

Westfälische Wilhelms-Universität Münster

Institut für Mathematische Statistik

Riffle Shuffle und Cut-Off-Effekt

Diplomarbeit

vorgelegt von

Walter Peter Sendfeld

Thema gestellt von

Professor Dr. G. Alsmeyer

15. März 2005

Inhaltsverzeichnis

Einleitung	v
1 Random Walks auf Gruppen	1
1.1 Random Walks auf endlichen Gruppen	1
1.2 Die Gruppe \mathcal{S}_n der Permutationen	6
1.3 Beschreibung eines Mischvorgangs	8
2 Der Riffle Shuffle	13
2.1 Modelle für den Riffle Shuffle	13
2.2 Aufsteigende Sequenzen	32
2.3 Eulersche Zahlen und aufsteigende Sequenzen	42
2.4 Konvergenz gegen die Gleichverteilung: Der Cut-Off-Effekt	49
2.5 Obere und untere Schranke für die Totalvariation	53
2.6 Nachweis des Cut-Off-Effekts für den Riffle Shuffle	58
Symbolverzeichnis	83
Literaturverzeichnis	85

Einleitung

Der *Riffle¹* *Shuffle²* oder *Dovetail³* *Shuffle* ist eine weit verbreitete Methode, einen Kartenstapel zu mischen. Dabei werden die Karten nach einer Spielrunde in zwei ungefähr gleich große Päckchen geteilt. Von diesen wird jeweils eines in die linke und eines in die rechte Hand genommen. Dann werden eine oder mehrere Karten abwechselnd aus beiden Händen fallengelassen und die Karten so sukzessive ineinandergeblättert. Durch wiederholtes Mischen soll die Reihenfolge der Karten so verändert werden, dass keine der am Spiel beteiligten Personen diese vorhersagen kann. Wie oft der Kartenstapel gemischt werden muss, um eine ausreichende Zufälligkeit der Kartenreihenfolge zu erreichen, ist die zentrale Frage. Am 09.01.1990 war in der New York Times [13] die Antwort in Form der folgenden Schlagzeile zu lesen:

„In card shuffling, 7 is winning number“.

Präziser formuliert, sind für einen Kartenstapel mit 52 Karten sieben Mischvorgänge nötig, um den Kartenstapel hinreichend zu mischen. Kartenspiele mit 52 Karten sind beispielsweise Blackjack, Bridge und Schafkopf. Dem Zeitungsartikel lagen die Forschungsergebnisse von David Bayer und Persi Diaconis zu Grunde. In ihrem im Jahre 1992 veröffentlichten Artikel „Trailing the Dovetail Shuffle to its lair“ (siehe [6]) stellten sie verschiedene wahrscheinlichkeitstheoretische Modelle für den Riffle Shuffle vor. Sie wiesen nach, dass für einen Kartenstapel mit n Karten $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ Mischvorgänge in einem gewissen Sinne ausreichen. Darüber hinaus

¹riffle (engl.): durchblättern

²shuffle (engl.): Mischen, Mischvorgang

³dovetail (engl.): Schwanzschwanz

erlaubten ihre Ergebnisse eine weitere erstaunliche Aussage: Einerseits sind weniger als $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ Mischvorgänge nicht ausreichend und andererseits mehr als $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ Mischvorgänge nicht notwendig, um einen Kartenstapel mit n Karten ausreichend zu mischen. Dieses Phänomen wird als *Cut-Off-Effekt* bezeichnet.

Das Ziel dieser Arbeit ist die wahrscheinlichkeitstheoretische Erfassung und Präzisierung des Riffle Shuffles und Cut-Off-Effekts. Wir orientieren uns dabei maßgeblich an den Ergebnissen von Bayer und Diaconis [6].

Das sukzessive Mischen von Karten wird als stochastischer Prozess $X = (X_m)_{m \geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ interpretiert. Dabei gibt die Zufallsgröße X_m für $m \geq 0$ die Kartenreihenfolge nach m sukzessiven Mischvorgängen in geeigneter Weise an. Wir werden zeigen, dass der konstruierte Prozess X ein Random Walk und zugleich eine zeitlich homogene, irreduzible und aperiodische Markov-Kette ist. Der Zustandsraum des Prozesses X , der aus allen möglichen Kartenreihenfolgen von n Karten besteht, ist dabei die Gruppe der Permutationen \mathcal{S}_n . Die eindeutig bestimmte stationäre Verteilung ist die Gleichverteilung $U_{\mathcal{S}_n}$ auf \mathcal{S}_n . Ein m -mal gemischter Kartenstapel wird als hinreichend gemischt angesehen, wenn alle Kartenreihenfolgen annähernd gleichwahrscheinlich sind, das heißt die Totalvariation $\|P^{X_m} - U_{\mathcal{S}_n}\|$ von P^{X_m} und der Gleichverteilung $U_{\mathcal{S}_n}$ auf \mathcal{S}_n hinreichend klein ist. Der Ergodensatz rechtfertigt dieses Vorgehen. Für $n = 52$ und $m = 1, \dots, 20$ führt eine Berechnung der Totalvariation zu der Darstellung in Abbildung 1, die den Cut-Off-Effekt veranschaulicht.

Mit Blick auf Abbildung 1 ist der Cut-Off-Effekt ein „Phasen-Übergang“: Vom ersten bis zum fünften Mischvorgang befindet sich die Totalvariation nah bei ihrem maximalen Wert von 1. Während der nächsten vier Mischvorgänge fällt sie jedoch rapide auf einen Wert nahe 0.

Der Nachweis des Cut-Off-Effekts erweist sich im Allgemeinen als äußerst schwierig. Bislang sind zwar zahlreiche Beispiele für Markov-Ketten mit Cut-Off-Effekt bekannt (siehe hierzu [2], [10], [20], [21] und [24]), jedoch ist keine allgemeine Theorie verfügbar.

In Kapitel 1 werden wir zunächst vom Zustandsraum \mathcal{S}_n abstrahieren und Random Walks auf beliebigen endlichen Gruppen vorstellen, die zugleich Markov-

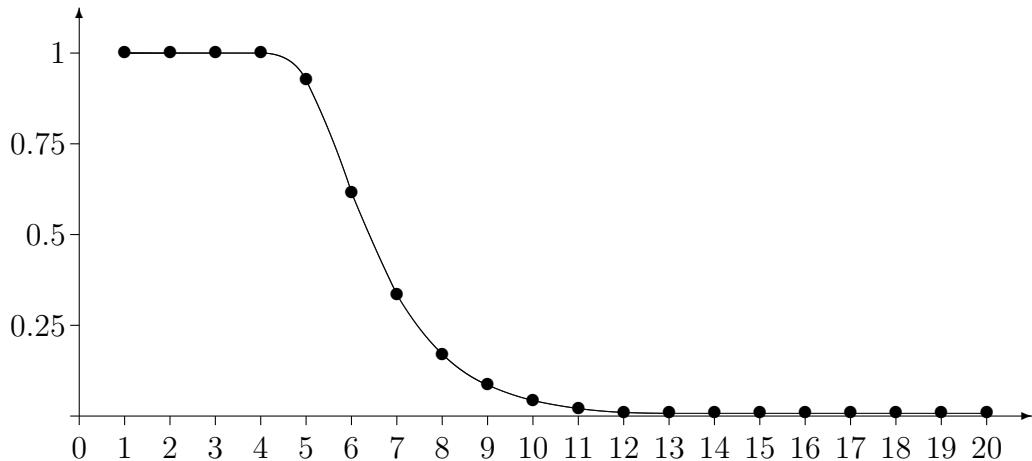


Abbildung 1: $m \mapsto \|P^{X_m} - U_{S_n}\|$ für $n = 52$ und $m = 1, \dots, 20$.

Ketten bilden. Ferner stellen wir ein wahrscheinlichkeitstheoretisches Modell für das sukzessive Kartenmischen vor, das von verschiedenen Autoren zur Modellierung von Mischvorgängen verwendet wird (siehe hierzu [1], [2], [6], [9], [14], [20] und [21]).

Kapitel 2 widmet sich ausschließlich dem Riffle Shuffle und dem Nachweis des Cut-Off-Effekts für den Riffle Shuffle in Form des folgenden Theorems (siehe Theorem 2.6.9 in Abschnitt 2.6).

Theorem. *Wird ein Stapel mit n Karten $m_n = (\lfloor \frac{3}{2} \log_2 n \rfloor + j)$ -mal, $j \in \mathbb{Z}$, $j \geq -\lfloor \frac{3}{2} \log_2 n \rfloor$, nach dem Riffle Shuffle gemischt, so gilt*

$$\|P^{X_{m_n}} - U_{S_n}\| \xrightarrow[n \rightarrow \infty]{} 1 - 2\Phi\left(-\frac{2^{-j}}{4\sqrt{3}}\right),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung sei.

In Kapitel 2 werden wir zunächst den Riffle Shuffle auf drei verschiedene Arten wahrscheinlichkeitstheoretisch modellieren. Diese Modelle führen schließlich zur selben Abbildung durch eine Markov-Kette bzw. einen Random Walk $X = (X_m)_{m \geq 0}$. Sie ermöglichen nicht nur die explizite Angabe der Verteilungen P^{X_m} , $m \geq 0$, sondern auch die der m -Schritt-Übergangswahrscheinlichkeiten der

Markov-Kette X für $m \geq 1$. Hierbei und im weiteren Vorgehen erweisen sich die *Eulerschen Zahlen* und ihre Verbindungen zur Kombinatorik und Wahrscheinlichkeitstheorie als zentrale Hilfsmittel (siehe hierzu [7], [8], [23] und [26]). Abschließend werden wir den Cut-Off-Effekt mathematisch präzisieren und die asymptotische Aussage des obigen Theorems herleiten.

Ich bedanke mich bei Herrn Professor Dr. G. Alsmeyer für die Vergabe dieser Diplomarbeit und die umfassende Betreuung. Er hat mein Interesse für das Thema geweckt und mir mit wertvollen Hinweisen über manche Hürde hinweggeholfen.

Kapitel 1

Random Walks auf Gruppen

Dieses Kapitel dient der Einführung in die Theorie der Random Walks auf Gruppen, die für die wahrscheinlichkeitstheoretische Modellierung von Mischvorgängen notwendig ist. Es werden ausschließlich endliche Gruppen betrachtet. Die Ergebnisse entstammen in weiten Teilen, sofern nicht anders angegeben, [1], [9], [20] und [21]. Begriffe aus der Gruppentheorie entnehmen wir [15]. Für die Theorie der Markov-Ketten wird auf [4] verwiesen.

1.1 Random Walks auf endlichen Gruppen

Sei G eine endliche Gruppe mit Verknüpfung „ \circ “ und neutralem Element id . Die *Ordnung* von G ist die Anzahl $|G|$ der Elemente von G . Sei $T \subset G$ und $T^- \stackrel{\text{def}}{=} \{x^{-1} \mid x \in T\}$. T heißt *Erzeuger von G* , falls

$$G = \bigcup_{n \in \mathbb{N}_0} \{x_1 \circ \dots \circ x_n \mid x_i \in T \cup T^-, i = 1, \dots, n\} \stackrel{\text{def}}{=} \langle T \rangle.$$

Für $n \in \mathbb{N}$ definieren wir $T^{\circ(n)} \stackrel{\text{def}}{=} \{x_1 \circ \dots \circ x_n \mid x_i \in T, i = 1, \dots, n\}$. Für eine endliche Gruppe G ist $G = \langle T \rangle$ wegen $\underbrace{x \circ \dots \circ x}_{|G|\text{-mal}} = \text{id}$ für alle $x \in G$ äquivalent zu

$$G = \bigcup_{n \in \mathbb{N}} T^{\circ(n)}.$$

Gegeben eine Teilmenge $T \subset G$ ist $\langle T \rangle$ eine Untergruppe von G und zwar die kleinste Untergruppe die T umfasst.

Für $x, y \in G$ schreiben wir auch kurz xy in der multiplikativen Schreibweise anstatt $x \circ y$. Das neutrale Element $\text{id} \in G$ nennen wir auch *Einselement*. Eine Untergruppe $N \subset G$ heißt *Normalteiler von G* , wenn $xN = Nx$ für alle $x \in G$ gilt. Die Mengen xN und Nx heißen *Links- bzw. Rechtsnebenklassen von N bezüglich x* . Sei H eine weitere Gruppe. Eine Abbildung $\varphi : G \longrightarrow H$ heißt *Gruppenhomomorphismus*, falls $\varphi(xy) = \varphi(x)\varphi(y)$ für alle $x, y \in G$ gilt.

Ist ein Wahrscheinlichkeitsmaß Q auf $(G, \mathfrak{P}(G))$ gegeben, so definieren wir den *Träger von Q* als

$$\text{supp}(Q) \stackrel{\text{def}}{=} \{x \in G \mid Q(\{x\}) > 0\}$$

und einen stochastischen Kern $\mathbb{K}_Q : G \times \mathfrak{P}(G) \longrightarrow [0, 1]$ durch

$$\mathbb{K}_Q(x, \{y\}) \stackrel{\text{def}}{=} Q(\{x^{-1}y\}) \text{ für } x, y \in G.$$

Für $x \in G$ sei δ_x die Dirac-Verteilung in x auf $(G, \mathfrak{P}(G))$.

1.1.1 Definition. Seien λ und Q Wahrscheinlichkeitsmaße auf $(G, \mathfrak{P}(G))$ und $(Y_m)_{m \geq 0}$ eine Folge G -wertiger, stochastisch unabhängiger Zufallsvariablen auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit $P^{Y_0} = \lambda$ und $P^{Y_m} = Q$ für alle $m \geq 1$. Die Folge $(X_m)_{m \geq 0}$ mit $X_m \stackrel{\text{def}}{=} Y_0 \circ \dots \circ Y_m$ für $m \geq 0$ heißt (Q, λ) -Random Walk auf G oder, falls $\lambda = \delta_{\text{id}}$, kurz Q -Random Walk auf G .

Dann ist $X = (X_m)_{m \geq 0}$ eine zeitlich homogene endliche Markov-Kette mit Zustandsraum G , Anfangsverteilung λ und Übergangskern \mathbb{K}_Q . Sind zwei Wahrscheinlichkeitsmaße Q_1 und Q_2 auf $(G, \mathfrak{P}(G))$ gegeben, so definieren wir die *Faltung* $Q_1 * Q_2$ von Q_1 und Q_2 durch

$$Q_1 * Q_2(\{x\}) \stackrel{\text{def}}{=} \sum_{y \in G} Q_1(\{y\})Q_2(\{y^{-1}x\}), \quad x \in G. \quad (1.1.1)$$

1.1.2 Bemerkung. Sei $X = (X_m)_{m \geq 0}$ ein (Q, λ) -Random Walk auf G . Dann gilt für alle $m \geq 1$

$$P^{X_m} = \lambda * \underbrace{Q * \dots * Q}_{m\text{-mal}} = \lambda * Q^{*(m)}, \quad (1.1.2)$$

wobei $Q^{*(1)} \stackrel{\text{def}}{=} Q$ und $Q^{*(k)} \stackrel{\text{def}}{=} Q * Q^{*(k-1)}$ für $k \geq 2$.

Im Fall $\lambda = \delta_{\text{id}}$ gilt für alle $m \geq 1$

$$P^{X_m} = Q^{*(m)}. \quad (1.1.3)$$

Beweis. Sei $x \in G$ beliebig. Dann gilt wegen der stochastischen Unabhängigkeit der Y_i , $i \geq 0$,

$$\begin{aligned} P(X_1 = x) &= P(Y_0 Y_1 = x) \\ &= P\left(\bigcup_{y \in G} \{Y_0 = y, Y_1 = y^{-1}x\}\right) \\ &= \sum_{y \in G} P(Y_0 = y) P(Y_1 = y^{-1}x) \\ &= \sum_{y \in G} \lambda(\{y\}) Q(\{y^{-1}x\}) = \lambda * Q(\{x\}). \end{aligned}$$

(1.1.2) und (1.1.3) folgen dann per Induktion aus obiger Gleichung unter Beachtung von $\delta_{\text{id}} * Q = Q$. \square

Aus der Definition des Übergangskerns \mathbb{K}_Q und (1.1.2) erhalten wir für alle $m \geq 1$

$$P^{X_m} = \lambda \underbrace{(\mathbb{K}_Q \circ \dots \circ \mathbb{K}_Q)}_{m\text{-mal}} \stackrel{\text{def}}{=} \lambda \mathbb{K}_Q^{(m)},$$

wobei $\lambda \mathbb{K}_Q(\cdot) \stackrel{\text{def}}{=} \int_G \mathbb{K}_Q(x, \cdot) \lambda(dx)$ und „ \circ “ hier die übliche Hintereinanderschaltung von Kernen bezeichne. Ferner gilt für alle $k, m \geq 0$

$$P^{X_{m+k}|X_m=\cdot} = \mathbb{K}_Q^{(k)} P^{X_m}\text{-f.s.},$$

wobei $\mathbb{K}_Q^{(0)}(x, \cdot) \stackrel{\text{def}}{=} \delta_x(\cdot)$. Die k -Schritt-Übergangswahrscheinlichkeiten seien für $k \in \mathbb{N}$ definiert durch

$$p_{x,y}^{(k)} \stackrel{\text{def}}{=} \mathbb{K}_Q^{(k)}(x, \{y\}), \quad x, y \in G.$$

Dann gilt für alle $k \in \mathbb{N}$ und $x, y \in G$

$$\begin{aligned} p_{x,y}^{(k)} &= \mathbb{K}_Q^{(k)}(x, \{y\}) \\ &= Q^{*(k)}(\{x^{-1}y\}) \end{aligned}$$

$$\begin{aligned}
&= Q^{*(k)}(\{\text{id } x^{-1}y\}) \\
&= \mathbb{K}_Q^{(k)}(\text{id}, \{x^{-1}y\}) \\
&= p_{\text{id}, x^{-1}y}^{(k)} \stackrel{\text{def}}{=} p_{x^{-1}y}^{(k)}. \tag{1.1.4}
\end{aligned}$$

Der nächste Satz stammt von Woess [25] und charakterisiert Irreduzibilität und Aperiodizität (siehe Definition 7.6. und 7.9. in [4]) der Markov-Kette $(X_m)_{m \geq 0}$.

