text{def}}{=} (n)$ ist eine Markov-Kette auf dem Zustandsraum $\mathfrak{Z}$, und es gilt

$$c(\nu; \nu^1, \ldots, \nu^k) = \sum_{\nu' \in \mathfrak{Z}} c(\nu'; \nu^1, \ldots, \nu^{k-1}) c(\nu; \nu', \nu^k), \quad k \geq 1.$$

Im Fall $\nu^1 = \ldots = \nu^k$ ist $\nu(A)$ sogar eine homogene Markov-Kette. Unter dieser Annahme ist sie im Sinne von Diaconis und Fill [7] der duale Markovprozess von dem Random Walk $(X_k)_{k\geq 0}$ auf $\mathfrak{S}_n$, der von Q_{ν^1} generiert wird und in $\mathrm{id} \in \mathfrak{S}_n$ startet. Hierzu identifiziere man $\nu \in \mathfrak{Z}$ mit der Teilmenge

$$\nu^* \stackrel{\text{def}}{=} \{\pi \in \mathfrak{S}_n \,:\, D(\pi) \subseteq \varphi^{-1}(\nu)\}.$$

In der Terminologie von [7] ist dann $(\nu(A^k)^*)_{k\geq 0}$ ein mengenwertiger, stark stationärer, dualer Markovprozess (engl. set-valued strong stationary dual) von $(X_k)_{k\geq 0}$, der in $(n)^* = \{\mathrm{id}\}$ startet und auf dem Zustandsraum

$$\mathcal{S}^* = \{\nu^* \,:\, \nu \in \mathfrak{Z}\}$$

operiert. Die Übergangsmatrix ist hierbei

$$P^*((\nu')^*, \nu^*) = c(\nu; \nu', \nu^1).$$

Die duale Kette besagt anschaulich gesprochen, wie weit wir von der stationären Verteilung (hier die Gleichverteilung U auf $\mathfrak{S}_\mathrm{n}$) entfernt sind. Es werden peu à peu immer größer werdende Teilmengen $\nu^* \subseteq \mathfrak{S}_\mathrm{n}$ durchlaufen, auf derer jede Permutation gleichwahrscheinlich und außerhalb denen jede Permutation mit Wahrscheinlichkeit Null angenommen wird. Zum Schluss gelangen wir in $\mathfrak{S}_\mathrm{n} \in \mathcal{S}^*$, den absorbierenden Zustand der Markov-Kette. Hier angelangt, ist unser Kartendeck perfekt durchmischt oder in der Terminologie von [7]

$$\Lambda(\mathfrak{S}_\mathrm{n}, \pi) \stackrel{\mathrm{def}}{=} P(X_k = \pi | \nu(A^k) = (1, \ldots, 1)) = U(\pi) = \frac{1}{n!}, \quad \pi \in \mathfrak{S}_\mathrm{n},\, k \geq 1.$$

Für ein ähnliches, komplett ausgearbeitetes Beispiel sei auf Alsmeyer [5, S.164ff] hingewiesen.

Wir kommen nun zu einigen interessanten Anwendungen von Satz 2.5.6. Als Erstes wird eine naheliegende Abänderung des TmTRS untersucht, bei der die Einordnung der $m \geq 1$ abgehobenen Karten unter Beibehaltung derer relativen Ordnung erfolgt. Intuitiv wird man zunächst davon ausgehen, dass dieses Mischverfahren langsamer als der TmTRS ist, da im letzteren die eingefügten m Karten noch beliebig untereinander permutieren können. In der Grenzwertbetrachtung $n \to \infty$ stellt sich diese Intuition aber als falsch heraus, wie die nachfolgende Abschätzung zeigen wird.

Wir wollen folglich $(m, n-m)$ Shuffles untersuchen. Nach Satz 2.5.6 ergibt sich mit $\nu^1 = \ldots = \nu^k = (m, n-m)$

$$\|Q_{\nu^k} * \ldots * Q_{\nu^1} - U\| \leq \binom{n}{2} \left(1 - \frac{2m(n-m)}{n(n-1)}\right)^k. \tag{2.161}$$

Obige Prosa folgt daher aus dem nächsten Lemma.

Lemma 2.5.8. *Für jedes fixierte $m \in \mathbb{N}$ und $c \in \mathbb{R}$ gilt mit $k_n \stackrel{\mathrm{def}}{=} \frac{n}{m}(\log n + c)$*

$$\binom{n}{2} \left(1 - \frac{2m(n-m)}{n(n-1)}\right)^{k_n} \longrightarrow \frac{1}{2} e^{-2c}, \quad n \to \infty.$$

Beweis. Wir schreiben

$$1 - \frac{2m(n-m)}{n(n-1)} = 1 - \frac{2}{\frac{n}{m}} \cdot \frac{n-m}{n-1} = 1 - \frac{2\alpha_n}{\frac{n}{m}}, \quad \alpha_n \stackrel{\mathrm{def}}{=} \frac{n-m}{n-1}.$$

Offenbar gilt $\alpha_n \leq 1$, $\forall n \in \mathbb{N}$ und $\alpha_n \uparrow 1$. Aus

$$\log\left(1 + \frac{x}{z}\right) \leq \frac{x}{z}, \quad \forall x > -z,\, z > 0$$

folgt die äquivalente Ungleichung

$$\left(1 + \frac{x}{z}\right)^z \leq e^x, \quad \forall x > -z,\, z > 0.$$

Mit $z = \frac{n}{m}$ und $x = -2\alpha_n$ erhalten wir für alle hinreichend großen n folglich

$$\left(1 - \frac{2\alpha_n}{\frac{n}{m}}\right)^{n/m} \leq e^{-2\alpha_n},$$

woraus

$$\left(1 - \frac{2m(n-m)}{n(n-1)}\right)^{k_n} \leq e^{-2\alpha_n k_n \frac{m}{n}}$$

folgt. Es gilt

$$\begin{aligned} n^2 \cdot e^{-2\alpha_n k_n \frac{m}{n}} &= n^2 \cdot e^{-2\alpha_n(\log n + c)} \\ &= n^2 \cdot e^{-2\frac{n-m}{n-1}\log n} e^{-2\frac{n-m}{n-1}c} \\ &= n^2 \cdot n^{-2\frac{n-m}{n-1}} e^{-2\frac{n-m}{n-1}c} \\ &= \left(n^{\frac{m-1}{n-1}}\right)^2 e^{-2\frac{n-m}{n-1}c}, \end{aligned}$$

woraus unter Berücksichtigung von $\binom{n}{2} = \frac{1}{2}(n^2 - n)$ und $\lim_{n\to\infty} n^{1/n} = 1$

$$\limsup_{n\to\infty} \binom{n}{2}\left(1 - \frac{2m(n-m)}{n(n-1)}\right)^{k_n} \leq \frac{1}{2}e^{-2c}$$

folgt.