1.1.3 Satz. *Sei G eine endliche Gruppe, Q ein Wahrscheinlichkeitsmaß auf $(G, \mathfrak{P}(G))$ und $X = (X_m)_{m \geq 0}$ ein (Q, λ) -Random Walk. Dann gilt*

(a) *X ist irreduzibel genau dann, wenn $\text{supp}(Q)$ ein Erzeuger von G ist.*

(b) *Sei X irreduzibel. Genau dann ist X aperiodisch, wenn $\text{supp}(Q)$ keine Teilmenge einer Nebenklasse eines beliebigen nicht-trivialen Normalteilers von G ist.*

Beweis. zu (a): „ \Rightarrow “ Sei X irreduzibel und $x \in G$. Dann existiert ein $k \in \mathbb{N}$, so dass $p_x^{(k)} > 0$. Wegen

$$\begin{aligned}
p_x^{(k)} &= Q^{*(k)}(\{x\}) \\
&= \sum_{z_1 \in G} Q(\{z_1\}) Q^{*(k-1)}(\{z_1^{-1}x\}) \\
&= \sum_{z_1, \dots, z_{k-1} \in G} Q(\{z_1\}) \cdot \dots \cdot Q(\{z_{k-1}\}) Q(\{z_{k-1}^{-1} \cdot \dots \cdot z_1^{-1}x\}) > 0
\end{aligned}$$

existieren $z_1, \dots, z_{k-1} \in G$ mit $z_1, \dots, z_{k-1}, z_{k-1}^{-1} \cdot \dots \cdot z_1^{-1}x \in \text{supp}(Q)$. Daraus folgt $x \in \langle \text{supp}(Q) \rangle$ und somit $G \subset \langle \text{supp}(Q) \rangle$.

„ \Leftarrow “ Nach (1.1.4) reicht es zu zeigen, dass für jedes $x \in G$ ein $k \in \mathbb{N}$ existiert mit $p_x^{(k)} > 0$. Sei $x \in G$ und $y \in \text{supp}(Q)$ beliebig. Wegen $\langle \text{supp}(Q) \rangle = G$ existiert ein $k \in \mathbb{N}$ und $y_1, \dots, y_{k-1} \in \text{supp}(Q)$ mit $y_1 \cdot \dots \cdot y_{k-1} = xy^{-1}$. Dann gilt $y = y_{k-1}^{-1} \cdot \dots \cdot y_1^{-1}x$ und

$$\begin{aligned}
p_x^{(k)} &= \sum_{z_1, \dots, z_{k-1} \in G} Q(\{z_1\}) \cdot \dots \cdot Q(\{z_{k-1}\}) Q(\{z_{k-1}^{-1} \cdot \dots \cdot z_1^{-1}x\}) \\
&\geq Q(\{y_1\}) \cdot \dots \cdot Q(\{y_{k-1}\}) Q(\{y_{k-1}^{-1} \cdot \dots \cdot y_1^{-1}x\}) \\
&= Q(\{y\}) \prod_{i=1}^{k-1} Q(\{y_i\}) > 0.
\end{aligned}$$

zu (b): Für den Beweis verweisen wir auf [25]. \square

Das nächste Lemma zeigt den Zusammenhang zwischen dem Träger der m -fachen Faltung von Q und der vom Träger von Q erzeugten Untergruppe.

1.1.4 Lemma. *Sei G eine endliche Gruppe und Q ein Wahrscheinlichkeitsmaß auf $(G, \mathfrak{P}(G))$. Dann gilt für alle $m \geq 0$*

$$\text{supp}(Q^{*(m)}) = (\text{supp}(Q))^{\circ(m)} \quad (1.1.5)$$

und somit

$$\langle \text{supp}(Q) \rangle = \bigcup_{m \geq 0} (\text{supp}(Q))^{\circ(m)} = \bigcup_{m \geq 0} \text{supp}(Q^{*(m)}). \quad (1.1.6)$$

Beweis. (1.1.6) folgt direkt aus (1.1.5), wir zeigen daher (1.1.5). Da für $m = 0, 1$ nichts zu zeigen ist, sei $m \geq 2$.

„ \subset “ Sei $x \in \text{supp}(Q^{*(m)})$ beliebig. Wegen

$$Q^{*(m)}(\{x\}) = \sum_{z_1, \dots, z_{m-1} \in G} Q(\{z_1\}) \cdot \dots \cdot Q(\{z_{m-1}\}) Q(\{z_{m-1}^{-1} \cdot \dots \cdot z_1^{-1} x\}) > 0$$

existieren $z_1, \dots, z_{m-1} \in G$ mit $z_1, \dots, z_{m-1}, z_{m-1}^{-1} \cdot \dots \cdot z_1^{-1} x \in \text{supp}(Q)$. Vermöge $(\text{supp}(Q))^{\circ(m)} = \{y_1 \cdot \dots \cdot y_m \mid y_1, \dots, y_m \in \text{supp}(Q)\}$ und

$$z_1 \cdot \dots \cdot z_{m-1} (z_{m-1}^{-1} \cdot \dots \cdot z_1^{-1} x) = x$$

erhalten wir $x \in (\text{supp}(Q))^{\circ(m)}$.

„ \supset “ Sei $x \in (\text{supp}(Q))^{\circ(m)}$ beliebig und $y_1, \dots, y_m \in \text{supp}(Q)$ mit $x = y_1 \cdot \dots \cdot y_m$. Dann gilt $y_{m-1}^{-1} \cdot \dots \cdot y_1^{-1} x = y_m$ und somit

$$\begin{aligned} Q^{*(m)}(\{x\}) &= \sum_{z_1, \dots, z_{m-1} \in G} Q(\{z_1\}) \cdot \dots \cdot Q(\{z_{m-1}\}) Q(\{z_{m-1}^{-1} \cdot \dots \cdot z_1^{-1} x\}) \\ &\geq Q(\{y_1\}) \cdot \dots \cdot Q(\{y_{m-1}\}) Q(\{y_{m-1}^{-1} \cdot \dots \cdot y_1^{-1} x\}) \\ &= \prod_{i=1}^m Q(\{y_i\}) > 0. \end{aligned}$$

Dies zeigt $x \in \text{supp}(Q^{*(m)})$. \square

Eine stationäre Verteilung der Markov-Kette X ist die Gleichverteilung $U_G \stackrel{\text{def}}{=} \frac{1}{|G|} \sum_{x \in G} \delta_x$ auf $(G, \mathfrak{P}(G))$, denn für alle $x \in G$ gilt

$$U_G * Q(\{x\}) = \sum_{y \in G} U_G(\{y\}) Q(\{y^{-1}x\}) = \frac{1}{|G|}.$$

Zusammen mit der Endlichkeit von G und Korollar 10.7. in [4] erhalten wir unmittelbar den folgenden Satz.

1.1.5 Satz. *Sei G eine endliche Gruppe, Q ein Wahrscheinlichkeitsmaß auf $(G, \mathfrak{P}(G))$, dessen Träger $\text{supp}(Q)$ die Gruppe G erzeugt, und $X = (X_m)_{m \geq 0}$ eine Markov-Kette mit Übergangskern \mathbb{K}_Q und beliebiger Anfangsverteilung. Dann ist X positiv rekurrent und besitzt die eindeutig bestimmte stationäre Verteilung U_G .*

Die *Totalvariation* oder der *Variationsabstand* zweier Maße Q_1 und Q_2 auf $(G, \mathfrak{P}(G))$ ist definiert als

$$\|Q_1 - Q_2\| \stackrel{\text{def}}{=} \sup_{A \subset G} |Q_1(A) - Q_2(A)| = \frac{1}{2} \sum_{x \in G} |Q_1(\{x\}) - Q_2(\{x\})|.$$

Ist ein (Q, λ) -Random Walk X auf G irreduzibel und aperiodisch, so konvergiert X nach Satz 1.1.5 und dem Ergodensatz (siehe Satz 11.1. in [4]) für jede Anfangsverteilung λ in Totalvariation gegen die Gleichverteilung U_G auf G , das heißt

$$\lim_{m \rightarrow \infty} \|P^{X_m} - U_G\| = \lim_{m \rightarrow \infty} \|\lambda * Q^{*(m)} - U_G\| = 0.$$

1.2 Die Gruppe \mathcal{S}_n der Permutationen

In diesem Abschnitt sei $n \in \mathbb{N}$ und \mathcal{S}_n die Gruppe der Permutationen von $\{1, \dots, n\}$, das heißt

$$\mathcal{S}_n = \{\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid \pi \text{ bijektiv}\}.$$

Eine Permutation $\pi \in \mathcal{S}_n$ wird gewöhnlich in der Form

$$\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}$$

dargestellt. Wir kürzen dies ab, indem wir nur die untere Zeile übernehmen und diese in eckige Klammern setzen: $\pi = [\pi(1), \dots, \pi(n)]$. Wir entnehmen die folgenden Definitionen und Aussagen über die Struktur von \mathcal{S}_n aus [15].

1.2.1 Definition. Eine Permutation π heißt *r-Zyklus*, wenn es paarweise verschiedene Zahlen $i_1, \dots, i_r \in \mathbb{N}_{\leq n} \stackrel{\text{def}}{=} \{1, \dots, n\}$ gibt, $r \geq 2$, mit

- (i) $\pi(i_j) = i_{j+1}$, $j = 1, \dots, r-1$, $\pi(i_r) = i_1$ und
- (ii) $\pi(i) = i$ für $i \neq i_1, \dots, i_r$.

In dieser Situation verwenden wir für π die Schreibweise $\pi = [i_1, \dots, i_r]_Z$.

1.2.2 Definition. Ein 2-Zyklus $\pi = [i, j]_Z \in \mathcal{S}_n$ heißt *Transposition*.

Eine Transposition $\pi = [i, j]_Z$ vertauscht also gerade die Zahlen i und j miteinander und lässt alle übrigen Zahlen unverändert. Wir bezeichnen die Menge aller Transpositionen in \mathcal{S}_n mit \mathcal{T}_n , das heißt

$$\mathcal{T}_n \stackrel{\text{def}}{=} \{[i, j]_Z \mid i, j \in \mathbb{N}_{\leq n}, i \neq j\}.$$

Die Menge der Transpositionen erzeugt \mathcal{S}_n . Es gilt also

$$\langle \mathcal{T}_n \rangle = \bigcup_{m \geq 0} \mathcal{T}_n^{\circ(m)} = \mathcal{S}_n. \quad (1.2.1)$$

Nach (1.2.1) ist jede Permutation das endliche Produkt von Transpositionen. Wir nennen eine Permutation *gerade* bzw. *ungerade*, falls sie sich als das Produkt einer geraden bzw. ungeraden Anzahl von Transpositionen darstellen lässt. Eine Permutation kann nicht gerade und ungerade zugleich sein. Wir bezeichnen die Menge aller geraden Permutationen mit \mathcal{A}_n .

Die Signumsfunktion $\text{sgn} : \mathcal{S}_n \longrightarrow \{-1, 1\}$, definiert durch

$$\text{sgn}(\pi) = (-1)^s,$$

falls $\pi = \sigma_1 \cdot \dots \cdot \sigma_s$ für $\sigma_1, \dots, \sigma_s \in \mathcal{T}_n$, ist ein Gruppenhomomorphismus. Für die Menge \mathcal{A}_n aller geraden Permutationen gilt

$$\mathcal{A}_n = \{\pi \in \mathcal{S}_n \mid \text{sgn}(\pi) = 1\}.$$

Wir erinnern noch einmal an die Definition des Normalteilers: Eine Untergruppe $N \subset \mathcal{S}_n$ heißt Normalteiler von \mathcal{S}_n , wenn $xN = Nx$ für alle $x \in \mathcal{S}_n$ gilt. Per Definition sind also $\{\text{id}\}$ und \mathcal{S}_n selbst stets Normalteiler von \mathcal{S}_n . $\{\text{id}\}$ und \mathcal{S}_n werden als *triviale Normalteiler* von \mathcal{S}_n bezeichnet. Die folgenden Ausführungen geben Auskunft über die weitere Normalteilerstruktur von \mathcal{S}_n für $n \in \mathbb{N}$.

1.2.3 Satz. *Für $n \geq 5$ ist \mathcal{A}_n der einzige nicht-triviale Normalteiler von \mathcal{S}_n .*

Die *Kleinsche Vierergruppe* \mathcal{V}_4 ist die Menge aller *Doppeltranspositionen* in \mathcal{S}_4 , das heißt

$$\mathcal{V}_4 = \{\text{id}, [1, 2]_Z[3, 4]_Z, [1, 3]_Z[2, 4]_Z, [1, 4]_Z[2, 3]_Z\}.$$

Die Kleinsche Vierergruppe ist ein nicht-trivialer Normalteiler von \mathcal{S}_4 . Die Normalteiler von \mathcal{S}_n sind für $n \in \mathbb{N}$ in der folgenden Übersicht aufgeführt:

$$\begin{aligned} n = 1 : \quad & \{\text{id}\} = \mathcal{S}_1, \\ n = 2 : \quad & \{\text{id}\} \subsetneq \mathcal{A}_2 = \mathcal{S}_2, \\ n = 3 : \quad & \{\text{id}\} \subsetneq \mathcal{A}_3 \subsetneq \mathcal{S}_3, \\ n = 4 : \quad & \{\text{id}\} \subsetneq \mathcal{V}_4 \subsetneq \mathcal{A}_4 \subsetneq \mathcal{S}_4, \\ n \geq 5 : \quad & \{\text{id}\} \subsetneq \mathcal{A}_n \subsetneq \mathcal{S}_n. \end{aligned}$$

1.3 Beschreibung eines Mischvorgangs

In diesem Abschnitt führen wir den Begriff der Mischmethode und ein wahrscheinlichkeitstheoretisches Modell für das sukzessive Mischen von Karten ein.

Wir beginnen unsere Betrachtungen mit einem Stapel von n Karten, $n \in \mathbb{N}$. Die Karten seien von 1 bis n nummeriert und nach aufsteigenden Werten von links nach rechts vor uns ausgelegt.

Wir identifizieren im Folgenden die Reihenfolge der Karten mit einem n -Tupel $x = (x_1, \dots, x_n) \in \mathbb{N}_{\leq n, \neq}^n$, wobei

$$\mathbb{N}_{\leq n, \neq}^n \stackrel{\text{def}}{=} \{(y_1, \dots, y_n) \in \{1, \dots, n\}^n \mid y_i \neq y_j \text{ für } i \neq j\}.$$

Dabei bedeutet $x_i = j$, dass sich Karte j an i -ter Stelle befindet. Ein *ungemischter Stapel* wird also durch das Tupel $(1, \dots, n)$ beschrieben. Führen wir nun einen *Mischvorgang* durch, das heißt, ändern wir die Reihenfolge der Karten, so können wir dies durch eine Permutation

$$\pi \in \mathcal{S}_n = \{\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid \sigma \text{ bijektiv}\}$$

beschreiben. Dabei bedeute $\pi(i) = j$, dass die Karte, die sich nach dem Mischvorgang an Position i befindet, vor dem Mischvorgang an Position j lag.

Führen wir $m \in \mathbb{N}$ Mischvorgänge nacheinander durch, gegeben durch die Permutationen π_1, \dots, π_m , so ist dies äquivalent zu einem einzelnen Mischvorgang gegeben durch die Permutation $\pi_1 \circ \dots \circ \pi_m$. Beginnen wir mit einem ungemischten Stapel, so gilt für die Reihenfolge $x^{(k)}$ der Karten nach dem k -ten Mischvorgang:

$$x^{(k)} = (\pi_1 \circ \dots \circ \pi_k(1), \dots, \pi_1 \circ \dots \circ \pi_k(n)), k = 1, \dots, m.$$

Das folgende Beispiel 1.3.1 verdeutlicht, dass die Unterscheidung zwischen den zu den einzelnen Mischvorgängen gehörigen Kartenreihenfolgen und Permutationen notwendig ist. Soll ein Kartenstapel m -mal sukzessive gemäß den Permutationen π_1, \dots, π_m gemischt werden, so führt, wie oben beschrieben, einmaliges Mischen gemäß der Permutation $\pi_1 \circ \dots \circ \pi_m$ zu derselben Reihenfolge. Entgegen der Intuition ist also $\pi_1 \circ \dots \circ \pi_m$ die „richtige“ Permutation und nicht etwa $\pi_m \circ \dots \circ \pi_1$. Dies liegt an der von uns gewählten Art, einen Mischvorgang durch eine Permutation zu beschreiben: Eine einzelne Permutation beschreibt nicht die Reihenfolge, in der die Karten nach dem Mischen vor uns liegen, sondern nur deren Reihenfolgeänderung und zwar unabhängig von den speziellen Kartenwerten. Die Permutation $\pi_1 \circ \dots \circ \pi_m$ beschreibt daher die sukzessiven Reihenfolgeänderungen beginnend beim letzten Mischvorgang: Für $i = 1, \dots, n$ gibt zunächst $\pi_m(i)$ die Postion an, an der die Karte an Position i vor dem Mischen gemäß π_m lag. Soll nun die gesamte Reihenfolgeänderung nach m -maligem Mischen angegeben werden, so müssen nacheinander π_{m-1} auf π_m , π_{m-2} auf $\pi_{m-1} \circ \pi_m$ usw. bis hin zu π_1 auf $\pi_2 \circ \dots \circ \pi_m$ angewendet werden. Die Reihenfolgeänderungen müssen also beginnend beim letzten Mischvorgang bis zum ersten Mischvorgang zurückverfolgt werden. Daraus resultiert dann die Permutation $\pi_1 \circ \dots \circ \pi_m$. Auch diesen Sachverhalt verdeutlicht das folgende Beispiel.

1.3.1 Beispiel. Wir beginnen mit einem ungemischten Stapel. Ändert sich die Reihenfolge der Karten durch einen ersten Mischvorgang von $(1, \dots, n)$ in $(n, 1, 2, \dots, n - 1)$, so ist $\pi_1 = [n, 1, 2, \dots, n - 1] = [1, n, n - 1, \dots, 3, 2]_Z$ die zugehörige Permutation. In einem zweiten Mischvorgang werde die Reihenfolge gemäß der Permutation $\pi_2 = [3, 2, 1, 4, \dots, n] = [1, 3]_Z$ geändert. Die Reihenfolge der Karten nach dem zweiten Mischvorgang lautet

$$x^{(2)} = (\pi_1 \circ \pi_2(1), \dots, \pi_1 \circ \pi_2(n)) = (2, 1, n, 3, \dots, n - 1).$$

Für die Reihenfolge der Karten nach dem zweiten Mischvorgang gilt also

$$\begin{aligned} x^{(2)} &\neq (\pi_2(1), \dots, \pi_2(n)) \quad \text{und} \\ x^{(2)} &\neq (\pi_2 \circ \pi_1(1), \dots, \pi_2 \circ \pi_1(n)). \end{aligned}$$

Die Wahl der Permutation, nach der wir einen Kartenstapel bei einem einzelnen Mischvorgang mischen, ist zufallsabhängig:

1.3.2 Definition. Eine *Mischmethode* Q ist eine Wahrscheinlichkeitsverteilung Q auf $(\mathcal{S}_n, \mathfrak{P}(\mathcal{S}_n))$.

Ein wahrscheinlichkeitstheoretisches Modell für das sukzessive Mischen eines Kartenstapels nach einer festgelegten Mischmethode Q ist der Q -Random Walk auf \mathcal{S}_n (siehe Definition 1.1.1): Wollen wir einen Stapel von n Karten m -mal nach der Mischmethode Q mischen, so wählen wir zunächst eine Realisation (π_1, \dots, π_m) der Zufallsvariable (Y_1, \dots, Y_m) . Als Ergebnis des Mischvorgangs erhalten wir die Realisation $\pi_1 \circ \dots \circ \pi_m$ der Zufallsvariable X_m . Wir können dabei sowohl bei Vorliegen eines ungemischten, als auch bei Vorliegen eines gemischten Staps vor Beginn des Mischvorgangs, etwa gemäß der Permutation $\pi \in \mathcal{S}_n$, das Dirac-Maß in der Identität δ_{id} als Anfangsverteilung wählen, denn für alle $m \in \mathbb{N}$ gilt

$$\begin{aligned} \|\delta_\pi * Q^{*(m)} - U_{\mathcal{S}_n}\| &= \frac{1}{2} \sum_{\sigma \in \mathcal{S}_n} \left| \delta_\pi * Q^{*(m)}(\{\sigma\}) - \frac{1}{n!} \right| \\ &= \frac{1}{2} \sum_{\sigma \in \mathcal{S}_n} \left| \sum_{\tau \in \mathcal{S}_n} \delta_\pi(\{\tau\}) Q^{*(m)}(\{\tau^{-1}\sigma\}) - \frac{1}{n!} \right| \\ &= \frac{1}{2} \sum_{\sigma \in \mathcal{S}_n} \left| Q^{*(m)}(\{\pi^{-1}\sigma\}) - \frac{1}{n!} \right| \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{\sigma \in S_n} \left| Q^{*(m)}(\{\sigma\}) - \frac{1}{n!} \right| \\
&= \|Q^{*(m)} - U_{S_n}\| \\
&= \|\delta_{id} * Q^{*(m)} - U_{S_n}\|.
\end{aligned}$$

Ferner führt eine etwaige mit der Wahl von δ_{id} als Anfangsverteilung verbundene Umnummerierung der Karten nicht zu einem Informationsverlust, da die zu einem Mischvorgang gehörige Permutation nicht die tatsächliche Reihenfolge der Karten, sondern nur deren Reihenfolgeänderung beschreibt.

Kapitel 2

Der Riffle Shuffle

In diesem Kapitel stellen wir den Riffle Shuffle und wahrscheinlichkeitstheoretische Modellierungen für die vom Riffle Shuffle auf \mathcal{S}_n induzierte Mischmethode vor. Mit Hilfe dieser Modelle wird es möglich sein, den Cut-Off-Effekt für den Riffle Shuffle nachzuweisen. Die Resultate in diesem Kapitel orientieren sich, sofern nicht anders angegeben, an [6].

2.1 Modelle für den Riffle Shuffle

Der *Riffle Shuffle* ist eine der meist verwendeten Methoden, einen Kartenstapel zu mischen. Dabei wird ein ungemischter Stapel von n Karten ungefähr halbiert, die beiden Hälften werden dann sukzessive ineinandergeblättert¹. Ein mathematisches Modell für den Riffle Shuffle wurde 1955 von Gilbert und Shannon [12] und unabhängig 1981 von Reeds [17] vorgestellt: Ein Kartenstapel von n Karten wird in zwei Päckchen A_1 und A_2 geteilt. Dabei sei die Wahrscheinlichkeit, dass A_1 die ersten k Karten enthält, durch $B(n, 1/2)(\{k\}) = \binom{n}{k}/2^n$, $k = 0, \dots, n$, gegeben. Die beiden Päckchen werden dann so ineinandergeblättert, dass die Wahrscheinlichkeit, dass eine Karte vom ersten bzw. zweiten Päckchen fällt, proportional zur Anzahl der verbliebenen Karten in den Päckchen ist. Das heißt, falls im ersten und

¹„ineinanderblättern“ bedeutet, die beiden Hälften so ineinanderzumischen, dass die Reihenfolge der Karten in den einzelnen Hälften erhalten bleibt.

zweiten Päckchen j_1 bzw. j_2 Karten verbleiben, so ist die Wahrscheinlichkeit, dass die nächste Karte von Päckchen i fällt $j_i/(j_1 + j_2)$, $i = 1, 2$. Nach Gilbert, Shannon und Reeds wird der Riffle Shuffle auch GSR-Shuffle genannt.

Der Riffle Shuffle lässt sich für $a \in \mathbb{N}_{\geq 2} \stackrel{\text{def}}{=} \{n \in \mathbb{N} \mid n \geq 2\}$ analog zu einem a -Riffle Shuffle erweitern, indem der Kartenstapel in a Päckchen A_1, \dots, A_a geteilt wird. Die Anzahl der Karten in den Päckchen sei dabei multinomialverteilt. Dies bedeutet, dass die Wahrscheinlichkeit, dass Päckchen A_1 die ersten k_1 Karten, Päckchen A_2 die nächsten k_2 Karten usw. enthält, für

$$(k_1, \dots, k_a) \in \sum_n^a \stackrel{\text{def}}{=} \left\{ (j_1, \dots, j_a) \in \{0, \dots, n\}^a \mid \sum_{k=1}^a j_k = n \right\}$$

durch

$$M(n, 1/a)(\{(k_1, \dots, k_a)\}) = \binom{n}{k_1, \dots, k_a} \frac{1}{a^n}.$$

gegeben sei. Die a Päckchen werden dann wie folgt ineinandergeblättert: Sind in den einzelnen Päckchen j_1, \dots, j_a Karten verblieben, so fällt die nächste Karte von Päckchen i mit der Wahrscheinlichkeit $j_i/(j_1 + \dots + j_a)$, $i = 1, \dots, a$. Wir bezeichnen einen a -Riffle Shuffle von n Karten auch als *GSR-(a, n)-Shuffle*. Im Folgenden stellen wir zwei Modelle vor, die, wie wir in Satz 2.1.7 sehen werden, die gleiche Mischmethode auf \mathcal{S}_n induzieren wie der GSR-(a, n)-Shuffle. Die verbalen Beschreibungen der Modelle sind bei verschiedenen Autoren zu finden, etwa in [1], [2], [6], [9], [14] und [16]. Sei im Folgenden $a \in \mathbb{N}_{\geq 2}$.

1. Maximum-Entropie-Modell: In diesem Modell gehen wir davon aus, dass alle möglichen Arten, einen Stapel in a Päckchen aufeinanderfolgender Karten zu teilen und diese dann ineinanderzublättern, gleich wahrscheinlich sind. Leere Päckchen sind zugelassen. Jedes nichtleere Päckchen muss aus den Karten $k, \dots, k+l$ für geeignetes $k \in \mathbb{N}_{\leq n}$ und $l \in \mathbb{N}_{0,n-1} \stackrel{\text{def}}{=} \{0, \dots, n-1\}$ bestehen. Dieses Modell gibt also jeder beliebigen Kombination aus Päckcheneinteilung und dem anschließenden Ineinanderblättern dieselbe Wahrscheinlichkeit und maximiert so die Entropie unter allen möglichen Wahrscheinlichkeitsverteilungen auf der Menge aller solchen Kombinationen.

Dieser Sachverhalt lässt sich folgendermaßen modellieren: Wir wählen zunächst ein Tupel $(x_1, \dots, x_n) \in \{0, \dots, a-1\}^n = \mathbb{N}_{0,a-1}^n$ nach der Laplace-Verteilung auf $\mathbb{N}_{0,a-1}^n$. Dann zählen wir alle Nullen, Einsen, Zweien usw. in (x_1, \dots, x_n) . Falls wir j_1 Nullen, j_2 Einsen usw. erhalten haben, bilden wir a Päckchen A_1, \dots, A_a mit den Karten 1 bis j_1 bzw. $j_1 + 1$ bis $j_1 + j_2$ usw. Schließlich verteilen wir die Karten aus Päckchen A_i unter Beibehaltung ihrer Reihenfolge auf die Positionen in (x_1, \dots, x_n) mit $x_k = i - 1$.

Wir verwenden hier und im folgenden Modell n -Tupel $x \in \{0, \dots, a-1\}^n$ und nicht etwa $x \in \{1, \dots, a\}^n$, da dies für die Darstellung unserer folgenden Ergebnisse in Lemma 2.1.8, Lemma 2.1.9, Lemma 2.1.10 und Satz 2.1.11 hilfreich ist.

Sei $U_{\mathbb{N}_{0,a-1}^n}$ die Laplace-Verteilung auf $\mathbb{N}_{0,a-1}^n$ und $x = (x_1, \dots, x_n) \in \mathbb{N}_{0,a-1}^n$. Für $k = 1, \dots, a$ definieren wir $j_{x,k} \stackrel{\text{def}}{=} |\{l \mid x_l = k-1\}|$ und

$$A_{x,k} \stackrel{\text{def}}{=} \left\{ \left(\sum_{i=1}^{k-1} j_{x,i} \right) + 1, \dots, \sum_{i=1}^k j_{x,i} \right\}.$$

Ferner definieren wir zu x eine Permutation $\pi_x \in \mathcal{S}_n$ induktiv durch $\pi_x(1) = \min A_{x,x_1+1}$ und

$$\pi_x(i) = \min \{A_{x,x_i+1} \setminus \{\pi_x(1), \dots, \pi_x(i-1)\}\} \quad (2.1.1)$$

für $i = 2, \dots, n$. Sei $X_{a,n}^M : (\mathbb{N}_{0,a-1}^n, \mathfrak{P}(\mathbb{N}_{0,a-1}^n)) \longrightarrow (\mathcal{S}_n, \mathfrak{P}(\mathcal{S}_n))$ die für $x \in \mathbb{N}_{0,a-1}^n$ durch $X_{a,n}^M(x) \stackrel{\text{def}}{=} \pi_x$ definierte Funktion. Die Definition der Permutation π_x findet sich für den Fall $a = 2$ in äquivalenter Form in Beispiel 4.17 in [1].

Wir können nun die *Maximum-Entropie-Mischmethode* $Q_{a,n}^M$ einführen, indem wir für $B \in \mathfrak{P}(\mathcal{S}_n)$ setzen

$$Q_{a,n}^M(B) \stackrel{\text{def}}{=} U_{\mathbb{N}_{0,a-1}^n}(X_{a,n}^M \in B).$$

Dann gilt für alle $\pi \in \mathcal{S}_n$

$$Q_{a,n}^M(\{\pi\}) = \frac{|\{x \in \{0, \dots, a-1\}^n \mid \pi_x = \pi\}|}{a^n}. \quad (2.1.2)$$

2.1.1 Beispiel. Ein Mischvorgang im Maximum-Entropie-Modell mit $a = 2$ und $n = 10$ zu $x = (0, 1, 1, 0, 1, 0, 1, 0, 0, 1) \in \mathbb{N}_{0,1}^{10} = \{0, 1\}^{10}$ (siehe Abbildung 2.1).

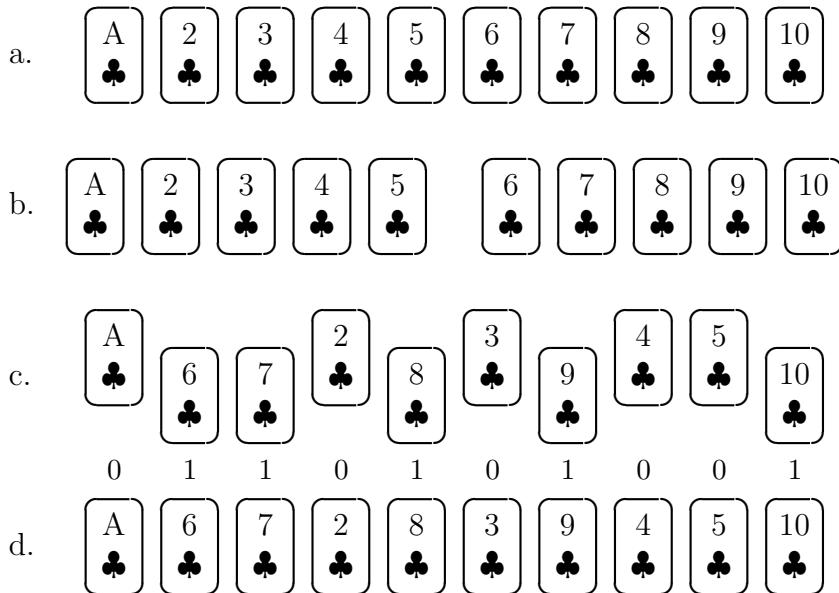


Abbildung 2.1: Mischvorgang im Maximum-Entropie-Modell.

a.: Wir beginnen mit einem geordneten Kartenstapel und identifizieren „Ass“ mit dem Wert 1.

b.: Der Stapel wird in zwei Päckchen geteilt, das erste Päckchen A_1 besteht aus den Karten 1 bis $j_{x,1} = 5$, das zweite Päckchen A_2 aus den Karten $j_{x,1} + 1 = 6$ bis $j_{x,1} + j_{x,2} = 10$.

c.,d.: Die beiden Päckchen werden ineinandergeblättert: Alle Karten aus Päckchen A_1 werden auf die Positionen in x mit $x_i = 0$ verteilt, $i = 1, \dots, 10$, und alle Karten aus Päckchen A_2 auf die Positionen in x mit $x_i = 1$, $i = 1, \dots, 10$. Der Stapel wurde also nach der Permutation $\pi_x = [1, 6, 7, 2, 8, 3, 9, 4, 5, 10]$ gemischt.

2. Inverses Modell: Das folgende Modell beschreibt einen *inversen Riffle Shuffle*. Wir wählen wieder ein Tupel $(x_1, \dots, x_n) \in \mathbb{N}_{0,a-1}^n$ nach der Laplace-Verteilung auf $\mathbb{N}_{0,a-1}^n$. Dann ordnen wir der Karte i den Wert x_i zu und legen alle Karten mit dem Wert 0 unter Beibehaltung ihrer Reihenfolge an den Anfang des Stapels, dahinter platzieren wir alle Karten mit dem Wert 1 usw., bis hin zu allen Karten mit dem Wert $a - 1$. Päckchen A_l besteht beim inversen Riffle Shuffle aus allen Karten, denen der Wert $l - 1$ zugeordnet wird, $l = 1, \dots, n$. Auch in diesem

Modell können leere Päckchen entstehen, sofern nicht alle Werte $0, \dots, a-1$ in x vertreten sind.

Sei $x = (x_1, \dots, x_n) \in \mathbb{N}_{0,a-1}^n$. Wir definieren eine Permutation $\pi_{x^-} \in \mathcal{S}_n$ induktiv durch $\pi_{x^-}(1) = \min \{k \mid x_k = \min\{x_1, \dots, x_n\}\}$ und

$$\pi_{x^-}(i) = \min \{k \mid x_k = l\} \setminus \{\pi_{x^-}(1), \dots, \pi_{x^-}(i-1)\} \quad (2.1.3)$$

für $j_{x,1} + \dots + j_{x,l} < i \leq j_{x,1} + \dots + j_{x,l+1}$, $l = 0, \dots, a-1$. Falls $j_{x,1}$ positiv ist, das heißt ein $k \in \{1, \dots, n\}$ existiert mit $x_k = 0$, so gilt $\pi_{x^-}(1) = \min \{k \mid x_k = 0\}$.

Sei $X_{a,n}^I : (\mathbb{N}_{0,a-1}^n, \mathfrak{P}(\mathbb{N}_{0,a-1}^n)) \longrightarrow (\mathcal{S}_n, \mathfrak{P}(\mathcal{S}_n))$ die für $x \in \mathbb{N}_{0,a-1}^n$ durch

$$X_{a,n}^I(x) \stackrel{\text{def}}{=} \pi_{x^-}$$

definierte Funktion. Wir definieren die Mischmethode $Q_{a,n}^I$ für $B \in \mathfrak{P}(\mathcal{S}_n)$ durch

$$Q_{a,n}^I(B) \stackrel{\text{def}}{=} U_{\mathbb{N}_{0,a-1}^n}(X_{a,n}^I = \pi^{-1}, \pi \in B). \quad (2.1.4)$$

Dann gilt für alle $\pi \in \mathcal{S}_n$

$$Q_{a,n}^I(\{\pi\}) = \frac{|\{x \in \{0, \dots, a-1\}^n \mid \pi_{x^-} = \pi^{-1}\}|}{a^n}. \quad (2.1.5)$$

2.1.2 Beispiel. Ein Mischvorgang im inversen Modell: Wie in Beispiel 2.1.1 seien $a = 2$, $n = 10$ und $x = (0, 1, 1, 0, 1, 0, 1, 0, 0, 1) \in \mathbb{N}_{0,1}^{10} = \{0, 1\}^{10}$ (siehe Abbildung 2.2).

a.: Wir beginnen mit einem geordneten Kartenstapel und identifizieren „Ass“ mit dem Wert 1.

b.: Karte i wird der Wert x_i zugeordnet, $i = 1, \dots, 10$.

c.,d.: Das erste Päckchen A_1 besteht aus den Karten 1, 4, 6, 8, 9, das zweite Päckchen aus den Karten 2, 3, 5, 7, 10. Alle Karten aus Päckchen A_1 werden an den Anfang des Stapels gelegt, dahinter alle Karten aus Päckchen A_2 . Der Stapel wurde gemäß der Permutation $\pi_{x^-} = [1, 4, 6, 8, 9, 2, 3, 5, 7, 10]$ gemischt.