Zur umgekehrten Abschätzung nutzen wir

$$\log(1+y) \geq y - y^2, \quad \forall y \in (-\frac{1}{2}, \infty).$$

Diese Ungleichung ergibt sich direkt aus dem Lagrangeschen Restglied der Taylorentwicklung von $y \mapsto \log(1+y)$ bis zur zweiten Ordnung im Punkt $y = 0$. Mit $y = \frac{x}{n}$, $x > -\frac{n}{2}$ folgt hieraus

$$\left(1 + \frac{x}{n}\right)^n \geq e^{x - \frac{x^2}{n}}.$$

Für hinreichend großes n haben wir also mit $x = -2\alpha_n m$

$$\left(1 - \frac{2\alpha_n}{\frac{n}{m}}\right)^{k_n} \geq \left(e^{-2\alpha_n - \frac{4\alpha_n^2 m}{n}}\right)^{\log n + c}.$$

Es gilt

$$e^{-\frac{4\alpha_n^2 m}{n}(\log n + c)} \to 1, \quad n \to \infty,$$

da mit $\alpha_n \uparrow 1$ auch $\alpha_n^2 \uparrow 1$ folgt. Der verbleibende Term $-2\alpha_n(\log n + c)$ ist der gleiche wie bei der ersten Abschätzung. Insgesamt folgt daher

$$\liminf_{n\to\infty} \binom{n}{2}\left(1 - \frac{2m(n-m)}{n(n-1)}\right)^{k_n} \geq \frac{1}{2}e^{-2c},$$

womit das Lemma bewiesen ist. Die bloße Kenntnis von $e^x = \lim_{n\to\infty}(1 + \frac{x}{n})^n$ reicht hier für einen rigorosen Beweis scheinbar nicht aus, daher dieser etwas längere Beweis. □

Im Folgenden sei $n \in \mathbb{N}$ eine gerade Zahl mit $n = 2m$, $m \in \mathbb{N}$. Wir werden nun den (m, m) Shuffle studieren. Dieser kommt dem bereits mehrfach heuristisch ange-

deuteten *Riffle-Shuffle* schon recht nahe, jedoch mit dem Unterschied, dass vor jedem Zusammenfügen das Kartendeck immer *genau* halbiert wird. Satz 2.5.6 liefert wieder eine Abschätzung nach oben, die besagt, dass höchstens $2\log_2 n$ Mischvorgänge nötig sind, um die Separation approximativ verschwinden zu lassen. Dieses folgt aus dem nächsten Lemma, da mit $\nu^1 = \ldots = \nu^k = (m, m)$

$$\|Q_{\nu^k} * \ldots * Q_{\nu^1} - U\| \leq \binom{n}{2}\left(1 - \frac{\frac{n}{2}}{n-1}\right)^k \tag{2.162}$$

gilt.

Lemma 2.5.9. *Für jedes fixierte $c \in \mathbb{R}$ gilt mit $k_n \stackrel{\text{def}}{=} 2\log_2 n + c$*

$$\binom{n}{2}\left(1 - \frac{\frac{n}{2}}{n-1}\right)^{k_n} \longrightarrow \left(\frac{1}{2}\right)^{c+1}, \quad n \to \infty. \tag{2.163}$$

Beweis. Wegen $\binom{n}{2} = \frac{1}{2}n(n-1)$ reicht es offensichtlich

$$n^2\left(1 - \frac{\frac{n}{2}}{n-1}\right)^{k_n} \longrightarrow \left(\frac{1}{2}\right)^{c}, \quad n \to \infty$$

nachzuweisen. Es gilt

$$\frac{1}{2} - \frac{1}{n-1} \leq 1 - \frac{\frac{n}{2}}{n-1} = \frac{\frac{n}{2}-1}{n-1} \leq \frac{1}{2}, \quad n = 2, 3, \ldots\ . \tag{2.164}$$

Einerseits gilt

$$n^2 \cdot \left(\frac{1}{2}\right)^{2\log_2 n + c} = \left(\frac{1}{2}\right)^{c}. \tag{2.165}$$

Andererseits gilt

$$n^2 \cdot \left(\frac{1}{2} - \frac{1}{n-1}\right)^{k_n} = \left(\frac{1}{2} - \frac{1}{n-1}\right)^{c} \overbrace{n^2\left(\frac{1}{2} - \frac{1}{n-1}\right)^{2\log_2 n}}^{\stackrel{\text{def}}{=}\Lambda_n}, \quad n = 4, 5, \ldots.$$

Mit $\log \Lambda_n \to 0$ folgte aus der Stetigkeit der Exponentialfunktion $\Lambda_n \to 1$. Ersteres werden wir nun nachweisen. Es gilt

$$\begin{aligned} \log \Lambda_n &= 2\log n + 2\log_2 n \log\left(\frac{1}{2} - \frac{1}{n-1}\right) \\ &= 2\log n\left(1 + \frac{\log(\frac{1}{2} - \frac{1}{n-1})}{\log 2}\right). \end{aligned}$$

$\log \Lambda_n \to 0$ folgt aus einer Anwendung von dem *Satz von L'Hospital*, da obige Terme ebenso als differenzierbare Funktionen auf $(4, \infty)$ aufgefasst werden können. Es handelt sich um eine Routinerechnung, die daher ausgelassen wird. Es folgt insgesamt

$$n^2\left(\frac{1}{2} - \frac{1}{n-1}\right)^{k_n} \to \left(\frac{1}{2}\right)^{c}, \quad n \to \infty. \tag{2.166}$$

(2.165) zusammen mit (2.164) und (2.166) ergibt die Behauptung. □

Bemerkung 2.5.10. Bei der Untersuchung der beiden letzten Mischverfahren wurde bei der Definition von k auch der Fall $k \in \mathbb{R} - \mathbb{N}$ in Kauf genommen. Aufgrund der Monotonie von (2.161) bzw. (2.162) in k bei Fixierung von m und n bzw. n bereitet die Analyse einer eventuellen Auf- bzw. Abrundung von k keinerlei Probleme. Siehe in diesem Zusammenhang auch das Ende des Beweises von Satz 2.5.13.