Um nun den Mischvorgang gemäß (2.1.4) zu erhalten, werden die Schritte a. bis d. in umgekehrter Reihenfolge durchgeführt, das heißt der geordnete Kartenstapel wird gemäß $\pi_{x^-}^{-1}$ gemischt. Der inverse Riffle Shuffle gemäß π_{x^-} invertiert also

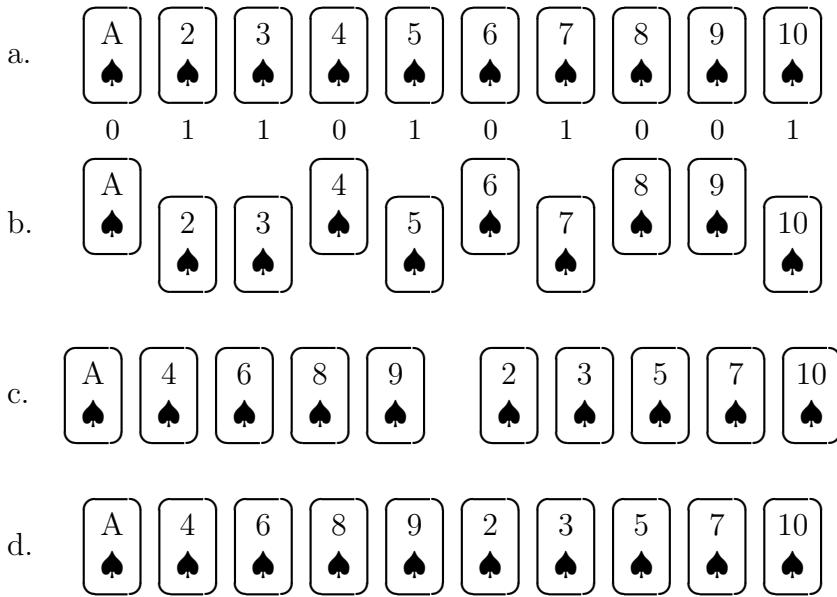


Abbildung 2.2: Mischvorgang im inversen Modell.

gerade den Mischvorgang im Maximum-Entropie-Modell gemäß π_x aus Beispiel 2.1.1. Er ordnet also den gemischten Kartenstapel zurück in die Ausgangslage. Das Mischen kann somit als „inverses Sortieren“ bezeichnet werden. Wir werden diese Beobachtung in Lemma 2.1.6 belegen.

Bayer und Diaconis [6], [9] nennen mit dem *geometrischen Modell* eine weitere Möglichkeit zur Modellierung des GSR- (a, n) -Shuffle. Es lässt sich auch hier zeigen, dass das geometrische Modell äquivalent zu den beiden vorangegangenen ist, in dem Sinne, dass es dieselbe Mischmethode auf \mathcal{S}_n induziert. Wir werden allerdings auf den technischen Beweis verzichten, da uns das geometrische Modell im weiteren Vorgehen keine Vorzüge gegenüber dem Maximum-Entropie-Modell bzw. dem inversen Modell bietet. Wir geben jedoch im Anschluss an die Modellierung eine kurze Beweisskizze an.

3. Geometrisches Modell: Wir wählen n Punkte x_1, \dots, x_n im Einheitsintervall voneinander unabhängig und gleichverteilt. Diese ordnen wir der Größe nach,

so dass $x_1 \leq \dots \leq x_n$. Die Abbildung

$$f_a : (x_1, \dots, x_n) \longmapsto (f_a^{(1)}(x_1), \dots, f_a^{(n)}(x_n)) = (ax_1 \pmod{1}, \dots, ax_n \pmod{1})$$

bildet das Intervall $[0, 1]^n$ auf sich selbst ab. f_a sortiert die Punkte x_1, \dots, x_n um und induziert so ein Wahrscheinlichkeitsmaß auf \mathcal{S}_n (siehe Abbildung 2.3).

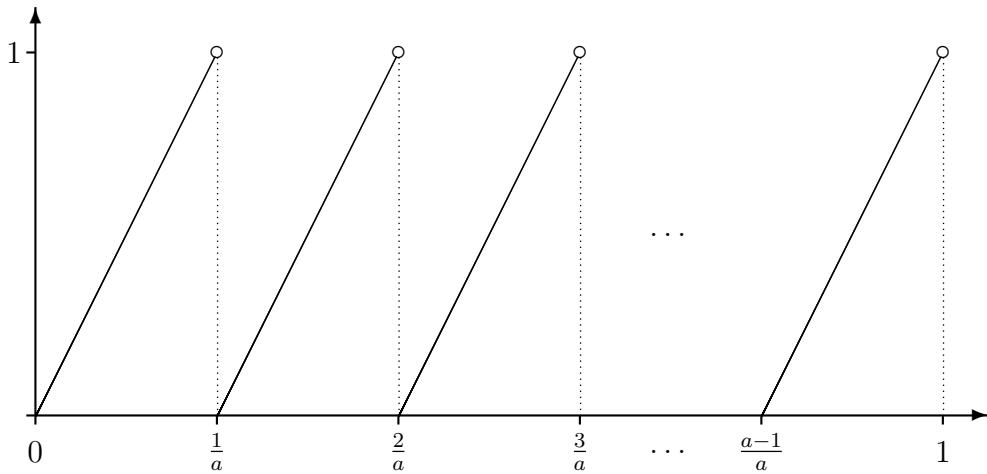


Abbildung 2.3: Die Funktion $f_a^{(1)} : [0, 1] \longrightarrow [0, 1]$ mit $f_a^{(1)}(x) = ax \pmod{1}$.

Seien X_1, \dots, X_n stochastisch unabhängig und identisch $R(0, 1)$ -verteilte Zufallsgrößen auf einem Wahrscheinlichkeitsraum $(\Omega_G, \mathfrak{A}_G, Q_G)$ und $T : [0, 1]^n \longrightarrow [0, 1]^n$ die Ordnungsstatistik mit

$$T(x_1, \dots, x_n) = (x_{(1)}, \dots, x_{(n)}) \text{ für } (x_1, \dots, x_n) \in [0, 1]^n.$$

Die Funktion $f_a : [0, 1]^n \longrightarrow [0, 1]^n$ mit $f_a = (f_a^{(1)}, \dots, f_a^{(n)})$ sei definiert durch

$$f_a^{(i)}(x_i) \stackrel{\text{def}}{=} ax_i \pmod{1} = ax_i - \lfloor ax_i \rfloor \quad (2.1.6)$$

für $x_i \in [0, 1]$ und $i = 1, \dots, n$, wobei $\lfloor x \rfloor$ die untere Gauss-Klammer von $x \in \mathbb{R}$ sei. Ferner sei $\mathcal{R} = (\mathcal{R}_1, \dots, \mathcal{R}_n) : [0, 1]^n \longrightarrow \mathbb{N}_{\leq n}^n$ die Rangstatistik mit $\mathcal{R}_j : [0, 1]^n \longrightarrow \mathbb{N}_{\leq n}$,

$$\mathcal{R}_j(x_1, \dots, x_n) = |\{i \mid x_i \leq x_j\}|$$

für $(x_1, \dots, x_n) \in [0, 1]^n$ und $j = 1, \dots, n$.

2.1.3 Bemerkung. Für die in (2.1.6) definierte Funktion gilt

- (a) f_a ist $\mathfrak{B}_{[0,1]}^n$ -messbar.
- (b) $f_a^{(1)}(X_1), \dots, f_a^{(n)}(X_n)$ sind stochastisch unabhängig und identisch $R(0, 1)$ -verteilt.

Beweis. zu (a): Für $i \leq n$ und $s, t \in [0, 1]$ mit $s < t$ gilt

$$f_a^{(i)-1}((s, t]) = \bigcup_{k=0}^{a-1} \left(\frac{s+k}{a}, \frac{t+k}{a} \right] \in \mathfrak{B}_{[0,1]}. \quad (2.1.7)$$

zu (b): Die stochastische Unabhängigkeit folgt aus (a) und der stochastischen Unabhängigkeit von X_1, \dots, X_n . Sei $i \leq n$ und $t \in (0, 1]$. Dann gilt

$$\begin{aligned} Q_G(f_a^{(i)}(X_i) \leq t) &= Q_G^{X_i}(f_a^{(i)} \in [0, t]) \\ &= Q_G^{X_i}\left(\bigcup_{k=0}^{a-1} \left[\frac{k}{a}, \frac{t+k}{a}\right]\right) \\ &= \sum_{k=0}^{a-1} \frac{t}{a} = t. \end{aligned} \quad (2.1.8)$$

Für $t \leq 0$ erhalten wir unmittelbar $Q_G(f_a^{(i)}(X_i) \leq t) = 0$ und für $t > 1$ analog zu (2.1.8)

$$Q_G(f_a^{(i)}(X_i) \leq t) = Q_G^{X_i}(f_a^{(i)} \in [0, t]) = Q_G^{X_i}(f_a^{(i)} \in [0, 1]) = 1.$$

Somit gilt $Q_G^{f_a^{(i)}(X_i)} = R(0, 1)$ für $i = 1, \dots, n$. \square

Wir definieren die Abbildung $X_a^G \stackrel{\text{def}}{=} \mathcal{R} \circ f_a \circ T \circ X$, $X = (X_1, \dots, X_n)$, und das Maß $Q_{a,n}^G : \mathfrak{P}(\mathcal{S}_n) \longrightarrow [0, 1]$ für $B \in \mathfrak{P}(\mathcal{S}_n)$ durch

$$Q_{a,n}^G(B) \stackrel{\text{def}}{=} Q_G(X_a^G \in B). \quad (2.1.9)$$

Damit durch diese Definition ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_n, \mathfrak{P}(\mathcal{S}_n))$ definiert wird, muss X_a^G f.s. \mathcal{S}_n -wertig sein. Wie die folgende Bemerkung zeigt, ist dies tatsächlich der Fall.

2.1.4 Bemerkung. $Q_{a,n}^G$ ist ein Wahrscheinlichkeitsmaß auf $(\mathcal{S}_n, \mathfrak{P}(\mathcal{S}_n))$.

Beweis. Es ist zu zeigen, dass X_a^G Q_G -f. s. \mathcal{S}_n -wertig ist, also

$$Q_{a,n}^G(\mathcal{S}_n) = Q_G(\mathcal{R} \circ f_a \circ T \circ X \in \mathbb{N}_{\leq n, \neq}^n) = 1. \quad (2.1.10)$$

Dies ist genau dann der Fall, wenn $Q_G(f_a \circ T \circ X \in [0, 1]_{\neq}^n) = 1$ gilt. Sei $(x_1, \dots, x_n) \in [0, 1]^n$, dann gilt

$$f_a(x_1, \dots, x_n) \in [0, 1]_{\neq}^n \iff f_a(x_{(1)}, \dots, x_{(n)}) \in [0, 1]_{\neq}^n$$

und somit

$$Q_G(f_a \circ T \circ X \in [0, 1]_{\neq}^n) = Q_G(f_a \circ X \in [0, 1]_{\neq}^n).$$

Da $f_a^{(1)}(X_1), \dots, f_a^{(n)}(X_n)$ nach Bemerkung 2.1.3 stochastisch unabhängig und stetig verteilt sind, ergibt sich $Q_G(f_a \circ X \in [0, 1]_{\neq}^n) = 1$ und daraus (2.1.10). \square

2.1.5 Lemma. Sei $a \in \mathbb{N}$ und $x, y \in [\frac{i-1}{a}, \frac{i}{a})$ für ein $i \in \mathbb{N}$, dann gilt

$$x < y \iff ax \pmod{1} < ay \pmod{1}.$$

Beweis. Seien $x_0, y_0 \in [i-1, i)$ für ein $i \in \mathbb{N}$. Dann gilt wegen $\lfloor x_0 \rfloor = \lfloor y_0 \rfloor$

$$\begin{aligned} x_0 < y_0 &\iff x_0 - y_0 < \lfloor x_0 \rfloor - \lfloor y_0 \rfloor \\ &\iff x_0 - \lfloor x_0 \rfloor < y_0 - \lfloor y_0 \rfloor \\ &\iff x_0 \pmod{1} < y_0 \pmod{1}. \end{aligned}$$

Die Behauptung folgt, indem wir $x_0 = ax$ und $y_0 = ay$ setzen. \square

Wir führen nun die Funktion $\mathcal{J}_a = (J_1, \dots, J_a) : [0, 1]^n \longrightarrow \sum_n^a$ ein, indem wir für $(x_1, \dots, x_n) \in [0, 1]^n$ und $i = 1, \dots, n$ definieren

$$J_i((x_1, \dots, x_n)) \stackrel{\text{def}}{=} |\{j \mid x_j \in [\frac{i-1}{a}, \frac{i}{a})\}|.$$

Mit Lemma 2.1.5 erhalten wir die folgende Interpretation eines durch eine Realisation $x = (x_1, \dots, x_n)$ der Zufallsvariablen $X = (X_1, \dots, X_n)$ gegebenen Mischvorgangs $(\mathcal{R} \circ f_a \circ T)((x_1, \dots, x_n))$ (siehe auch Abbildung 2.3):

Zunächst können wir die Punkte x_1, \dots, x_n und $f_a^{(1)}(x_{(1)}), \dots, f_a^{(n)}(x_{(n)})$ jeweils als paarweise verschieden voraussetzen, da zum einen X_1, \dots, X_n stetig verteilt sind und zum anderen $f_a \circ T \circ X$ fast sicher $[0, 1]^n$ -wertig ist. Zusätzlich nehmen wir noch $x_1, \dots, x_n < 1$ an. Nach Anwendung der Ordnungsstatistik T auf (x_1, \dots, x_n) sortiert die Funktion f_a das Tupel $(x_{(1)}, \dots, x_{(n)})$ um. Nach Lemma 2.1.5 bleibt dabei die Rangfolge aller Punkte, die im selben Intervall $[\frac{i-1}{a}, \frac{i}{a})$ für ein $i \leq a$ liegen, unverändert. Gegeben $i, j \leq a$ mit $i \neq j$ und Punkte

$$x_{(i_1)}, \dots, x_{(i_{r_1})} \in \left[\frac{i-1}{a}, \frac{i}{a} \right) \text{ und } x_{(j_1)}, \dots, x_{(j_{r_2})} \in \left[\frac{j-1}{a}, \frac{j}{a} \right)$$

mit der Eigenschaft

$$x_{(i_1)} < \dots < x_{(i_{r_1})} \text{ und } x_{(j_1)} < \dots < x_{(j_{r_2})},$$

kann durch Anwendung der Funktion f_a ihre Rangfolge geändert werden. Da jedoch die Rangfolge der Punkte innerhalb des Intervalls $[\frac{i-1}{a}, \frac{i}{a})$ bzw. $[\frac{j-1}{a}, \frac{j}{a})$ jeweils durch f_a nach Lemma 2.1.5 nicht geändert wird, „verzahnen“ sich die beiden Ketten

$$f_a^{(i_1)}(x_{(i_1)}), \dots, f_a^{(i_{r_1})}(x_{(i_{r_1})}) \text{ und } f_a^{(j_1)}(x_{(j_1)}), \dots, f_a^{(j_{r_2})}(x_{(j_{r_2})}).$$

Daraus folgt, dass wir die Anzahl der Karten in den Päckchen A_1, \dots, A_a bei dem durch $(\mathcal{R} \circ f_a \circ T)((x_1, \dots, x_n))$ gegebenen Mischvorgang mit der Anzahl der Punkte x_1, \dots, x_n in den einzelnen Intervallen $[0, \frac{1}{a}), [\frac{1}{a}, \frac{2}{a}), \dots, [\frac{a-1}{a}, 1)$ identifizieren können. Die Anzahl der Karten in den Päckchen A_1, \dots, A_a ist also gerade durch $J_1(x), \dots, J_a(x)$ gegeben. Genauer ist Päckchen A_k dann für $k = 1, \dots, a$ gegeben durch

$$A_k = \left\{ \sum_{i=1}^{k-1} J_i(x) + 1, \dots, \sum_{i=1}^k J_i(x) \right\}.$$

Es lässt sich zeigen, dass das geometrische Modell die gleiche Mischmethode auf \mathcal{S}_n erzeugt, wie das Maximum-Entropie-Modell. Hierzu betrachte man für $a \in \mathbb{N}_{\geq 2}$ und $x \in [0, 1]$ die a -adische Entwicklung von x . Eine solche ist gegeben durch eine Folge $(z_k)_{k \geq 1} \in \{0, \dots, a-1\}^{\mathbb{N}}$ mit

$$x = \sum_{k \geq 1} z_k a^{-k}.$$

Diese Darstellung ist im Allgemeinen nicht eindeutig bestimmt, es existieren jedoch für jedes $x \in [0, 1]$ höchstens zwei verschiedene a -adische Entwicklungen. Sei $Y = (Y_{j,k})_{j \leq n, k \geq 1}$ eine Zufallsvariable mit stochastisch unabhängigen und auf $\{0, \dots, a-1\}$ Laplace-verteilten Komponenten. Dann gilt für $j = 1, \dots, n$

$$P^{\sum_{k \geq 1} Y_{j,k} a^{-k}} = R(0, 1).$$

Somit kann das geometrische Modell wie folgt beschrieben werden: Wir wählen zunächst eine Realisation

$$(y_1, \dots, y_n) = ((y_{1,k})_{k \geq 1}, \dots, (y_{n,k})_{k \geq 1})$$

der Zufallsvariablen Y . Im geometrischen Modell entspricht dies der Wahl von n Punkten

$$x_1 = \sum_{k \geq 1} y_{1,k} a^{-k}, \dots, x_n = \sum_{k \geq 1} y_{n,k} a^{-k}.$$

Die Ordnung $x_{(1)}, \dots, x_{(n)}$ der Punkte x_1, \dots, x_n der Größe nach, entspricht der lexikographischen Ordnung $y_{(1)}, \dots, y_{(n)}$ der Folgen y_1, \dots, y_n . Die Anwendung von f_a auf $(x_{(1)}, \dots, x_{(n)})$ findet ihr Analogon in der Anwendung der Shift-Operation σ auf $y_{(1)}, \dots, y_{(n)}$, die für $(z_k)_{k \geq 1} \in \{0, \dots, a-1\}^{\mathbb{N}}$ definiert ist durch

$$\sigma((z_k)_{k \geq 1}) = (z_k)_{k \geq 2}.$$

Ermitteln wir nun die Ränge der Punkte $f_a^{(1)}(x_{(1)}), \dots, f_a^{(n)}(x_{(n)})$, so zeigt sich, dass die Ränge im Wesentlichen von den ersten Komponenten von $\sigma(y_{(1)}), \dots, \sigma(y_{(n)})$ abhängen. Da diese aber vereinfacht gesprochen wiederum modulo Ordnung und Anwendung der Shift-Operation aus einer Realisation von n stochastisch unabhängigen und über $\{0, \dots, a-1\}$ Laplace-verteilten Zufallsgrößen stammen, zeigt sich dann die Äquivalenz von geometrischem Modell und Maximum-Entropie-Modell.

Wir wenden uns nun wieder dem Maximum-Entropie-Modell und dem inversen Modell zu. Das folgende Lemma bestätigt die Beobachtung aus Beispiel 2.1.2, dass der zu $x \in \mathbb{N}_{0,a-1}^n$ gehörige Mischvorgang gemäß π_x aus dem Maximum-Entropie-Modell das Inverse des Mischvorgangs π_{x^-} aus dem inversen Modell ist.