Es wird im Folgenden die Definition eines *Riffle-Shuffles* angegeben. Hierbei werden wir drei äquivalente Beschreibungen aufführen, wobei jeweils ein Kartendeck mit n Karten gegeben ist.

Beschreibung 1. Zunächst werde ein von dem späteren Mischvorgang unabhängiges Experiment zur Bestimmung einer Zahl $0 \leq c \leq n$ mit $P(C = c) = \frac{1}{2^n}\binom{n}{c}$ durchgeführt. Für die folgende, recht plastische Beschreibung stelle man sich einen Kartenspieler vor, der vor einem Tisch sitzt und einen *Riffle-Shuffle* ausführen möchte. Dieser Spieler nehme die oberen c Karten des Kartendecks in die linke Hand und die verbleibenden $n-c$ Karten in die rechte Hand. Diese beiden Teilmengen des Kartendecks werden nun innerhalb von n Schritten sukzessive wieder zu einem Kartendeck zusammengefügt.

Im ersten Schritt wird ein Bernoulli Experiment ausgeführt, das mit einer Wahrscheinlichkeit von $\frac{c}{n}$ die unterste Karte der Kartenkonstellation in der linken Hand wählt. Im Komplementärereignis wird dementsprechend die unterste Karte der rechten Hand gewählt. Die gewählte Karte wird auf den Tisch gelegt. Wir nehmen an, dass die Karte in der linken Hand gewählt wurde. In dieser Hand befinden sich daher nun $c - 1$ Karten. Im zweiten Schritt wird die Unterste hiervon mit einer Wahrscheinlichkeit von $\frac{c-1}{n-1}$ auf die Karte, welche wir zuvor auf den Tisch gelegt haben, heraufgelegt. Hierzu wird wieder ein von allen bisherigen Experimenten unabhängiges Bernoulli Experiment durchgeführt. Im Komplementärereignis, das mit einer Wahrscheinlichkeit von $\frac{(n-1)-(c-1)}{n-1}$ eintritt, wird die entsprechende Karte der rechten Hand gewählt. Dieses Vorgehen wird bis zur n-ten Karte fortgesetzt, wobei die Wahrscheinlichkeit, welche Karte gewählt wird, immer proportional zur Kartenanzahl in der entsprechenden Hand ist. Nach dem n-ten Schritt sind beide Hände leer, und auf dem Tisch befindet sich das im Sinne eines *Riffle-Shuffles* permutierte n Kartendeck.

Beschreibung 2. Wir heben von dem n Kartendeck c Karten ab, wobei c wieder die Realisierung einer mit den Parameter $\frac{1}{2}$ binomialverteilten Zufallsvariablen sei. Beide Teildecks sollen nun wieder zu einem zusammengefasst werden, wobei die relative Ordnung der Karten beider Teildecks nach dem Zusammenfügen, d.h. Mischen, erhalten bleiben soll. Hierfür gibt es offensichtlich $\binom{n}{c}$ Möglichkeiten. Von diesen wird eine rein zufällig ausgewählt.

Beschreibung 3. Während Beschreibungen *eins* und *zwei* $\pi \in \mathfrak{S}_n$ generiert haben, generiert diese Beschreibung $\pi^{-1} \in \mathfrak{S}_n$. Es werde hierzu n mal unabhängig voneinander eine faire Münze geworfen. Die i-te Karte wird mit dem Ausgang des i-ten Experimentes beschriftet, z.B. Null für Kopf und Eins für Zahl. Als Nächstes werden alle Karten mit einer Null aus dem Kartendeck herausgezogen und unter Beibehaltung ihrer relativen Ordnung oben auf das Kartendeck heraufgelegt.

Lemma 2.5.11. *Alle drei Beschreibungen generieren dieselbe Wahrscheinlichkeitsverteilung, die sogenannte Gilbert, Shannon, Reeds Verteilung.*

Beweis. Beschreibungen *zwei* und *drei* sind äquivalent, da die binären Markierungen der Karten in Beschreibung *drei* eine binomialverteilte Anzahl von Nullen erzeugt

und, bedingt auf diese, alle möglichen Platzierungen von Nullen und Einsen gleichwahrscheinlich sind. Es ist ebenso klar, dass die relative Ordnung der Karten in Beschreibung *drei* beibehalten wird, denn $\pi^{-1} \in \mathfrak{S}_n$ beschreibt gerade, an welchen Positionen die Karten nach dem Mischvorgang liegen sollen. Direkt nach Konstruktion in Beschreibung *drei* folgt

$$\pi^{-1}(1) < \pi^{-1}(2) < \ldots < \pi^{-1}(c), \quad \pi^{-1}(c+1) < \pi^{-1}(c+2) < \ldots < \pi^{-1}(n).$$

Obige Relationen besagen gerade, dass die relative Ordnung der Karten beibehalten wird. Ferner sei noch erwähnt, dass, falls $\pi^{-1}(c) > \pi^{-1}(c+1)$, wir genau zwei aufsteigende Sequenzen haben und ansonsten nur eine, was offenbar nur im Fall $\pi = \pi^{-1} = \mathrm{id}$ möglich ist.

Beschreibungen *eins* und *zwei* sind äquivalent, da ein Kartenmischvorgang in Beschreibung *eins* nach gegebenem c durch ein Element der Menge $\{L, R\}^n$ charakterisiert ist, wobei die Einträge eines jeden Tupels indizieren, ob in dem entsprechenden Schritt die linke oder rechte Hand gewählt wird. In diesem Modell ist die Wahrscheinlichkeit einer Permutation $\pi \in \mathfrak{S}_n$, die mit diesem Mischvorgang erzeugt werden kann, gerade das Produkt der Wahrscheinlichkeiten der jeweiligen Bernoulliexperimente in den einzelnen Schritten. Nach einer eventuellen Umordung der Faktoren sehen wir, dass diese gerade

$$\frac{c!(n-c)!}{n!} = \frac{1}{\binom{n}{c}}$$

beträgt, womit die Äquivalenz erwiesen ist. □

Lemma 2.5.12. *Q_R bezeichne obige Gilbert, Shannon, Reeds Verteilung. Dann gilt*

$$Q_R(\pi) = \begin{cases} \frac{1}{2^n} & \text{falls } \pi \text{ genau zwei aufsteigende Sequenzen besitzt,} \\ \frac{n+1}{2^n} & \text{falls } \pi = id. \end{cases} \tag{2.167}$$

Insgesamt kann ein Riffle-Shuffle genau $2^n - n$ verschiedene Permutationen generieren.

Beweis. Offenbar gibt es genau $n+1$ Möglichkeiten, als Mischergebnis die Identität zu erzeugen. Im Sinne von Beschreibung *zwei* werden hierzu $0 \leq i \leq n$ Karten abgehoben und wieder genauso auf den unteren Stapel zurückgelegt. Jede andere erreichbare Permutation kann allerdings nur auf *genau eine* Art und Weise erreicht werden. Da es insgesamt 2^n Mischvorgänge gibt (jede Karte wird in Beschreibung 3 entweder nach oben gelegt oder nicht) und jeder davon per Definition des *Riffle-Shuffles* gleichwahrscheinlich sein soll, folgt sofort (2.167).

Die Anzahl der durch einen *Riffle-Shuffle* erreichbaren Permutationen ist offensichtlich

$$1 + \sum_{c=0}^{n} \left(\binom{n}{c} - 1 \right) = 2^n - n,$$

wobei jeder Summand repräsentativ für alle *Riffle-Shuffles*, bei denen genau c Karten abgehoben wurden, steht. Hierbei muss jedes Mal $\mathrm{id} \in \mathfrak{S}_n$ abgezogen werden, damit es nicht mehrfach mitgezählt wird. □

Im weiteren Verlauf wird Beschreibung *drei* benutzt werden, wobei dortige Beschreibung schon gleich als Generation von π und nicht von π^{-1} aufgefasst werde, was wegen Lemma 2.5.1 erlaubt ist.

Folgender Satz zeigt, dass es $2\log_2 n$ *Riffle-Shuffles* bedarf, um ein vorsortiertes n Kartendeck im Sinne asymptotisch verschwindender *Separation* zu durchmischen. Ein *Cutoff-Effekt* ist wieder klar zu erkennen.

Satz 2.5.13. *Sei $\eta \in \mathbb{R}$ fixiert und $k_n \stackrel{\text{def}}{=} 2\log_2 n + \eta$. Dann gilt*

$$1 - e^{-2^{-\eta-1}} \leq \liminf_{n\to\infty} \operatorname{sep}(Q_R^{*\lfloor k_n \rfloor}, U) \leq \limsup_{n\to\infty} \operatorname{sep}(Q_R^{*\lfloor k_n \rfloor}, U) \leq 1 - e^{-2^{-\eta}}. \quad (2.168)$$

Beweis. Ähnlich wie in (2.157) können wir k voneinander unabhängige, sukzessive *Riffle-Shuffles* realisieren. Die zugehörige Matrix $A^k \in \mathbb{N}^{n\times k}$ entsteht hierbei durch $n \cdot k$ unabhängig voneinander ausgeführte, faire Münzwürfe, wobei wir jedem Matrixeintrag $(A_{ij}^k)_{1\leq i\leq n,\, 1\leq j\leq k}$ ein Münzwurfexperiment zuordnen und $A_{ij}^k = 1$ bzw. $A_{ij}^k = 2$ setzen, falls zugehöriges Experiment Kopf bzw. Zahl ergibt. So erhalten wir sukzessive Matrizen $A^1, A^2, A^3, \ldots$, wobei $(A_{ij}^{k+1})_{1\leq i\leq n,\, 1\leq j\leq k} = (A_{ij}^k)_{1\leq i\leq n,\, 1\leq j\leq k}$ gilt. A^{k+1} ist also die Fortsetzung von A^k und wird *nicht* komplett neu durch Münzwurfexperimente kreiert. Nur die $k+1$-te Spalte wird durch n weitere Bernoulli Experimente erzeugt. Es handelt sich also sozusagen um einen $\nu = (\nu^1, \ldots, \nu^k)$ Shuffle, wobei $\nu^i = (\nu_1^i, n - \nu_1^i)$ und ν_1^i die Realisierung einer $B(n, \frac{1}{2})$ verteilten Zufallsvariablen ist. Es ist allerdings *kein* ν Shuffle im Sinne von Definition 2.5.2 (3), da dort jedes ν^i fixiert ist. Dennoch lässt sich der Beweis von Satz 2.5.6 imitieren.

Wir definieren auf $(\Omega, \mathfrak{A}, P)$ hierzu wieder nach vertrautem Algorithmus $(X_k)_{k\geq 0}$ als Funktion von $(A^k)_{k\geq 1}$. $(X_k)_{k\geq 0}$ ist daher per Konstruktion ein *Random Walk* auf $\mathfrak{S}_n$, wobei die Übergangsmatrix in der Notation von Lemma 2.5.1 von $\widetilde{Q_R}$ generiert wird und $X_0 = \mathrm{id}$ gilt. Sei weiter

$$\begin{aligned} \mathcal{F}_0 &\stackrel{\text{def}}{=} \{\emptyset, \Omega\}, \\ \mathcal{F}_k &\stackrel{\text{def}}{=} \sigma(A^m,\, 1 \leq m \leq k), \quad k \geq 1. \end{aligned}$$

Es ist dann $(X_k)_{k\in\mathbb{N}_0}$ offenbar eine DMK bzgl. der Filtration $(\mathcal{F}_k)_{k\in\mathbb{N}_0}$. Wir definieren wieder T als den frühsten Zeitpunkt, zu dem alle n Zeilen von A^k verschieden sind, d.h.

$$T \stackrel{\text{def}}{=} \inf\{k \geq 1 \,:\, A_i^k \neq A_j^k, \quad \forall i \neq j\}.$$