2.1.6 Lemma. Sei $x \in \mathbb{N}_{0,a-1}^n$. Dann gilt $\pi_x^{-1} = \pi_{x^-}$.

Beweis. Sei $x = (x_1, \dots, x_n) \in \mathbb{N}_{0,a-1}^n$ beliebig. Wir nehmen o.B.d.A. $j_{x,i} \neq 0$ für $i = 1, \dots, a$ an. Sonst setzen wir $b \stackrel{\text{def}}{=} |\{x_1, \dots, x_n\} \cap \mathbb{N}_{0,a-1}|$ und definieren $x' = (x'_1, \dots, x'_n) \in \mathbb{N}_{0,b-1}^n$ durch

$$x'_i \stackrel{\text{def}}{=} x_i - |\{k \in \{0, \dots, n\} \mid k < x_i, k \neq x_1, \dots, x_n\}|, \quad i = 1, \dots, n.$$

Dann gilt $j_{x',i} \neq 0$ für $i = 1, \dots, b$. Ferner erhalten wir $A_{x',k} = A_{x,l_k} \neq \emptyset$ für $k = 1, \dots, b$, wobei l_k induktiv definiert ist durch $l_1 \stackrel{\text{def}}{=} \min\{i \mid A_{x,i} \neq \emptyset\}$ und

$$l_k \stackrel{\text{def}}{=} \min\{i > l_{k-1} \mid A_{x,i} \neq \emptyset\}$$

für $k = 2, \dots, b$. Dann gilt $\pi_x = \pi_{x'}$ wegen der Definition (2.1.1) von π_x .

Sei also $j_{x,i} \neq 0$ für $i = 1, \dots, a$. Es existieren $k_0, \dots, k_a \in \mathbb{N}$ mit

$$0 = k_0 < k_1 < \dots < k_{a-1} < k_a = n \quad \text{und}$$

$$A_{x,l+1} = \{k_l + 1, \dots, k_{l+1}\}$$

für $l = 0, \dots, a-1$. Wir definieren $i_1, \dots, i_n \in \mathbb{N}_{\leq n}$ durch

$$\begin{aligned} i_{k_l+1} &\stackrel{\text{def}}{=} \min\{i \mid x_i = l\}, \quad l = 0, \dots, a-1 \quad \text{und} \\ i_j &\stackrel{\text{def}}{=} \min \left\{ \{i \mid x_i = l\} \setminus \{i_{k_l+1}, \dots, i_{j-1}\} \right\} \end{aligned}$$

für $k_l + 1 < j \leq k_{l+1}$ und $l = 0, \dots, a-1$.

Wir zeigen nun mit einer Induktion $\pi_x(i_j) = j$ für $j = 1, \dots, n$. Aus $i_1 = \min\{i \mid x_i = 0\}$ folgt $x_{i_1} = 0$ und somit

$$\begin{aligned} \pi_x(i_1) &= \min \left\{ A_{x,x_{i_1}+1} \setminus \{\pi_x(1), \dots, \pi_x(i_1-1)\} \right\} \\ &= \min \left\{ A_{x,1} \setminus \{\pi_x(1), \dots, \pi_x(i_1-1)\} \right\} = \min A_{x,1} = 1. \end{aligned} \tag{2.1.11}$$

Sei $m \leq n$ beliebig und $\pi_x(i_j) = j$ für $j < m$. Wir zeigen $\pi_x(i_m) = m$. Sei $l \in \{0, \dots, a-1\}$, so dass $k_l + 1 \leq m \leq k_{l+1}$. Falls $m = k_l + 1$, so folgt wegen

$$i_m = \min\{i \mid x_i = l\}$$

analog zu (2.1.11) $\pi_x(i_m) = m$. Gilt andererseits $k_l + 1 < m \leq k_{l+1}$, so folgt aus $i_m = \min \{ \{i \mid x_i = l\} \setminus \{i_{k_l+1}, \dots, i_{m-1}\} \}$ die Beziehung

$$i_m > i_{k_l+1}, \dots, i_{m-1}.$$

Ferner gilt $\pi_x(j) \notin A_{x,l+1}$ für alle $j < i_m$ mit $j \neq i_{k_l+1}, \dots, i_{m-1}$. Dann folgt zusammen mit $x_{i_m} = l$ und der Induktionsvoraussetzung

$$\begin{aligned} \pi_x(i_m) &= \min \{A_{x,x_{i_m}+1} \setminus \{\pi_x(1), \dots, \pi_x(i_m-1)\}\} \\ &= \min \{A_{x,l+1} \setminus \{\pi_x(1), \dots, \pi_x(i_m-1)\}\} \\ &= \min \{A_{x,l+1} \setminus \{\pi_x(i_{k_l+1}), \dots, \pi_x(i_{m-1})\}\} \\ &= \min \{A_{x,l+1} \setminus \{k_l + 1, \dots, m - 1\}\} \\ &= \min \{\{k_l + 1, \dots, k_{l+1}\} \setminus \{k_l + 1, \dots, m - 1\}\} \\ &= m. \end{aligned}$$

Also gilt $\pi_x(i_j) = j$ für $j = 1, \dots, n$. Aus der Definition der i_j folgt unmittelbar

$$i_j = \min \{ \{i \mid x_i = l\} \setminus \{i_1, \dots, i_{j-1}\} \}$$

für $k_l + 1 \leq j \leq k_{l+1}$ und $l = 0, \dots, a - 1$. Wegen $\pi_x^{-1}(j) = i_j$ für $j = 1, \dots, n$ erhalten wir zusammen mit (2.1.3)

$$\pi_x^{-1}(1) = i_1 = \min \{k \mid x_k = 0\} = \pi_{x^-}(1).$$

Daraus folgt wiederum mit (2.1.3) induktiv für $k_l + 1 \leq j \leq k_{l+1}$ und $l = 0, \dots, a - 1$

$$\begin{aligned} \pi_x^{-1}(j) &= i_j \\ &= \min \{ \{k \mid x_k = l\} \setminus \{i_1, \dots, i_{j-1}\} \} \\ &= \min \{ \{k \mid x_k = l\} \setminus \{\pi_x^{-1}(1), \dots, \pi_x^{-1}(j-1)\} \} \\ &= \min \{ \{k \mid x_k = l\} \setminus \{\pi_{x^-}(1), \dots, \pi_{x^-}(j-1)\} \} \\ &= \pi_{x^-}(j) \end{aligned}$$

und somit $\pi_x^{-1} = \pi_{x^-}$. □

Wir kommen nun zu dem bereits angekündigten Resultat, dass das Maximum-Entropie-Modell und das inverse Modell die gleiche Mischmethode wie der GSR- (a, n) -Shuffle erzeugen (siehe Lemma 1, Abschnitt 3 in [6] und Abschnitt 3 in [16]).

2.1.7 Satz. *Das Maximum-Entropie-Modell und das inverse Modell erzeugen die gleiche Mischmethode, das heißt für alle $n \in \mathbb{N}$ gilt*

$$Q_{a,n}^M = Q_{a,n}^I.$$

Ferner beschreiben beide Modelle den GSR- (a, n) -Shuffle.

Beweis. Sei $n \in \mathbb{N}$. $Q_{a,n}^M = Q_{a,n}^I$ folgt unter Benutzung von (2.1.2) und (2.1.5) direkt aus $\pi_x^{-1} = \pi_{x-}$ für alle $x \in \mathbb{N}_{0,a-1}^n$. Es reicht also zu zeigen, dass das Maximum-Entropie-Modell den GSR- (a, n) -Shuffle beschreibt.

Wir teilen den GSR- (a, n) -Shuffle in zwei Schritte auf: Im ersten Schritt werden die Päckchengrößen nach einer $M(n, 1/a)$ -Verteilung gewählt, das heißt die Wahrscheinlichkeit, dass die Päckchen A_1, \dots, A_a j_1, \dots, j_a Karten enthalten, ist für $(j_1, \dots, j_a) \in \sum_n^a$ gegeben durch

$$M(n, 1/a)(\{(j_1, \dots, j_a)\}) = \binom{n}{j_1, \dots, j_a} \frac{1}{a^n}.$$

Im zweiten Schritt blättern wir die Päckchen in der zu Beginn dieses Kapitels auf Seite 14 beschriebenen Weise ineinander.

Gegeben die Päckchengrößen $(j_1, \dots, j_a) \in \sum_n^a$ sind dann alle möglichen Mischvorgänge gleichwahrscheinlich: Die Wahrscheinlichkeit, dass die Karten in einer beliebigen Reihenfolge sukzessive von den Päckchen A_1, \dots, A_a fallen, ist wegen $j_1 + \dots + j_a = n$ gegeben durch

$$\frac{j_1(j_1 - 1) \cdot \dots \cdot 1 \cdot \dots \cdot j_a(j_a - 1) \cdot \dots \cdot 1}{(j_1 + \dots + j_a)(j_1 + \dots + j_a - 1) \cdot \dots \cdot 1} = \frac{j_1! \cdot \dots \cdot j_a!}{n!} = \binom{n}{j_1, \dots, j_a}^{-1}.$$

Die Wahrscheinlichkeit für einen beliebigen GSR- (a, n) -Shuffle mit Päckchengrößen j_1, \dots, j_a im ersten Schritt ist dann gegeben durch

$$\binom{n}{j_1, \dots, j_a} \frac{1}{a^n} \binom{n}{j_1, \dots, j_a}^{-1} = \frac{1}{a^n}.$$

Das heißt alle möglichen GSR- (a, n) -Shuffle sind gleichwahrscheinlich mit Wahrscheinlichkeit $1/a^n$. Im Maximum-Entropie-Modell entsprechen der erste und zweite Schritt gerade der Wahl eines $x \in \mathbb{N}_{0,a-1}^n$, denn $j_{x,1}, \dots, j_{x,a} \in \sum_n^a$ legen die

Päckchengrößen fest und x selbst die Reihenfolge, in der die Karten ineinandergeblättert werden. Die Wahrscheinlichkeit, dass ein GSR- (a, n) -Shuffle in der Permutation $\pi \in \mathcal{S}_n$ resultiert, ist dann gegeben durch die Anzahl aller $x \in \mathbb{N}_{0,a-1}^n$ mit $\pi_x = \pi$, dividiert durch die Anzahl a^n aller möglichen Mischvorgänge. Die gesuchte Wahrscheinlichkeit entspricht also gerade

$$\frac{|\{x \in \mathbb{N}_{0,a-1}^n \mid \pi_x = \pi\}|}{a^n} = Q_{a,n}^M(\{\pi\}),$$

was wir beweisen mussten. \square

Die folgenden Resultate dieses Abschnitts orientieren sich an den Ausführungen in Abschnitt 7.2 in [16]. Sie sind allerdings in der hier formulierten Weise in der verwendeten Literatur nicht zu finden, dies gilt insbesondere für den Beweis von Satz 2.1.11.

Wir bezeichnen nun für ein $x \in \mathbb{N}_{0,a-1}^n$ einen Mischvorgang, der durch die Permutation π_x beschrieben wird, als *GSR- (a, n) -Shuffle* und einen Mischvorgang, der durch π_x beschrieben wird, als *inverser GSR- (a, n) -Shuffle*. Ferner setzen wir $Q_{a,n} \stackrel{\text{def}}{=} Q_{a,n}^M = Q_{a,n}^I$.

Seien $a, b \in \mathbb{N}_{\geq 2}$. Für das weitere Vorgehen definieren wir eine totale Ordnung auf $\mathbb{N}_{0,a-1} \times \mathbb{N}_{0,b-1}$, genannt „lexikographische Ordnung nach der zweiten Komponente“: Für $(x_i, y_i) \in \mathbb{N}_{0,a-1} \times \mathbb{N}_{0,b-1}$, $i = 1, 2$, setzen wir

$$(x_1, y_1) \leq (x_2, y_2) \iff (x_1 \leq x_2, y_1 = y_2) \text{ oder } y_1 < y_2. \quad (2.1.12)$$

2.1.8 Lemma. Seien $a, b \in \mathbb{N}_{\geq 2}$ und $(x_i, y_i) \in \mathbb{N}_{0,a-1} \times \mathbb{N}_{0,b-1}$, $i = 1, 2$. Dann gilt

$$(x_1, y_1) \leq (x_2, y_2) \iff ay_1 + x_1 \leq ay_2 + x_2. \quad (2.1.13)$$

Beweis. „ \Rightarrow “ Falls gilt $x_1 \leq x_2$ und $y_1 = y_2$, so folgt direkt

$$ay_1 + x_1 \leq ay_2 + x_2.$$

Gilt andererseits $y_1 < y_2$, so folgt $x_1 - x_2 \leq a$ wegen $x_1, x_2 \in \{0, \dots, a-1\}$ und somit

$$\begin{aligned} y_1 < y_2 &\Rightarrow ay_1 < ay_2 \\ &\Rightarrow a \leq ay_2 - ay_1 \end{aligned}$$

$$\begin{aligned} &\Rightarrow x_1 - x_2 \leq ay_2 - ay_1 \\ &\Rightarrow ay_1 + x_1 \leq ay_2 + x_2. \end{aligned}$$

„ \Leftarrow “ Im Fall $ay_1 + x_1 = ay_2 + x_2$ folgt wegen der Eindeutigkeit der Division mit Rest $x_1 = x_2$ und $y_1 = y_2$. Es gilt also $(x_1, y_1) = (x_2, y_2)$. Falls andererseits $ay_1 + x_1 < ay_2 + x_2$ ist, so gilt entweder $x_1 < x_2$ und $y_1 = y_2$ oder $y_1 < y_2$, also in beiden Fällen per Definition $(x_1, y_1) \leq (x_2, y_2)$. \square

Für $x \in \mathbb{N}_{0,a-1}^n$, $y \in \mathbb{N}_{0,b-1}^n$ definieren wir

$$x \circ y \stackrel{\text{def}}{=} ay_{\pi_x} + x = (ay_{\pi_x(1)} + x_1, \dots, ay_{\pi_x(n)} + x_n). \quad (2.1.14)$$

Dann gilt entsprechend

$$y \circ x = bx_{\pi_y} + y = (bx_{\pi_y(1)} + y_1, \dots, bx_{\pi_y(n)} + y_n).$$

Den Nutzen dieser Definition, die in ähnlicher Weise in Abschnitt 7.2 in [16] zu finden ist, zeigt das folgende Lemma.

2.1.9 Lemma. Seien $a, b \in \mathbb{N}_{\geq 2}$, $x \in \mathbb{N}_{0,a-1}^n$ und $y \in \mathbb{N}_{0,b-1}^n$. Dann gilt

$$\pi_{x^-} \circ \pi_{y^-} = \pi_{(x,y_{\pi_x})^-} = \pi_{(x \circ y)^-}, \quad (2.1.15)$$

wobei $\pi_{(x,y_{\pi_x})^-}$ entsprechend (2.1.3) gemäß der in (2.1.12) definierten Ordnung bestimmt ist.

Beweis. $\pi_{x^-} \circ \pi_{y^-}$ beschreibt einen inversen GSR- (a, n) -Shuffle gemäß π_{x^-} gefolgt von einem inversen GSR- (b, n) -Shuffle gemäß π_{y^-} : Das heißt zunächst wird Karte i der Wert x_i zugeordnet, $i = 1, \dots, n$. Dann werden alle Karten in die Reihenfolge $(\pi_{x^-}(1), \dots, \pi_{x^-}(n))$ sortiert, das heißt alle Karten mit dem Wert 0 werden unter Beibehaltung ihrer Reihenfolge an den Anfang des Stapels gelegt, gefolgt von allen Karten mit dem Wert 1, bis hin zu allen Karten mit dem Wert $a - 1$. Anschließend wird der Karte an Position $\pi_{x^-}(i)$ der Wert y_i zugeordnet, $i = 1, \dots, n$, und die Karten in der bekannten Weise nach aufsteigenden Werten von 0 bis $b - 1$ in (y_1, \dots, y_n) sortiert. Wir notieren dabei die Werte x_1, \dots, x_n und y_1, \dots, y_n auf den einzelnen Karten (siehe (2.1.16)), indem wir zunächst auf die Karte an Position j

den Wert x_j schreiben und anschließend gemäß π_{x^-} mischen. Dann schreiben wir auf die Karte an Position j den Wert y_j rechts von dem Wert $x_{\pi_{x^-}(j)}$, der nach dem Mischen gemäß π_{x^-} bereits auf dieser Karte steht. Anschließend mischen wir die Karten gemäß π_{y^-} . Die Karte an Position j trägt also jeweils die folgenden Werte:

$$x_j \xrightarrow{\pi_{x^-}} x_{\pi_{x^-}(j)} \xrightarrow{y_j} (x_{\pi_{x^-}(j)}, y_j) \xrightarrow{\pi_{y^-}} (x_{\pi_{x^-} \circ \pi_{y^-}(j)}, y_{\pi_{y^-}(j)}). \quad (2.1.16)$$

Wegen $x_{\pi_{x^-}(1)} \leq \dots \leq x_{\pi_{x^-}(n)}$ gilt

$$y_{\pi_{y^-}(j)} \leq y_{\pi_{y^-}(j+1)} \implies x_{\pi_{x^-} \circ \pi_{y^-}(j)} \leq x_{\pi_{x^-} \circ \pi_{y^-}(j+1)}.$$

Daraus folgt, dass die Werte $(x_{\pi_{x^-} \circ \pi_{y^-}(1)}, y_{\pi_{y^-}(1)}), \dots, (x_{\pi_{x^-} \circ \pi_{y^-}(n)}, y_{\pi_{y^-}(n)})$ der Karten nach dem Mischen gemäß $\pi_{x^-} \circ \pi_{y^-}$ lexikographisch nach der zweiten Komponente sortiert sind. Es gilt also

$$(x_{\pi_{x^-} \circ \pi_{y^-}(1)}, y_{\pi_{y^-}(1)}) \leq \dots \leq (x_{\pi_{x^-} \circ \pi_{y^-}(n)}, y_{\pi_{y^-}(n)}).$$

Notieren wir nun vor dem Mischen gemäß $\pi_{x^-} \circ \pi_{y^-}$ den Wert $(x_j, y_{(\pi_{x^-})^{-1}(j)})$ auf Karte j , so trägt die Karte, die sich nach dem Mischen gemäß $\pi_{x^-} \circ \pi_{y^-}$ an Position j befindet, den Wert

$$(x_{\pi_{x^-} \circ \pi_{y^-}(j)}, y_{(\pi_{x^-})^{-1} \circ \pi_{x^-} \circ \pi_{y^-}(j)}) = (x_{\pi_{x^-} \circ \pi_{y^-}(j)}, y_{\pi_{y^-}(j)}).$$

Wir können also in einem einzelnen Mischvorgang alle Karten nach aufsteigenden Werten $(x_1, y_{(\pi_{x^-})^{-1}(1)}), \dots, (x_n, y_{(\pi_{x^-})^{-1}(n)})$ sortieren und erhalten dieselbe Permutation wie beim sukzessiven Mischen nach π_{x^-} und π_{y^-} , das heißt es gilt

$$\pi_{x^-} \circ \pi_{y^-} = \pi_{(x, y_{(\pi_{x^-})^{-1}})} = \pi_{(x, y_{\pi_x})},$$

wobei die letzte Gleichheit aus $(\pi_{x^-})^{-1} = \pi_x$ folgt. Da die Rangfolge der Tupel $(x_1, y_{\pi_x(1)}), \dots, (x_n, y_{\pi_x(n)})$ und der Werte $ay_{\pi_x(1)} + x_1, \dots, ay_{\pi_x(n)} + x_n$ jedoch nach Lemma 2.1.8 identisch ist, gilt

$$\pi_{(x, y_{\pi_x})} = \pi_{(x \circ y)^-}$$

und somit insgesamt (2.1.15). □

Für den Beweis des nächsten Satzes benötigen wir noch ein weiteres technisches Lemma, das zwei wichtige Aussagen über die in (2.1.14) definierten Verknüpfungen $x \circ y$ und $y \circ x$, $x \in \mathbb{N}_{0,a-1}^n$, $y \in \mathbb{N}_{0,b-1}^n$, beinhaltet.