T ist auch hier eine P-f.s. endliche $(\mathcal{F}_k)_{k\in\mathbb{N}_0}$ Stoppzeit. Genau wie im Satz 2.5.6 ist es entscheidend, dass T eine *stark stationäre Zeit* ist, woraus nach Satz 1.2.13 zusammen mit $P^{X_k} = \big(\widetilde{Q_R}\big)^{*k}$, $k \in \mathbb{N}_0$ unter Berücksichtigung von Lemma 2.5.1 erneut

$$\operatorname{sep}(Q_R^{*k}, U) = \operatorname{sep}(\big(\widetilde{Q_R}\big)^{*k}, U) \leq P(T > k), \quad k \in \mathbb{N}_0 \quad (2.169)$$

folgt. Im Gegensatz zu (2.159) werden wir hier auf die subadditive Abschätzung verzichten und direkt den Ausdruck $P(T > k)$ auswerten. Hierzu bemerken wir zunächst, dass T eine *schnellste stark stationäre Zeit* ist, m.a.W. in (2.169) sogar Gleichheit besteht. Dies liegt daran, dass

$$\tau = \begin{pmatrix} 1 & 2 & \ldots & n \\ n & n-1 & \ldots & 1 \end{pmatrix} \in \mathfrak{S}_n \quad (2.170)$$

frühstens zum Zeitpunkt T erreicht werden kann oder formaler

$$\{X_k = \tau\} \subseteq \{T \leq k\}, \quad k \in \mathbb{N}_0$$

gilt. Hieraus folgt

$$P(X_k = \tau) = P(X_k = \tau, T \leq k),$$

was nach einem kurzen Blick auf den Beweis von Satz 1.2.13 schließlich die Gleichheit in (2.169) ergibt.

Insgesamt bleibt hier folglich nur noch $P(T > k)$ zu berechnen. Diese Aufgabe ist offensichtlich äquivalent damit, die Wahrscheinlichkeit $P(T \leq k)$ zu bestimmen, n Kugeln auf 2^k Zellen ohne Doppeltbesetzungen zu verteilen, wenn jede Kugel unabhängig von einer jeden anderen Kugel mit völliger Willkür in die 2^k Zellenanordnung gelegt wird. Jede Zelle entspricht hierbei einer möglichen Realisierung einer Zeile von A^k. Jede Zeilenrealisierung ist wegen der unabhängigen, fairen, bereits zuvor erwähnten $k \cdot n$ Münzwurfexperimente gleichwahrscheinlich. Die n Zeilen bilden ferner eine Familie unabhängiger Zufallsvariablen. Insgesamt haben wir also nichts anderes als das klassische Geburtstagsproblem. Hierbei interpretiere man die 2^k Zellen als mögliche Geburtstage und n als die Menschenanzahl, die auf zeitgleiche Geburtstage untersucht werden soll. Das Ergebnis ist wohlbekannt und lautet

$$\operatorname{sep}(Q_R^{*k}, U) = P(T > k) = 1 - P(T \leq k) = 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right), \quad k \in \mathbb{N}. \quad (2.171)$$

Der Leser beachte, dass für $2^k < n$ obiges Produkt verschwindet und $\operatorname{sep}(Q_R^{*k}, U) = 1$ folgt. Es bleibt für die Grenzbetrachtung (2.168) nur noch

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{2^{k_n}}\right) \longrightarrow e^{-2^{-(\eta+1)}}, \quad k_n = 2\log_2 n + \eta,\ n \to \infty$$

zu zeigen, da offenbar

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{2^{k-1}}\right) \leq \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^{\lfloor k \rfloor}}\right) \leq \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right), \quad \forall n \in \mathbb{N},\ k \in \mathbb{R} : 2^{k-1} > n$$

gilt. Der etwas kompliziert erscheinende Ausdruck (2.168) resultiert also ausschließlich daher, dass wir nur eine *ganze* Anzahl von Mischvorgängen absolvieren können. Der Beweis ergibt sich hiermit aus folgendem Lemma. □

Lemma 2.5.14. *Sei $\eta \in \mathbb{R}$ fixiert und $k_n = 2\log_2 n + \eta$. Dann gilt*

$$\prod_{i=1}^{n-1} \left(1 - \frac{i}{2^{k_n}}\right) \longrightarrow e^{-2^{-(\eta+1)}}, \quad n \to \infty. \quad (2.172)$$

Beweis. Der linke Term in (2.172) ist für hinreichend großes n offenbar strikt positiv. Wegen der Stetigkeit der Exponentialfunktion reicht es daher

$$\sum_{i=1}^{n-1} \log\left(1 - \frac{i}{2^{k_n}}\right) \longrightarrow -2^{-(\eta+1)}, \quad n \to \infty$$

zu zeigen. Aus $\log(1+x) \leq x$, $\forall x \in (-1, \infty)$ folgt einerseits für alle hinreichend großen $n \in \mathbb{N}$

$$\sum_{i=1}^{n-1} \log\left(1 - \frac{i}{2^{k_n}}\right) \leq -\sum_{i=1}^{n-1} \frac{i}{2^{k_n}} = -\frac{1}{2^{k_n}} \cdot \frac{1}{2}(n-1)n = -\frac{1}{2^{\eta+1}} \cdot \frac{1}{n^2}(n-1)n, \quad (2.173)$$

woraus sich

$$-2^{-(\eta+1)} \geq \limsup_{n\to\infty} \log\left(\prod_{i=1}^{n-1}\left(1-\frac{i}{2^{k_n}}\right)\right)$$

ergibt. Aus $\log(1+x) \geq x - x^2$, $\forall x \in (-\frac{1}{2}, \infty)$ (siehe den zweiten Teil des Beweises von Lemma 2.5.8) folgt andererseits

$$\sum_{i=1}^{n-1} \log\left(1-\frac{i}{2^{k_n}}\right) \geq -\sum_{i=1}^{n-1}\left(\frac{i}{2^{k_n}}+\frac{i^2}{2^{2k_n}}\right) \geq -\frac{1}{2^{k_n}}\sum_{i=1}^{n-1} i - \frac{n^3}{2^{2k_n}}$$

für alle *hinreichend großen* $n \in \mathbb{N}$, da dann

$$\begin{aligned} 1-\frac{i}{2^{k_n}} &> 0, \quad i = 1,\ldots,n-1 \text{ wegen } 2^{k_n} = 2^\eta \cdot n^2 \text{ und} \\ n^3 &\geq \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6} = \frac{1}{3}n^3 + O(n^2) \cong \frac{1}{3}n^3 \end{aligned}$$

gilt. Wegen $2^{2k_n} = 2^{2\eta} \cdot n^4$ folgt

$$\frac{n^3}{2^{2k_n}} = 2^{-2\eta}\frac{1}{n} \to 0, \quad n \to \infty,$$

was insgesamt unter erneuter Betrachtung der Gleichungen in (2.173)

$$\begin{aligned} -2^{-(\eta+1)} &\leq \liminf_{n\to\infty} \log\left(\prod_{i=1}^{n-1}\left(1-\frac{i}{2^{k_n}}\right)\right) \\ &\leq \limsup_{n\to\infty} \log\left(\prod_{i=1}^{n-1}\left(1-\frac{i}{2^{k_n}}\right)\right) \\ &\leq -2^{-(\eta+1)} \end{aligned}$$

ergibt, also dieses Lemma und damit endgültig Satz 2.5.13 beweist. □

Bemerkung 2.5.15. Es kann (2.171) auch direkt aus der Verteilung eines *Riffle-Shuffles* gefolgert werden. Es gilt nämlich

$$Q_R^{*k}(\pi) = \frac{\binom{2^k+n-r(\pi)}{n}}{2^{kn}}, \quad k \in \mathbb{N}, \tag{2.174}$$

wobei $r(\pi)$ die Anzahl der aufsteigenden Sequenzen von $\pi \in \mathfrak{S}_n$ bezeichnet. Eine Herleitung von (2.174) befindet sich in Bayer, Diaconis [6]. Sie ist elementar und folgt letztendlich aus der Betrachtung des Faltungsverhaltens sogenannter *a-Shuffles*, $a \in \mathbb{N}$, einer naheliegenden Verallgemeinerung, bei der nicht zwei, sondern a Kartendecks auf eine dem hier definierten *Riffle-Shuffle* ($a = 2$) analoge Art und Weise zusammengefügt werden. Wir verzichten auf weitere Details und kommen durch folgende Umformungen direkt auf den Punkt.

$$\begin{aligned} \operatorname{sep}(Q_R^{*k}, U) &= \max_{\pi\in\mathfrak{S}_n}\left\{1-\frac{Q_R^{*k}(\pi)}{U(\pi)}\right\} \\ &= 1 - \min_{\pi\in\mathfrak{S}_n}\{n!Q_R^{*k}(\pi)\} \end{aligned}$$

$$
\begin{aligned}
&= 1 - \frac{1}{2^{kn}} \min_{\pi \in \mathfrak{S}_n} \left\{ \frac{(2^k + n - r(\pi))!}{(2^k - r(\pi))!} \right\} \\
&= 1 - \frac{1}{2^{kn}} \min_{\pi \in \mathfrak{S}_n} \{(2^k - r(\pi) + 1) \cdot \ldots \cdot (2^k - r(\pi) + n)\} \\
&= 1 - \frac{1}{2^{kn}} \prod_{i=0}^{n-1} (2^k - i) \\
&= 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^k}\right),
\end{aligned}
$$

wobei dem Element $\tau \in \mathfrak{S}_n$ mit $r(\tau) = n$ aus (2.170) auch hier wieder eine besondere Bedeutung zukommt. Dieses bildet gerade das Maximum in obiger Separation.

Wir konnten einen *Cutoff-Effekt* bzgl. der Separation relativ leicht nachweisen, und machen darauf aufmerksam, dass ebenso ein *Cutoff-Effekt* bzgl. des Variationsabstandes besteht, dieser aber wesentlich schwerer nachzuweisen ist. Diesbezüglich wird das Ergebnis von Bayer, Diaconis [6] zitiert.

Theorem 2.5.16. *Für jedes fixierte $c \in \mathbb{R}$ gilt*

$$
\|Q_R^{*k_n} - U\| = 1 - 2\Phi\left(\frac{-2^{-c}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{1/4}}\right), \quad n \to \infty, \tag{2.175}
$$

wobei $k_n \stackrel{\text{def}}{=} \frac{3}{2}\log_2 n + c$ und $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt$.

Es werden also $\frac{3}{2}\log_2 n$ Shuffles benötigt, um den Variationsabstand approximativ verschwinden zu lassen. Aus der allgemeinen Theorie über Random Walks auf Gruppen (siehe Abschnitt 1.3, Satz 1.3.7) kann hieraus gefolgert werden, dass auf jeden Fall eine Größenordnung von $2 \cdot \frac{3}{2}\log_2 n$ Shuffles ausreicht, um den Separationsabstand approximativ Null anzunähern. Wir wissen jetzt aber, dass tatsächlich schon $2\log_2 n$ Shuffles ausreichen.

Zum Beweis von Theorem 2.5.16 sei hier noch angemerkt, dass der Ansatz in einer direkten Berechnung von $\|Q_R^{*k} - U\|$ mit Hilfe von (2.174) besteht. Das Problem ist es, diesen Ausdruck geschickt abzuschätzen. Da für einen *Riffle-Shuffle* die Anzahl der aufsteigenden Sequenzen von zentraler Bedeutung ist, wie schon (2.174) zeigt, stößt man zwangsläufig auf die sogenannten Eulerzahlen (siehe Knuth [14, S.35ff]). Ganz entscheidend ist, dass diese in Zusammenhang mit Wahrscheinlichkeiten gewisser Ereignisse von Summen unabhängiger, identisch verteilter Zufallsvariablen gebracht werden können (siehe Tanny [18]). So kommt der zentrale Grenzwertsatz ins Spiel, und die Gaußform in (2.175) wird verständlich.

Zum Schluss wird noch eine weitere Abänderung des *Riffle-Shuffles* untersucht, die sich in dem hier definierten *Riffle-Shuffle* genau dadurch unterscheidet, dass vor dem Zusammenfügen der beiden Kartenstapel der zuvor abgehobene gründlich durchmischt wird. Präziser handelt es sich um einen Q_μ Shuffle mit $\mu = \mathrm{Bin}(n, \frac{1}{2})$ in der Notation von (2.5). Wir sind hier in der Lage, einen *Cutoff-Effekt* bzgl. des Variationsabstandes nachzuweisen. Intuitiv würde man sicher erwarten, dass obige Abänderung wesentlich schneller gegen die Gleichverteilung auf $\mathfrak{S}_n$ konvergiert als der *Riffle-Shuffle* in (2.167). In der Tat ist diese Abänderung auch schneller, allerdings nur um einen konstanten Faktor, da der *Cutoff-Effekt* bei $\log_2 n$ Mischschritten auftritt, währenddessen dieser beim *Riffle-Shuffle* erst nach $\frac{3}{2}\log_2 n$ Schritten erscheint. Grob formuliert ist der *Riffle-Shuffle* folglich *nur* 50% langsamer als die

Variante, bei der die zurückgesteckten Karten ihre relative Ordnung *nicht* beibehalten müssen. Obige Prosa wird nun formalisiert.