2.1.10 Lemma. *Seien $a, b \in \mathbb{N}_{\geq 2}$, $x, x' \in \mathbb{N}_{0,a-1}^n$ und $y, y' \in \mathbb{N}_{0,b-1}^n$. Dann gelten die folgenden Aussagen:*

- (i) $x \circ y = x' \circ y' \iff x = x', y = y' \iff y \circ x = y' \circ x'$,
- (ii) $\mathbb{N}_{0,ab-1}^n = \{y \circ x \mid (x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n\}$.

Beweis. zu (i): Wir führen den Beweis der zweiten Äquivalenz, da wir diese im Beweis von Satz 2.1.11 benutzen werden. Die erste Äquivalenz wird analog bewiesen. Aus $y \circ x = y' \circ x'$ folgt zunächst für $i = 1, \dots, n$

$$bx_{\pi_y(i)} + y_i = bx'_{\pi_{y'}(i)} + y'_i.$$

Wegen $x_i \in \{0, \dots, a-1\}$ und $y_i \in \{0, \dots, b-1\}$ für $i = 1, \dots, n$ und $a, b \geq 2$ folgt dann aus der Eindeutigkeit der Division mit Rest

$$bx_{\pi_y(i)} = bx'_{\pi_{y'}(i)} \text{ und } y_i = y'_i \quad (2.1.17)$$

für $i = 1, \dots, n$. Damit ergibt sich $y = y'$ und daraus unmittelbar $\pi_y = \pi_{y'}$. Nach (2.1.17) erhalten wir dann $bx = bx'$. Insgesamt gilt also $x = x'$ und $y = y'$. Die umgekehrte Richtung ist wegen $\pi_y = \pi_{y'}$ trivial.

zu (ii): „ \subset “ Sei $z = (z_1, \dots, z_n) \in \mathbb{N}_{0,ab-1}^n = \{0, \dots, ab-1\}^n$ beliebig. Wir definieren $y = (y_1, \dots, y_n)$ durch

$$y_i \stackrel{\text{def}}{=} z_i \pmod{b}, \quad i = 1, \dots, n,$$

und $x = (x_1, \dots, x_n)$ durch

$$x_{\pi_y(i)} \stackrel{\text{def}}{=} \frac{z_i - y_i}{b}, \quad i = 1, \dots, n.$$

Dann gilt $x \in \mathbb{N}_{0,a-1}^n$, $y \in \mathbb{N}_{0,b-1}^n$ und

$$\begin{aligned} y \circ x &= bx_{\pi_y} + y \\ &= (bx_{\pi_y(1)} + y_1, \dots, bx_{\pi_y(n)} + y_n) \end{aligned}$$

$$\begin{aligned}
&= \left(b \cdot \frac{z_1 - y_1}{b} + y_1, \dots, b \cdot \frac{z_n - y_n}{b} + y_n \right) \\
&= (z_1, \dots, z_n) = z.
\end{aligned}$$

Daraus folgt $\mathbb{N}_{0,ab-1}^n \subset \{y \circ x \mid (x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n\}$.

„ \supset “ Für die umgekehrte Inklusion notieren wir

$$0 \leq bx_{\pi_y(i)} \leq b(a-1) \text{ und } 0 \leq y_i \leq b-1$$

für alle $(x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n$ und $i = 1, \dots, n$. Dann folgt für alle $(x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n$ und $i = 1, \dots, n$

$$0 \leq bx_{\pi_y(i)} + y_i \leq b(a-1) + b-1 = ab-1.$$

Es gilt also $y \circ x \in \mathbb{N}_{0,ab-1}^n$ für alle $(x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n$. \square

Der folgende Satz zeigt, dass das sukzessive Mischen nach den Mischnmethoden $Q_{a,n}$ und $Q_{b,n}$ als einmaliges Mischen nach der Mischnmethode $Q_{ab,n}$ interpretiert werden kann. Diese einfache Beziehung erlaubt eine erhebliche Vereinfachung des Riffle Shuffles, denn sie ermöglicht es, dass eine beliebige Anzahl sukzessiver Mischnvorgänge mit beliebigen Päckchenanzahlen als einmaliger Mischnvorgang aufgefasst werden kann. Wir präzisieren diese Aussage im anschließenden Korollar.

2.1.11 Satz. Seien $a, b \in \mathbb{N}_{\geq 2}$ und $n \in \mathbb{N}$. Ein GSR-(a, n)-Shuffle gefolgt von einem GSR-(b, n)-Shuffle ist äquivalent zu einem GSR-(ab, n)-Shuffle, das heißt

$$Q_{a,n} * Q_{b,n} = Q_{ab,n}.$$

Beweis. Seien $a, b \in \mathbb{N}_{\geq 2}$ und $n \in \mathbb{N}$ beliebig. Nach Bemerkung 1.1.2 und den vorangehenden Erläuterungen ist zu zeigen $Q_{a,n} * Q_{b,n} = Q_{ab,n}$. Sei $\pi \in \mathcal{S}_n$ beliebig. Dann gilt wegen Lemma 2.1.9 und Lemma 2.1.10 unter Verwendung des inversen Modells

$$\begin{aligned}
Q_{ab,n}(\{\pi\}) &= |\{z \in \mathbb{N}_{0,ab-1}^n \mid \pi_{z^-} = \pi^{-1}\}|/(ab)^n \\
&= |\{(x, y) \in \mathbb{N}_{0,a-1}^n \times \mathbb{N}_{0,b-1}^n \mid \pi_{(y \circ x)^-} = \pi^{-1}\}|/(ab)^n \\
&= |\{x \in \mathbb{N}_{0,a-1}^n, y \in \mathbb{N}_{0,b-1}^n \mid \pi_{y^-} \circ \pi_{x^-} = \pi^{-1}\}|/(ab)^n
\end{aligned}$$

$$\begin{aligned}
&= \left| \bigcup_{\sigma \in \mathcal{S}_n} \{x \in \mathbb{N}_{0,a-1}^n, y \in \mathbb{N}_{0,b-1}^n \mid \pi_{x^-} = \sigma^{-1}, \pi_{y^-} = \pi^{-1}\sigma\} \right| / (ab)^n \\
&= \sum_{\sigma \in \mathcal{S}_n} \frac{|\{x \in \mathbb{N}_{0,a-1}^n \mid \pi_{x^-} = \sigma^{-1}\}|}{a^n} \cdot \frac{|\{y \in \mathbb{N}_{0,b-1}^n \mid \pi_{y^-} = \pi^{-1}\sigma\}|}{b^n} \\
&= \sum_{\sigma \in \mathcal{S}_n} Q_{a,n}(\{\sigma\}) Q_{b,n}(\{(\pi^{-1}\sigma)^{-1}\}) = \sum_{\sigma \in \mathcal{S}_n} Q_{a,n}(\{\sigma\}) Q_{b,n}(\{\sigma^{-1}\pi\}) \\
&= Q_{a,n} * Q_{b,n}(\{\pi\}).
\end{aligned}$$

Hierbei gehen in der zweiten Zeile entscheidend (i) und (ii) aus Lemma 2.1.10 ein. Es gilt also $Q_{ab,n} = Q_{a,n} * Q_{b,n}$. \square

Aus Satz 2.1.11 ergibt sich direkt das folgende Korollar.

2.1.12 Korollar. Seien $m, n \in \mathbb{N}$, $a_1, \dots, a_m \in \mathbb{N}_{\geq 2}$, $\pi \in \mathcal{S}_n$ und $a \stackrel{\text{def}}{=} \prod_{i=1}^m a_i$. Dann gilt

$$Q_{a_1,n} * \dots * Q_{a_m,n} = Q_{a_{\pi(1)},n} * \dots * Q_{a_{\pi(m)},n} = Q_{a,n}.$$

2.2 Aufsteigende Sequenzen

In diesem Abschnitt geben wir die Wahrscheinlichkeit, dass ein GSR- (a, n) -Shuffle in der Permutation $\pi \in \mathcal{S}_n$ resultiert, explizit an. Dafür ist es notwendig, den Begriff der aufsteigenden Sequenz einzuführen. Unter Benutzung von Korollar 2.1.12 ist es dadurch sogar möglich, die Verteilung $Q_{a,n}^{*(m)}$ für alle $m \geq 1$ zu bestimmen. Mit Blick auf (1.1.4) sind damit schon die m -Schritt Übergangswahrscheinlichkeiten des $Q_{a,n}$ -Random Walks auf \mathcal{S}_n gegeben.

2.2.1 Definition. Sei $\pi \in \mathcal{S}_n$. Eine *aufsteigende Sequenz* oder kurz *Sequenz* von π ist eine Kette $((i_1, \pi(i_1)), \dots, (i_k, \pi(i_k)))$, so dass gilt

- (i) $i_1, \dots, i_k \in \mathbb{N}_{\leq n}$, $i_1 < \dots < i_k$,
- (ii) $\pi(i_j) + 1 = \pi(i_{j+1})$ für $j = 1, \dots, k - 1$,
- (iii) $\pi(l) \neq \pi(i_1) - 1$ für $l = 1, \dots, i_1 - 1$ und $\pi(l) \neq \pi(i_k) + 1$ für $l = i_k + 1, \dots, n$.

Eine aufsteigende Sequenz von $\pi \in \mathcal{S}_n$ kann somit auch als *maximal aufsteigende Teilkette von π* bezeichnet werden. Jede Permutation π zerfällt vollständig in aufsteigende Sequenzen und ist durch Angabe aller ihrer aufsteigenden Sequenzen eindeutig festgelegt. Wir sagen, die Sequenz $((i_{k+1}, \pi(i_{k+1})), \dots, (i_l, \pi(i_l)))$ folgt auf die Sequenz $((i_1, \pi(i_1)), \dots, (i_k, \pi(i_k)))$, falls $\pi(i_k) + 1 = \pi(i_{k+1})$ gilt. Aufgrund von Definition 2.2.1 gilt dann notwendigerweise $i_{k+1} > i_k$, das heißt die beiden Sequenzen lassen sich nicht zu einer einzelnen Sequenz vereinen.

Aufsteigende Sequenzen sind eng mit den Sprungstellen einer Permutation verbunden.

2.2.2 Definition. Sei $\pi \in \mathcal{S}_n$ und $i \in \{1, \dots, n-1\}$. Die Permutation π hat eine *Sprungstelle* oder einen *Descent* in i , falls $\pi(i) > \pi(i+1)$.

Sei $\mathcal{R}_{n,k} \stackrel{\text{def}}{=} \{\pi \in \mathcal{S}_n \mid \pi \text{ hat } k \text{ aufsteigende Sequenzen}\}$, $k = 1, \dots, n$, und $\mathcal{D}_{n,k} \stackrel{\text{def}}{=} \{\pi \in \mathcal{S}_n \mid \pi \text{ hat } k \text{ Descente}\}$, $k = 1, \dots, n-1$. Wir definieren die Funktionen $R_n, D_n : \mathcal{S}_n \longrightarrow \mathbb{N}_{\leq n}$ für $\pi \in \mathcal{S}_n$ durch

$$\begin{aligned} R_n(\pi) = r &\iff \pi \in \mathcal{R}_{n,r} \text{ und} \\ D_n(\pi) = r &\iff \pi \in \mathcal{D}_{n,r}. \end{aligned}$$

Wir setzen von nun an $R = R_n$ und $D = D_n$. Aufgrund von Definition 2.2.1 und 2.2.2 gilt $R(\mathcal{S}_n) = \{1, \dots, n\}$, $D(\mathcal{S}_n) = \{1, \dots, n-1\}$ und für $\pi \in \mathcal{S}_n$ $D(\pi) = |\{i \leq n-1 \mid \pi(i) > \pi(i+1)\}|$.

2.2.3 Beispiel. Für die Permutation $\pi = [4, 1, 2, 5, 7, 3, 6, 8] \in \mathcal{S}_8$ gilt $R(\pi) = 3$, denn π hat die drei aufeinanderfolgenden aufsteigenden Sequenzen

$$((2, 1), (3, 2), (6, 3)), ((1, 4), (4, 5)), (7, 6)) \text{ und } ((5, 7), (8, 8)).$$

Für die Inverse $\pi^{-1} = [2, 3, 6, 1, 4, 7, 5, 8]$ gilt $D(\pi) = 2$. Sie hat Descente in 3 und 6.

Aufschluss über den Zusammenhang zwischen aufsteigenden Sequenzen und Descenten gibt das folgende Lemma.

2.2.4 Lemma. Sei $\pi \in \mathcal{S}_n$ und $r \in \mathbb{N}_{\leq n}$. π besitzt genau dann r aufsteigende Sequenzen, wenn π^{-1} $r - 1$ Descente hat, das heißt

$$R(\pi) = r \iff D(\pi^{-1}) = r - 1.$$

Beweis. Für $r = 1$ folgt die Behauptung direkt aus

$$R(\pi) = 1 \iff \pi = \text{id} \iff D(\pi) = 0.$$

Sei also $r \geq 2$.

„ \Rightarrow “ Wir wählen $k_0, \dots, k_r \in \mathbb{N}$ mit

$$0 = k_0 < k_1 < \dots < k_{r-1} < k_r = n$$

und paarweise verschiedene $i_1, \dots, i_n \in \mathbb{N}_{\leq n}$, so dass die r aufsteigenden Sequenzen von π gegeben sind durch

$$\begin{aligned} ((i_1, 1)), \dots, (i_{k_1}, k_1)), ((i_{k_1+1}, k_1 + 1)), \dots, (i_{k_2}, k_2)), \dots, \\ ((i_{k_{r-1}+1}, k_{r-1} + 1)), \dots, (i_n, n)). \end{aligned}$$

Dann gilt $\pi(i_j) = j$ und $\pi^{-1}(j) = i_j$ für $j = 1, \dots, n$. Die Anzahl der Descente von π^{-1} ist gegeben durch

$$\begin{aligned} D(\pi^{-1}) &= |\{j \leq n-1 \mid \pi^{-1}(j) > \pi^{-1}(j+1)\}| \\ &= |\{j \leq n-1 \mid i_j > i_{j+1}\}|. \end{aligned}$$

Aufgrund der Definition der aufsteigenden Sequenz gilt $i_j < i_{j+1}$ für $j = k_l + 1, \dots, k_{l+1}$, $l = 0, \dots, r-1$ und

$$i_j > i_{j+1} \text{ für } j = k_l \text{ und } l = 1, \dots, r-1.$$

Daraus folgt unmittelbar $D(\pi^{-1}) = r - 1$.

„ \Leftarrow “ Seien $k_1, \dots, k_{r-1} \in \mathbb{N}$ mit

$$0 \stackrel{\text{def}}{=} k_0 < k_1 < \dots < k_{r-1} < k_r \stackrel{\text{def}}{=} n$$

die Sprungstellen von π^{-1} und $i_1, \dots, i_n \in \mathbb{N}_{\leq n}$, so dass $\pi^{-1}(j) = i_j$ für $j = 1, \dots, n$. Dann gilt

$$\begin{aligned}\pi^{-1}(k_l) &> \pi^{-1}(k_l + 1) \text{ für } l = 1, \dots, r - 1 \text{ und} \\ \pi^{-1}(k_l + 1) &< \dots < \pi^{-1}(k_{l+1}) \text{ für } l = 0, \dots, r - 1.\end{aligned}$$

Daraus folgt wegen $\pi^{-1}(j) = i_j$, $j = 1, \dots, n$,

$$\begin{aligned}i_{k_1} &> i_{k_1+1}, i_{k_2} > i_{k_2+1} \dots, i_{k_{r-1}} > i_{k_{r-1}+1} \text{ und} \\ i_1 &< \dots < i_{k_1}, i_{k_1+1} < \dots < i_{k_2}, \dots, i_{k_{r-1}+1} < \dots < i_n.\end{aligned}$$

Wegen $\pi(i_j) = j$ für $j = 1, \dots, n$ sind die aufsteigenden Sequenzen von π dann gegeben durch

$$\begin{aligned}&((i_1, 1)), \dots, (i_{k_1}, k_1)), ((i_{k_1+1}, k_1 + 1)), \dots, (i_{k_2}, k_2)), \dots, \\ &((i_{k_{r-1}+1}, k_{r-1} + 1)), \dots, (i_n, n)).\end{aligned}$$

Also ist $R(\pi) = r$. □

Bei einem GSR- (a, n) -Shuffle eines ungemischten Kartenstapels helfen uns die aufsteigenden Sequenzen, die ursprünglichen Päckchen wiederzuentdecken: Sie sind nach dem Ineinanderblättern als aufsteigende Sequenzen zu erkennen (siehe Beispiel 2.1.1). Dies wird im Beweis des folgenden Theorems (siehe Theorem 3, Abschnitt 3 in [6]) deutlich. Wie eingangs dieses Abschnitts beschrieben, erweist es sich als Schlüssel zur Bestimmung der Verteilung $Q_{a,n}^{*(m)}$ und somit zur Berechnung der Totalvariation $\|Q_{a,n}^{*(m)} - U_{S_n}\|$ für $m \geq 1$.

2.2.5 Theorem. *Sei $n \in \mathbb{N}$, $a \in \mathbb{N}_{\geq 2}$ und $\pi \in \mathcal{S}_n$. Die Wahrscheinlichkeit, dass ein GSR- (a, n) -Shuffle in der Permutation π resultiert, ist gegeben durch*

$$Q_{a,n}(\{\pi\}) = \frac{\binom{n+a-R(\pi)}{n}}{a^n}, \quad (2.2.1)$$

wobei $\binom{m}{n} \stackrel{\text{def}}{=} 0$ für $m < n$.