Lemma 2.5.17. *Sei* $0 < p_1 \leq p_2 < 1$. *Dann gilt* $\mathrm{Bin}(n, p_2) \leq \mathrm{Bin}(n, p_1)$ *im Sinne von Definition 2.3.16.*

Beweis. Seien f_1 bzw. f_2 die Dichten bzgl. des Zählmaßes von $\mathrm{Bin}(n, p_1)$ bzw. $\mathrm{Bin}(n, p_2)$. Dann ist

$$i \mapsto \frac{f_1(i)}{f_2(i)} = \frac{p_1^i}{p_2^i} \cdot \frac{(1-p_1)^{n-i}}{(1-p_2)^{n-i}}$$

eine fallende Funktion. Die Behauptung folgt damit aus Lemma 2.3.17. □

Nun können wir den gewünschten Satz formulieren und beweisen.

Satz 2.5.18. *Sei* $\mu = \mathrm{Bin}(n, p)$ *für ein* $p \in (0, 1)$. *Es gilt dann mit* $k_n \overset{\mathrm{def}}{=} -\frac{\log n}{\log(1-p)} + c$ *für fixiertes* $c \in \mathbb{R}$ *und* v *wie in (2.86)*

$$v((1-p)^c) \leq \liminf_{n\to\infty} \|Q_\mu^{*\lfloor k_n \rfloor} - U\| \leq \limsup_{n\to\infty} \|Q_\mu^{*\lfloor k_n \rfloor} - U\| \leq v((1-p)^{c-1}). \quad (2.176)$$

Beweis. Aus Lemma 2.1.10 folgt

$$\overbrace{\mu \sharp \ldots \sharp \mu}^{s-\mathrm{mal}} = \mathrm{Bin}(n, 1-(1-p)^s).$$

Sei $\mu_r \overset{\mathrm{def}}{=} \mathrm{Bin}(n, 1-(1-p)^r)$, $r \in \mathbb{R}^{>0}$. Dann ist nach Lemma 2.5.17 $r \mapsto \mu_r$ eine fallende Funktion, falls wir die Menge der Verteilungen auf $\{0, \ldots, n\}$ mit der partiellen Ordnung bzgl. stochastischer Dominiertheit betrachten. Nach Satz 2.3.18 gilt folglich

$$\|Q_{\mu_{k_n}} - U\| \leq \|Q_{\mu_{\lfloor k_n \rfloor}} - U\| \leq \|Q_{\mu_{k_n - 1}} - U\|. \quad (2.177)$$

Es ist $(1-p)^{k_n} = \frac{(1-p)^c}{n}$. Wegen

$$\mu_{k_n}(n-i) = \mathrm{Bin}\left(n, \frac{(1-p)^c}{n}\right)\{i\}, \quad i = 0, \ldots, n$$

für hinreichend große $n \in \mathbb{N}$ gilt

$$\mu_{k_n}(n-i) \longrightarrow \mathrm{Poi}((1-p)^c)\{i\}, \quad n \to \infty.$$

Aus Theorem 2.3.4 folgt daher

$$\lim_{n\to\infty} \|Q_{\mu_{k_n}} - U\| = v((1-p)^c).$$

Die Behauptung ergibt sich daher direkt aus (2.177). □

Bemerkung 2.5.19. Obiger Satz wurde verbal nur für $p = \frac{1}{2}$ diskutiert. Der allgemeine Fall bereitet aber keine weiteren Komplikationen, weshalb er hier hinzugezogen wurde. Es sei noch erwähnt, dass die Gleichung

$$-\frac{1}{\log(1-p)} = \frac{3}{2} \cdot \frac{1}{\log 2}$$

die eindeutige Lösung $p = 1 - 2^{-\frac{2}{3}} \cong 0.37$ besitzt. Heben wir im Schnitt also immer etwas mehr als ein Drittel der Karten ab, so sind wir in derselben Größenordnung

wie der *Riffle-Shuffle*. Trotzdem muss man sich natürlich darüber im Klaren sein, dass $v(e^{-c})$ und der Gaußförmige Übergang in Theorem 2.5.16 verschieden sind. Wir haben also immer noch zwei unterschiedliche Arten des Phasenübergangs, und durch unseren modifizierten Parameter $p \widetilde{=} 0.37$ wird lediglich die Übergangsschwelle des einen Phasenübergangs quasi künstlich in die des anderen hineingeschoben.

Literaturverzeichnis

[1] Aldous, D., Diaconis, P. (1986). Shuffling cards and stopping times. *Amer. Math. Monthly* **2**, 333-348.

[2] Aldous, D., Diaconis, P. (1987). Strong uniform times and finite random walks. *Advances in applied mathematics* **8**, 69-97.

[3] Alsmeyer, G. (2005). Wahrscheinlichkeitstheorie (4. Auflage). Skripten zur Mathematischen Statistik, Nr.**30**. *Universität Münster*

[4] Alsmeyer, G. (2004). Mathematische Statistik (2. Auflage). Skripten zur Mathematischen Statistik, Nr.**36**. *Universität Münster*

[5] Alsmeyer, G. (2004). Stochastische Prozesse, Teil 1 (3. Auflage). Skripten zur Mathematischen Statistik, Nr.**33**. *Universität Münster*

[6] Bayer, D., Diaconis, P. (1992). Trailing the dovetail shuffle to its lair. *The Annals of Applied Probability* **2**, 294-313.

[7] Diaconis, P., Fill, J.A. (1990). Strong stationary times via a new form of duality. *The Annals of Probability* **4**, 1483-1522.

[8] Diaconis, P., Fill, J.A., Pitman, J. (1992). Analysis of top to random shuffles. *Combinatorics, Probability and Computing* **1**, 135-155.