Beweis. Seien $n \in \mathbb{N}$, $\pi \in \mathcal{S}_n$ beliebig und $R(\pi) = r$. Ausgehend vom Maximum-Entropie- bzw. inversen Modell ist die gesuchte Wahrscheinlichkeit bestimmt durch

die Anzahl aller möglichen Arten, einen ungemischten Stapel in a Päckchen A_1, \dots, A_a zu unterteilen, so dass π ein möglicher Mischvorgang ist. Dies gilt, da die Reihenfolge, in der die Karten ineinandergeblättert werden, bereits eindeutig durch π festgelegt ist. Wir wählen $k_0, \dots, k_r \in \mathbb{N}$ mit

$$0 = k_0 < k_1 < \dots < k_{r-1} < k_r = n,$$

so dass die r aufsteigenden Sequenzen von π gegeben sind durch

$$\begin{aligned} ((i_1, 1)), \dots, (i_{k_1}, k_1)), ((i_{k_1+1}, k_1 + 1)), \dots, (i_{k_2}, k_2)), \dots, \\ ((i_{k_{r-1}+1}, k_{r-1} + 1)), \dots, (i_n, n)). \end{aligned}$$

Sei nun A_1, \dots, A_a eine Päckchenfolge, so dass π ein möglicher Mischvorgang ist. Da sich die Karten in den einzelnen Päckchen vor und nach dem Ineinanderblättern in derselben Reihenfolge befinden, besteht jede Sequenz in π aus aufeinanderfolgenden Päckchen A_k, \dots, A_{k+l} mit geeignetem $k \in \mathbb{N}_{\leq a}$ und $l \in \{0, \dots, a-1\}$. Die Vereinigung $A_k \cup \dots \cup A_{k+l}$ von aufeinanderfolgenden Päckchen A_k, \dots, A_{k+l} bezeichnen wir als *Paket*. Zwei aufeinanderfolgende Sequenzen $((i_{k_{j-1}}, k_{j-1}), \dots, (i_{k_j}, k_j))$ und $((i_{k_j+1}, k_j + 1), \dots, (i_{k_{j+1}}, k_{j+1}))$ müssen in verschiedenen Paketen aufeinanderfolgender Päckchen liegen, da es sonst nicht möglich ist, die Karte $k_j + 1$ vor der Karte k_j einzusortieren. Hieraus folgt unmittelbar $Q_{a,n}(\{\pi\}) = 0$ für $r > a \wedge n$. Sei also im Folgenden $r \leq a \wedge n$. Dann existieren $j_1, \dots, j_r \in \{1, \dots, a\}$, mit $1 \leq j_1 < \dots < j_{r-1} < j_r = a$ und

$$\begin{aligned} A_1 \cup \dots \cup A_{j_1} = \{1, \dots, k_1\}, A_{j_1+1} \cup \dots \cup A_{j_2} = \{k_{1+1}, \dots, k_2\}, \dots, \\ A_{j_{r-1}+1} \cup \dots \cup A_a = \{k_{r-1} + 1, \dots, n\}. \end{aligned} \tag{2.2.2}$$

Wir zählen nun alle Möglichkeiten ab, einen ungemischten Stapel in a Päckchen A'_1, \dots, A'_a zu unterteilen, so dass π ein möglicher Mischvorgang ist. Wir zeigen unsere Behauptung auf zwei verschiedene Arten, indem wir zwei unterschiedliche kombinatorische Abzählargumente verwenden.

1. Abzählargument (Kombination ohne Wiederholung): Wir codieren zunächst jede Päckchenfolge A'_1, \dots, A'_a eindeutig durch ein $(n + a - 1)$ -Tupel bestehend aus n Sternchen „*“ und $a - 1$ Trennstrichen „|“. Enthalten die Päckchen

A'_1, \dots, A'_a l_1, \dots, l_a Karten, $(l_1, \dots, l_a) \in \sum_n^a$, so besetzen wir den l_1, \dots, l_a entsprechend viele Positionen mit Sternchen und trennen die so entstandenen a Blöcke durch $a - 1$ Trennstriche:

$$A'_1, \dots, A'_a \leftrightarrow (\underbrace{*, \dots, *}_{l_1\text{-mal}} | \underbrace{*, \dots, *}_{l_2\text{-mal}} | \dots | \underbrace{*, \dots, *}_{l_a\text{-mal}}). \quad (2.2.3)$$

Um nun alle Möglichkeiten abzuzählen, einen ungemischten Stapel so in a Päckchen A'_1, \dots, A'_a zu unterteilen, dass π ein möglicher Mischvorgang ist, genügt es schon, alle Möglichkeiten zu bestimmen, die $a - 1$ Trennstriche auf $n + a - 1$ Positionen zu verteilen, so dass die gemäß (2.2.3) bestimmte Päckchenfolge zum Mischvorgang π gehört. Die n Sternchen nehmen dann gerade die verbleibenden Positionen ein. Die Reihenfolge, in der die einzelnen Karten ineinandergeblättert werden, ist durch π eindeutig festgelegt ist. Daher führt eine Päckchenfolge A'_1, \dots, A'_a genau dann zum Mischvorgang π , wenn sie von der Form (2.2.2) ist. Dann sind aber $r - 1$ Pakete bereits durch π festgelegt, das heißt die Positionen der $r - 1$ Trennstriche zwischen diesen Paketen sind vorgegeben: Vor dem ersten Trennstrich müssen sich k_1 Sternchen befinden, vor dem zweiten Trennstrich k_2 Sternchen usw. Es bleiben also noch $(a - 1) - (r - 1) = a - r$ Trennstriche, die beliebig auf die $n + a - 1$ Positionen verteilt werden können. Hierbei ist zu beachten, dass tatsächlich jede der $n + a - 1$ Positionen von den $a - r$ Trennstrichen eingenommen werden kann, da die Positionen der übrigen $r - 1$ Trennstriche zwar festgelegt sind, aber von den vorher schon vergebenen Positionen abhängen: Werden zunächst $a - r$ Trennstriche verteilt, so erhalten wir ein $(n + a - 1)$ -Tupel, bestehend aus $a - r$ Trennstrichen und $(n + a - 1) - (a - r) = n + r - 1$ Leerstellen „ \square “, der Form

$$(\underbrace{\square, \dots, \square}_{m_1\text{-mal}} | \underbrace{\square, \dots, \square}_{m_2\text{-mal}} | \dots | \underbrace{\square, \dots, \square}_{m_a\text{-mal}}),$$

mit geeigneten $m_1, \dots, m_a \in \sum_n^a$. Die n Sternchen und die übrigen $r - 1$ Trennstriche müssen dann so auf die Leerstellen-Positionen verteilt werden, dass sich vor dem ersten der $r - 1$ Trennstriche k_1 Sternchen befinden, vor dem zweiten Trennstrich k_2 Sternchen usw. Wir kennzeichnen die $r - 1$ festgelegten Trennstriche

durch „ \downarrow “:

$$\underbrace{(*, \dots, * | \dots | *, \dots, *)}_{k_1\text{-mal } „*“} \downarrow \underbrace{(*, \dots, * | \dots, | *, \dots, *)}_{k_2\text{-mal } „*“} \downarrow \dots \downarrow \underbrace{(*, \dots, * | \dots | *, \dots, *)}_{k_r\text{-mal } „*“}.$$

Es gibt $\binom{n+a-r}{a-r} = \binom{n+a-r}{n}$ Möglichkeiten die $a-r$ Trennstiche auf $n+a-r$ Positionen zu verteilen und somit $\binom{n+a-r}{n}$ Möglichkeiten den Kartenstapel so in a Päckchen zu teilen, dass π ein möglicher Mischvorgang ist. Somit gilt also

$$Q_{a,n}(\{\pi\}) = \frac{\binom{n+a-r}{n}}{a^n},$$

da es insgesamt a^n mögliche GSR- (a, n) -Shuffle gibt.

2. Abzählargument (Kombination mit Wiederholung): Die n Karten seien nach aufsteigenden Werten von links nach rechts vor uns ausgelegt. Zwischen zwei Karten, sowie vor der ersten und nach der letzten Karte sei jeweils ein Fach aufgestellt (siehe Abbildung 2.4). Insgesamt benötigen wir hierzu $n+1$ Fächer.



Abbildung 2.4: Fächerverteilung: $n+1$ Fächer zwischen n Karten.

Wir können nun jede Päckchenfolge A'_1, \dots, A'_a eindeutig durch $a-1$ unterscheidbare Kugeln, die wir auf die $n+1$ Fächer verteilen, identifizieren: Beginnt bei Karte j ein Päckchen, so legen wir eine Kugel in das Fach links von Karte j . Bleibt ein Päckchen leer, so legen wir eine Kugel in das Fach links von der Karte, bei der das nächste nichtleere Päckchen beginnt. Ist das letzte Päckchen oder sind mehrere der letzten Päckchen leer, so legen wir eine oder der Anzahl der leeren letzten Päckchen entsprechend viele Kugeln in das letzte Fach. Dabei benötigen wir nur $a-1$ und nicht etwa a Kugeln, da bei Karte 1 immer mindestens ein Päckchen beginnt. Enthalten die Päckchen also A'_1, \dots, A'_a l_1, \dots, l_a

Karten, $(l_1, \dots, l_a) \in \sum_n^a$, das heißt

$$A'_1 = \{1, \dots, l_1\}, A'_2 = \{l_1 + 1, \dots, l_1 + l_2\}, \dots, A'_a = \{l_1 + \dots + l_{a-1} + 1, \dots, n\},$$

so verteilen wir die Kugeln den Werten l_1, \dots, l_a entsprechend auf die Fächer: Die erste Kugel legen wir in das Fach links von Karte $l_1 + 1$, die zweite Kugel in das Fach links von Karte $l_1 + l_2 + 1$ usw., bis hin zur $(a - 1)$ -Kugel, die wir in das Fach links von Karte $l_1 + \dots + l_{a-1} + 1$ legen, bzw in das letzte Fach, falls $l_1 + \dots + l_{a-1} + 1 = n + 1$. Hierbei ist zu beachten, dass für den Fall $l_1 = 0$ das erste Päckchen leer bleibt, das heißt $A'_1 = \emptyset$, und die erste der $a - 1$ Kugeln in das Fach links von Karte 1 gelegt wird. Ebenso können für den Fall, dass nicht sämtlich l_1, \dots, l_a von Null verschieden sind, mehrere Kugeln in einem Fach liegen (siehe Abbildung 2.5: Das erste und dritte Päckchen A_1 und A_3 sind leer, das zweite Päckchen A_2 besteht aus den Karten 1, 2, 3).

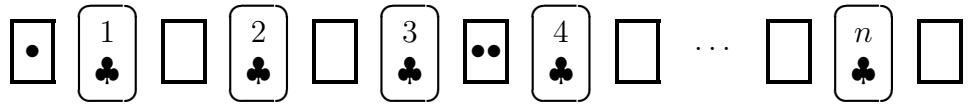


Abbildung 2.5: Beispiel für eine Kugelverteilung.

Eine Päckchenfolge A'_1, \dots, A'_a führt genau dann zum Mischvorgang π , wenn sie von der Form (2.2.2) ist, das heißt

$$\begin{aligned} A'_1 \cup \dots \cup A'_{j_1} &= \{1, \dots, k_1\}, A'_{j_1+1} \cup \dots \cup A'_{j_2} = \{k_{1+1}, \dots, k_2\}, \dots, \\ A'_{j_{r-1}+1} \cup \dots \cup A'_a &= \{k_{r-1} + 1, \dots, n\}, \end{aligned}$$

wobei $0 < k_1 < \dots < k_{r-1} < n$ und $1 \leq j_1 < \dots < j_{r-1} < j_r = a$. Dann sind $r - 1$ Pakete bereits durch π festgelegt, das heißt die Fachbelegung von $r - 1$ Kugeln, die den Beginn der einzelnen Pakete festlegen, sind vorgegeben: Die erste Kugel liegt in dem Fach links von Karte $k_1 + 1$, die zweite Kugel liegt in dem Fach links von Karte $k_2 + 1$ usw. und die letzte Kugel im Fach links von Karte $k_{r-1} + 1$. Es bleiben also noch $(a - 1) - (r - 1) = a - r$ Kugeln, die beliebig auf die $n + 1$ Fächer verteilt werden können, dabei können mehrere Kugeln in dasselbe Fach gelegt werden, da auch leere Päckchen zugelassen sind. Es gibt insgesamt $\binom{(n+1)+a-r-1}{(n+1)-1} = \binom{n+a-r}{n}$

Möglichkeiten die $a - r$ ununterscheidbaren Kugeln auf $n + 1$ Fächer zu verteilen und somit $\binom{n+a-r}{n}$ Möglichkeiten den Kartenstapel so in a Päckchen zu teilen, dass π ein möglicher Mischvorgang ist. Insgesamt zeigt dies wiederum

$$Q_{a,n}(\{\pi\}) = \frac{\binom{n+a-r}{n}}{a^n},$$

da es insgesamt a^n mögliche GSR- (a, n) -Shuffle gibt. \square

Für $a \in \mathbb{N}_{\geq 2}$ ist der Träger von $Q_{a,n}$ nach Theorem 2.2.5 gegeben durch

$$\text{supp}(Q_{a,n}) = \{\pi \in \mathcal{S}_n \mid R(\pi) \in \{1, \dots, a \wedge n\}\}. \quad (2.2.4)$$

Ferner ergibt sich zusammen mit Korollar 2.1.12 die folgende Verallgemeinerung von Theorem 2.2.5, in der die Verteilung $Q_{a,n}^{*(m)}$ explizit angegeben ist.

2.2.6 Korollar. (i) Seien $m, n \in \mathbb{N}$, $a_1, \dots, a_m \in \mathbb{N}_{\geq 2}$, $a \stackrel{\text{def}}{=} \prod_{i=1}^m a_i$ und $\pi \in \mathcal{S}_n$. Dann gilt

$$Q_{a_1,n} * \dots * Q_{a_m,n}(\{\pi\}) = \frac{\binom{n+a-R(\pi)}{n}}{a^n}. \quad (2.2.5)$$

(ii) Seien $m, n \in \mathbb{N}$, $a \in \mathbb{N}_{\geq 2}$ und $\pi \in \mathcal{S}_n$. Dann gilt

$$Q_{a,n}^{*(m)}(\{\pi\}) = \frac{\binom{n+a^m-R(\pi)}{n}}{a^{mn}}. \quad (2.2.6)$$

Aus Korollar 2.2.6 (2.2.6) und Lemma 1.1.4 erhalten als weitere triviale Folgerung das folgende Korollar.

2.2.7 Korollar. Seien $m, n \in \mathbb{N}$ und $a \in \mathbb{N}_{\geq 2}$. Eine Permutation π ist ein mögliches Ergebnis von m sukzessiven GSR- (a, n) -Mischvorgängen genau dann, wenn $R(\pi) \in \{1, \dots, a^m \wedge n\}$, das heißt

$$\text{supp}(Q_{a,n}^{*(m)}) = \{\pi \in \mathcal{S}_n \mid R(\pi) \in \{1, \dots, a^m \wedge n\}\} = (\text{supp}(Q_{a,n}))^{\circ(m)}.$$

Für $a \in \mathbb{N}_{\geq 2}$ ist die Funktion $r \mapsto \binom{n+a-r}{n}/a^n$ auf $\{1, \dots, a \wedge n\}$ streng monoton fallend. Wegen $R(\text{id}) = 1$ ist daher $Q_{a,n}(\{\text{id}\}) > Q_{a,n}(\{\pi\})$ für alle $\pi \in \mathcal{S}_n \setminus \{\text{id}\}$. Gegeben einen $Q_{a,n}$ -Random Walk $X = (X_m)_{m \geq 0}$ auf \mathcal{S}_n , ist demnach in jedem Mischvorgang das Mischen gemäß der Identität id wahrscheinlicher als das Mischen nach jeder anderen Permutation. Allerdings führt wiederholtes Mischen nach dem GSR- (a, n) -Shuffle tatsächlich dazu, dass ein Kartenstapel „gut durchgemischt“ wird:

2.2.8 Satz. Seien $n \in \mathbb{N}$ und $a \in \mathbb{N}_{\geq 2}$ fest und $(X_m)_{m \geq 0}$ ein $Q_{a,n}$ -Random Walk auf \mathcal{S}_n . Dann konvergieren die Verteilungen von X_m in Totalvariation gegen die Gleichverteilung $U_{\mathcal{S}_n}$ auf \mathcal{S}_n , das heißt

$$\lim_{m \rightarrow \infty} \|P^{X_m} - U_{\mathcal{S}_n}\| = \lim_{m \rightarrow \infty} \|Q_{a,n}^{*(m)} - U_{\mathcal{S}_n}\| = 0.$$

Beweis. $X = (X_m)_{m \geq 0}$ ist eine endliche Markov-Kette. Nach dem Ergodensatz reicht es daher zu zeigen, dass X (a) irreduzibel und (b) aperiodisch ist. Da für $n = 1$ nichts zu zeigen ist, sei $n \geq 2$ und $a \geq 2$.

zu (a): Jede Transposition $[i, i+1]_Z \in \mathcal{S}_n$ mit $i \leq n-1$ zerfällt in zwei Sequenzen

$$((1, 1), \dots, (i-1, i-1), (i+1, i)) \text{ und } ((i, i+1), (i+2, i+2), \dots, (n, n)),$$

was $[i, i+1]_Z \in \text{supp}(Q_{a,n})$ vermöge Korollar 2.2.7 impliziert.