[9] Feller, W. (1968). An introduction to probability theory and its applications, Vol. 1 (3. edition). *published by Wiley*

[10] Fischer, G. (2002). Lineare Algebra: Eine Einführung für Studienanfänger (13. Auflage). *Vieweg Verlag*

[11] Forster, O. (2001). Differential- und Integralrechnung einer Veränderlichen (6. Auflage). *Vieweg Verlag*

[12] Henze, N. (2000). Stochastik für Einsteiger (3. Auflage). *Vieweg Verlag*

[13] Holst, L. (1980). On matrix occupancy, committee, and capture-recapture problems. *Scand J Statist* **7**, 139-146.

[14] Knuth, D. E. (1998). The art of computer programming, Vol. 3 (2. edition). *published by addison-wesley*

[15] Lang, S. (2005). Algebra (3. edition). *published by Springer*

[16] Pierce, R.S. (1982). Associative algebras (1. edition). *published by Springer*

[17] Rudin, W. (1999). Reelle und komplexe Analysis (3. Auflage). *Oldenbourg Verlag*

[18] Tanny, S. (1973). A probabilistic interpretation of Eulerian numbers. *Duke Math. J.* **40**, 717-722. [Correction (1974) **41** 689.]

[19] Woess, W. (1980). Aperiodische Wahrscheinlichkeitsmaße auf topologischen Gruppen. *Mh. Math.* **90**, 339-345.

Symbolverzeichnis

$\mathbb{N}$	Menge der natürlichen Zahlen $1, 2, 3, \ldots$
$\mathbb{N}_0$	$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$
$\lfloor x \rfloor$	größte ganze Zahl $\leq x$, Gaußklammer
$\lceil x \rceil$	kleinste ganze Zahl $\geq x$
$\log_2$	Logarithmus zur Basis 2
$x^{(k)}$	Produkt $x(x-1) \cdot \ldots \cdot (x-k+1)$
$L\lambda$	Abkürzung für $\log \lambda$
$\Delta(\lambda)$	Abkürzung für $\frac{\lambda}{L\lambda} - \lfloor \frac{\lambda}{L\lambda} \rfloor$
$x \wedge y$	$\min\{x, y\}$
$x \vee y$	$\max\{x, y\}$
U	Gleichverteilung auf $\mathfrak{S}_\mathrm{n}$
δ_k, $0 \leq k \leq n$	Diracverteilung auf $\{0, 1, \ldots, n\}$ mit Masse im Punkt k
$\mathrm{Bin}(n, p)$	Binomialverteilung zu den Parametern n, p
$\mathrm{Poi}(\lambda)$, $\lambda > 0$	Poissonverteilung zu dem Parameter λ
$\mathrm{Geo}(\theta)$, $0 < \theta \leq 1$	geometrische Verteilung zu dem Parameter θ
Q_m	Verteilung eines Top-m-to-Random-Shuffles
Q_μ	Verteilung eines μ Shuffles, μ Verteilung auf $\{0, 1, \ldots, n\}$
Q_ν	Verteilung eines ν Shuffles, ν Komposition von n
Q_R	Gilbert, Shannon, Reeds Verteilung des Riffle-Shuffles
P_k^n	Verteilung der Anzahl *unbesetzten* Zellen einer n Zellenstruktur nach k-facher zufälliger Belegung einer Kugel
$P_{m_1,\ldots,m_k}$	Verteilung der Anzahl *unbesetzten* Zellen einer n Zellenstruktur gemäß dem Experiment von Satz 2.1.9
$\nu \sharp \mu$	Verteilung der Anzahl *besetzten* Zellen einer n Zellenstruktur gemäß dem Experiment vor Lemma 2.1.3
$\mu \leq \nu$	ν ist stochastisch kleiner als μ
P_λ	zugrundeliegendes Wahrscheinlichkeitsmaß einer Markov-Kette mit Anfangsverteilung λ
$\|A - B\|$	Variationsabstand zweier Verteilungen A und B
$d(k)$, $k \geq 0$	Variationsabstand zwischen der stationären Verteilung einer Markov-Kette und ihrer Verteilung zum Zeitpunkt k
$\mathrm{sep}(A, B)$	Separationsabstand zweiter Verteilungen A und B
$s(k)$, $k \geq 0$	Separationsabstand zwischen der stationären Verteilung einer Markov-Kette und ihrer Verteilung zum Zeitpunkt k
$\mathfrak{S}_\mathrm{n}$	symmetrische Gruppe
$\mathfrak{A}_n$	alternierende Gruppe
$\mathfrak{V}_4$	Kleinsche Vierergruppe

$\mathcal{P}(\mathfrak{S}_n)$	Potenzmenge von $\mathfrak{S}_n$
$\mathrm{sgn}(\pi)$	Signum der Permutation $\pi \in \mathfrak{S}_n$
$\langle\sigma\rangle$	zyklische Gruppe, die von $\sigma \in \mathfrak{S}_n$ erzeugt wird
$L(\mathfrak{S}_n)$	Menge der Abbildung von $\mathfrak{S}_n$ auf $\mathbb{Q}$
$\mathbb{Q}[\mathfrak{S}_n]$	$\mathbb{Q}$-Algebra, die unter Definition 2.4.4 beschrieben wird
$\mathbb{M}$	Menge der gruppenzirkulären Matrizen
$Q_2 \circ Q_1$	Faltung zweier Verteilungen Q_1, Q_2 auf einer Gruppe G
$\langle v_1, \ldots, v_r\rangle_{\mathbb{Q}}$	kleinster $\mathbb{Q}$-Vektorraum, der die Vektoren $v_1, \ldots, v_r$ enthält
$\mathrm{Sp}(A)$	Spur der Matrix A
$\mathbb{P}_j$, $0 \leq j \leq n$	Übergangsmatrix des Top-j-to-Random-Shuffles
$\mathbb{P}$	Übergangsmatrix des Top-to-Random-Shuffles (Kapitel 2.4)
$L(\pi)$	Länge der aufsteigenden Sequenz von π, die n enthält
$F(\pi)$	erster Abstieg von π
$G(\pi)$	letzter Abstieg von π

Abkürzungsverzeichnis

DMK	diskrete Markov-Kette
TmTRS	Top-m-to-Random-Shuffle
T1TRS	Top-to-Random-Shuffle
ITmTRS	inverser Top-m-to-Random-Shuffle
IBmTRS	inverser Bottom-m-to-Random-Shuffle

Ich versichere, dass ich diese Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder Sinn nach entnommen wurden, habe ich in jedem Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht.

Münster, 15. März 2010

(Christian Palmes)