Sei $[i, j]_Z$ eine Transposition mit $|i - j| > 1$. Wegen $[i, j]_Z = [j, i]_Z$ können wir $i < j$ annehmen. Dann hat $[i, j]_Z$ die drei aufeinanderfolgenden Sequenzen

$$((1, 1), \dots, (i-1, i-1), (j, i)), ((i+1, i+1), \dots, (j-1, j-1)) \text{ und} \\ ((i, j), (j+1, j+1), \dots, (n, n)).$$

Definieren wir nun zwei Permutationen $\pi_1, \pi_2 \in \mathcal{S}_n$ durch

$$\begin{aligned} \pi_2 &\stackrel{\text{def}}{=} \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & i & i+2 & \cdots & j & i+1 & j+1 & \cdots & n \end{pmatrix} \text{ und} \\ \pi_1 &\stackrel{\text{def}}{=} \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & j & i & \cdots & j-2 & j-1 & j+1 & \cdots & n \end{pmatrix}, \end{aligned}$$

so gilt

$$\pi_1 \circ \pi_2 = \begin{pmatrix} 1 & \cdots & i-1 & i & i+1 & \cdots & j-1 & j & j+1 & \cdots & n \\ 1 & \cdots & i-1 & j & i+1 & \cdots & j-1 & i & j+1 & \cdots & n \end{pmatrix} = [i, j]_Z$$

und $R(\pi_1) = R(\pi_2) = 2$. Also folgt wiederum mit Korollar 2.2.7

$$[i, j]_Z \in (\text{supp}(Q_{a,n}))^{\circ(2)} = \{\pi \circ \sigma \mid \pi, \sigma \in \text{supp}(Q_{a,n})\}.$$

Insgesamt gilt dann $\mathcal{T}_n \subset (\text{supp}(Q_{a,n}))^{\circ(2)}$ und somit wegen (1.2.1) $\langle \text{supp}(Q_{a,n}) \rangle = \mathcal{S}_n$. Also ist X nach Satz 1.1.3 (a) irreduzibel.

zu (b): Wegen $R(\text{id}) = 1$ folgt aus Korollar 2.2.7 $\text{id} \in \text{supp}(Q_{a,n})$ und daher

$$p_{\pi,\pi}^{(m)} = p_{\pi^{-1}\pi}^{(m)} = p_{\text{id}}^{(m)} = Q_{a,n}^{*(m)}(\{\text{id}\}) > 0$$

für alle $\pi \in \mathcal{S}_n$ und für alle $m \in \mathbb{N}$. Damit ist X aperiodisch.

Wir können die Aperiodizität von X auch auf andere Weise mit Hilfe von Satz 1.1.3 (b) und der Ergebnisse aus Abschnitt 1.2 zeigen: Wir zeigen $\text{supp}(Q_{a,n}) \not\subset \pi N$ für alle $\pi \in \mathcal{S}_n$ und alle nicht-trivialen Normalteiler von \mathcal{S}_n . Mit Blick auf die Liste der Normalteiler von \mathcal{S}_n in Abschnitt 1.2 ist für $n = 2$ nichts zu zeigen. Für $n = 3$ und $n \geq 5$ ist \mathcal{A}_n der einzige nicht-triviale Normalteiler von \mathcal{S}_n . Es existieren $\sigma_1, \sigma_2 \in \text{supp}(Q_{a,n})$ mit $\text{sgn}(\sigma_1) = 1$ und $\text{sgn}(\sigma_2) = -1$, etwa $\sigma_1 = \text{id}$ und $\sigma_2 = [1, 2]_Z$. Wegen $\text{sgn}(\rho) = 1$ für alle $\rho \in \mathcal{A}_n$ und da sgn ein Gruppenhomomorphismus ist, gilt jedoch für jedes $\pi \in \mathcal{S}_n$ entweder $\text{sgn}(\sigma) = 1$ für alle $\sigma \in \pi \mathcal{A}_n$, falls $\text{sgn}(\pi) = 1$, oder $\text{sgn}(\sigma) = -1$ für alle $\sigma \in \pi \mathcal{A}_n$, falls $\text{sgn}(\pi) = -1$. Daraus folgt $\text{supp}(Q_{a,n}) \not\subset \pi \mathcal{A}_n$ für $n = 3$ und $n \geq 5$. Für $n = 4$ gilt ebenfalls $\text{sgn}(\sigma) = 1$ für alle $\sigma \in \mathcal{V}_4$ bzw. \mathcal{A}_4 . Insgesamt gilt also $\text{supp}(Q_{a,n}) \not\subset \pi N$ für alle $\pi \in \mathcal{S}_n$ und alle nicht-trivialen Normalteiler von \mathcal{S}_n . Da X nach (a) irreduzibel ist, folgt die Aperiodizität von X also aus Satz 1.1.3 (b). \square

2.3 Eulersche Zahlen und aufsteigende Sequenzen

Im vorherigen Abschnitt haben wir gezeigt, dass die aufsteigenden Sequenzen der Schlüssel zur Angabe der Wahrscheinlichkeiten $Q_{a,n}^{*(m)}(\{\pi\})$ für $\pi \in \mathcal{S}_n$ und $m \geq 1$ sind. Aufsteigende Sequenzen sind eng mit den *Eulerschen Zahlen* verbunden. Diese Verbindung werden wir im Folgenden aufzeigen und für verschiedene Resultate nutzen.

Sei $n \in \mathbb{N}$. Für $r = 1, \dots, n$ bezeichnen wir die Anzahl aller Permutationen in \mathcal{S}_n mit r aufsteigenden Sequenzen mit $\mathcal{A}_{n,r}$, also $\mathcal{A}_{n,r} = |\mathcal{R}_{n,r}|$. Wir setzen $\mathcal{A}_{n,x} = 0$ für $x \neq 1, \dots, n$. Gegeben sei im Folgenden ein $Q_{a,n}$ -Random Walk $X = (X_m)_{m \geq 0}$. Die stochastische Folge $R(X) = (R(X_m))_{m \geq 0}$ gibt dann die Anzahlen der Sequenzen der Markov-Kette X an. Mit Korollar 2.2.6 erhalten wir die Verteilung von $R(X_m)$,

$m \geq 0$. Für $m \geq 0$ und $r = 1, \dots, n$ gilt

$$\begin{aligned} P(R(X_m) = r) &= P^{X_m}(\{\pi \in \mathcal{S}_n \mid R(\pi) = r\}) \\ &= Q_{a,n}^{*(m)}(\mathcal{R}_{n,r}) \\ &= \frac{\mathcal{A}_{n,r}}{a^{mn}} \binom{n+a^m-r}{n}. \end{aligned} \quad (2.3.1)$$

Summieren wir (2.3.1) über $r = 1, \dots, n$, so erhalten wir mit $m = 1$ für alle $n, a \in \mathbb{N}$

$$a^n = \sum_{r=1}^n \mathcal{A}_{n,r} \binom{a+n-r}{n}. \quad (2.3.2)$$

Gleichung (2.3.2) ist die sogenannte *Worpitzky-Identität*. Nach [26] entsprechen die $\mathcal{A}_{n,r}$ dann gerade den *Eulerschen Zahlen*. Verschiedene Eigenschaften der Eulerschen Zahlen und Aussagen über ihre Verbindungen zur Wahrscheinlichkeitstheorie und Kombinatorik finden sich ferner in [7], [8], [19], [22] und [23]. Für $r = 1, \dots, n$ erhalten wir aus der Definition der Eulerschen Zahlen die Rekursion

$$\mathcal{A}_{n,r} = r\mathcal{A}_{n-1,r} + (n+1-r)\mathcal{A}_{n-1,r-1},$$

wobei $\mathcal{A}_{0,0} \stackrel{\text{def}}{=} 1$, $\mathcal{A}_{n,0} \stackrel{\text{def}}{=} 0$ und $\mathcal{A}_{n-1,n} \stackrel{\text{def}}{=} 0$.

Wegen $\mathcal{R}_{n,1} = \{\text{id}\}$, $\mathcal{R}_{n,n} = \{[n, n-1, \dots, 1]\}$ und $|\mathcal{S}_n| = n!$ folgt unmittelbar

$$\begin{aligned} \mathcal{A}_{n,1} &= \mathcal{A}_{n,n} = 1 \text{ und} \\ \sum_{r=1}^n \mathcal{A}_{n,r} &= n!. \end{aligned}$$

Aus [26] erhalten wir ferner für $r = 1, \dots, n$ die Symmetrie-Eigenschaft

$$\mathcal{A}_{n,r} = \mathcal{A}_{n,n-r+1},$$

sowie die geschlossene Form der Eulerschen Zahlen

$$\mathcal{A}_{n,r} = \sum_{i=0}^{r-1} (-1)^i \binom{n+1}{i} (r-i)^n. \quad (2.3.3)$$

In Tabelle 2.1 sind die Werte der Eulerschen Zahlen für $1 \leq r \leq n \leq 8$ aufgelistet.

Die Eulerschen Zahlen besitzen neben der Verbindung zur Kombinatorik eine weitere wichtige wahrscheinlichkeitstheoretische Interpretation:

	$r=1$	2	3	4	5	6	7	8
n=1	1							
2	1	1						
3	1	4	1					
4	1	11	11	1				
5	1	26	66	26	1			
6	1	57	302	302	57	1		
7	1	120	1191	2416	1191	120	1	
8	1	247	4293	15619	15619	4293	247	1

Tabelle 2.1: Eulersche Zahlen $\mathcal{A}_{n,r}$ für $1 \leq r \leq n \leq 8$.

2.3.1 Satz. Gegeben seien stochastisch unabhängige und identisch $R(0, 1)$ -verteilte Zufallsgrößen X_1, \dots, X_n . Dann entspricht $\frac{1}{n!} \mathcal{A}_{n,k}$ für $k = 1, \dots, n$ der Wahrscheinlichkeit, dass die Summe von X_1, \dots, X_n im Intervall $(k-1, k]$ liegt, das heißt

$$P(S_n \in (k-1, k]) = \frac{1}{n!} \mathcal{A}_{n,k}, \quad (2.3.4)$$

wobei $S_n \stackrel{\text{def}}{=} \sum_{i=1}^n X_i$.

Beweis. Den folgenden kurzen Beweis entnehmen wir [23]. Sei F_n die Verteilungsfunktion von S_n . Dann gilt nach Theorem 1 in Abschnitt I.9 aus [11] für alle $x \geq 0$

$$F_n(x) = \frac{1}{n!} \sum_{i=0}^{\lfloor x \rfloor \wedge n} (-1)^i \binom{n}{i} (x-i)^n,$$

wobei $\lfloor x \rfloor$ die untere Gauss-Klammer von x sei. Sei nun $1 \leq k \leq n$. Dann folgt aus obiger Gleichung unter Beachtung von $\binom{n+1}{i} = \binom{n}{i} + \binom{n}{i-1}$ für $i = 1, \dots, n$ und $\binom{n}{-1} = 0$ zusammen mit (2.3.3)

$$\begin{aligned} F_n(k) - F_n(k-1) &= \frac{1}{n!} \sum_{i=0}^k (-1)^i \binom{n}{i} (k-i)^n - \frac{1}{n!} \sum_{i=0}^{k-1} (-1)^i \binom{n}{i} ((k-1)-i)^n \\ &= \frac{1}{n!} \sum_{i=0}^k (-1)^i \binom{n}{i} (k-i)^n - \frac{1}{n!} \sum_{i=1}^k (-1)^{i-1} \binom{n}{i-1} (k-i)^n \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n!} \sum_{i=0}^k (-1)^i \binom{n}{i} (k-i)^n + \frac{1}{n!} \sum_{i=0}^k (-1)^i \binom{n}{i-1} (k-i)^n \\
&= \frac{1}{n!} \sum_{i=0}^k (-1)^i \binom{n+1}{i} (k-i)^n \\
&= \frac{1}{n!} \sum_{i=0}^{k-1} (-1)^i \binom{n+1}{i} (k-i)^n \\
&= \frac{1}{n!} \mathcal{A}_{n,k}. \tag{2.3.5}
\end{aligned}$$

□

Ist M eine Markov-Kette und φ eine auf dem Zustandsraum von M definierte Funktion, so überträgt sich die Markov-Eigenschaft von M im Allgemeinen nicht auf $\varphi(M)$. In Satz 2.3.3 zeigen wir, dass die stochastische Folge $(R(X_m))_{m \geq 0}$, also die Folge der Sequenzanzahlen der Markov-Kette $(X_m)_{m \geq 0}$, jedoch wiederum eine Markov-Kette bildet. Wir benötigen hierzu zunächst ein hinreichendes Kriterium aus [18], das Voraussetzungen nennt, unter denen eine Funktion angewandt auf einen Markov-Prozess (in stetiger Zeit) wiederum einen Markov-Prozess bildet.

Seien S und S' endliche Mengen und $(P_t)_{t \in [0, \infty)}$ eine Familie von Wahrscheinlichkeitsmaßen auf $(S, \mathfrak{P}(S))$, $\varphi : (S, \mathfrak{P}(S)) \rightarrow (S', \mathfrak{P}(S'))$ und $B(S')$ die Menge aller reellen Funktionen auf S' . φ heißt *vollständig* für die Familie $(P_t)_{t \in [0, \infty)}$, falls für $f \in B(S')$ gilt

$$\int_{S'} f(s') P_t^\varphi(ds') = 0 \text{ für alle } t \in [0, \infty) \implies f = 0.$$

2.3.2 Lemma. Seien S, S' wie oben, $\varphi : (S, \mathfrak{P}(S)) \rightarrow (S', \mathfrak{P}(S'))$ und $(M_t)_{t \in [0, \infty)}$ ein Markov-Prozess in stetiger Zeit mit Zustandsraum S und Anfangsverteilung λ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$. Es gelte ferner:

- (i) Für alle $t \in [0, \infty)$ und $y \in S'$ ist die regulär bedingte Verteilung von M_t gegeben $\varphi(M_t) = y$ unabhängig von t .
- (ii) φ ist vollständig für $(P^{M_t})_{t \in [0, \infty)}$.

Dann ist $(\varphi(M_t))_{t \in [0, \infty)}$ ein Markov-Prozess mit Anfangsverteilung $P^{\varphi(M_0)}$.

Um Lemma 2.3.2 auf unsere Situation eines $Q_{a,n}$ -Random Walks X mit Folge der Sequenzanzahlen $R(X)$ anwenden zu können, betten wir die stochastische Folge $R(X) = (R(X_m))_{m \geq 0}$ in einen stochastischen Prozess $(R(M_t))_{t \in [0, \infty)}$ in stetiger Zeit ein. Dazu sei $(N_t)_{t \in [0, \infty)}$ ein von $(X_m)_{m \geq 0}$ stochastisch unabhängiger und zeitlich homogener Poisson-Prozess. Für $t > 0$ ist N_t dann eine mit dem Parameter t Poisson-verteilte Zufallsgröße, das heißt für $m \in \mathbb{N}_0$ gilt

$$P(N_t = m) = Poi(t)(\{m\}) = e^{-t} \frac{t^m}{m!}. \quad (2.3.6)$$

Für $t = 0$ ist $N_0 = \delta_0$. Wir definieren den stochastischen Prozess $M = (M_t)_{t \in [0, \infty)}$ durch

$$M_t \stackrel{\text{def}}{=} X_{N_t}, \quad t \in [0, \infty).$$

Nach Abschnitt 1.3 in [21] ist M dann ein zeitlich homogener Markov-Prozess in stetiger Zeit. Die zeitliche Homogenität von M folgt dabei aus der zeitlichen Homogenität von X . Zu einem Zeitparameter $t \in [0, \infty)$ befindet sich der Prozess in dem Zustand, den die Markov-Kette X zum Zeitpunkt N_t annimmt. Wir betrachten nun die zu M gehörige Folge der Sequenzanzahlen $R(M) = (R(M_t))_{t \in [0, \infty)}$. Wegen $R(\mathcal{S}_n) = \mathbb{N}_{\leq n}$ ist der stochastische Prozess $R(M)$ $\mathbb{N}_{\leq n}$ -wertig. Für $t \in [0, \infty)$ und $r = 1, \dots, n$ gilt wegen der stochastischen Unabhängigkeit von $(R(X_m))_{m \geq 0}$ und $(N_t)_{t \in [0, \infty)}$ und (2.3.1)

$$\begin{aligned} P(R(M_t) = r) &= P(R(X_{N_t}) = r) \\ &= P\left(\bigcup_{m \geq 0} \{N_t = m, R(X_m) = r\}\right) \\ &= \sum_{m \geq 0} P(N_t = m) P(R(X_m) = r) \\ &= \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} P(R(X_m) = r) \\ &= \mathcal{A}_{n,r} \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} \cdot \frac{1}{a^{mn}} \binom{n + a^m - r}{n}. \end{aligned} \quad (2.3.7)$$

Für $t = 0$ erhalten wir wegen $N_0 = \delta_0$, $X_0 = \text{id}$ P -f.s. und $R(\text{id}) = 1$ speziell

$$P(R(M_0) = r) = P(R(X_0) = r) = \delta_1(\{r\}). \quad (2.3.8)$$

Wir berechnen nun für $t \in [0, \infty)$ und $r = 1, \dots, n$ die regulär bedingte Verteilung von M_t gegeben $R(M_t) = r$. Sei $\pi \in \mathcal{S}_n$ beliebig. Dann erhalten wir analog zur Rechnung in (2.3.7)

$$\begin{aligned} P(M_t = \pi, R(M_t) = r) &= \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} P(X_m = \pi, R(X_m) = r) \\ &= \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} P(X_m = \pi) P(R(\pi) = r) \\ &= \delta_r(\{R(\pi)\}) \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} \cdot \frac{1}{a^{mn}} \binom{n + a^m - r}{n}. \end{aligned}$$

Wegen $\delta_r(\{R(\pi)\}) = \delta_{\mathcal{R}_{n,r}}(\{\pi\})$ gilt dann

$$P(M_t = \pi, R(M_t) = r) = \delta_{\mathcal{R}_{n,r}}(\{\pi\}) \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} \cdot \frac{1}{a^{mn}} \binom{n + a^m - r}{n}. \quad (2.3.9)$$

Unter Benutzung von (2.3.7) und (2.3.9) ergibt sich

$$P(M_t = \pi \mid R(M_t) = r) = \frac{1}{\mathcal{A}_{n,r}} \delta_{\mathcal{R}_{n,r}}(\{\pi\}), \quad (2.3.10)$$

das heißt die regulär bedingte Verteilung von M_t gegeben $R(M_t) = r$ ist unabhängig von t . Wir kommen nun zu dem bereits angekündigten Resultat.

2.3.3 Satz. *Sei $n \in \mathbb{N}$, $a \in \mathbb{N}_{\geq 2}$ und $X = (X_m)_{m \geq 0}$ ein $Q_{a,n}$ -Random Walk. Dann ist $R(X) = (R(X_m)_{m \geq 0})$ eine zeitlich homogene Markov-Kette mit Zustandsraum $\mathbb{N}_{\leq n}$ und Anfangsverteilung $P^{R(X_0)} = \delta_1$.*

Beweis. Sei $n \in \mathbb{N}$ und $a \in \mathbb{N}_{\geq 2}$ und M der oben zu X konstruierte Markov-Prozess in stetiger Zeit. Mit Lemma 2.3.2 zeigen wir, dass $R(M)$ eine Markov-Kette mit Anfangsverteilung $P^{R(M_0)}$ ist. Nach (2.3.10) ist für $t \in [0, \infty)$ und $r = 1, \dots, n$ die regulär bedingte Verteilung von M_t gegeben $R(M_t) = r$ unabhängig von t .

Wir zeigen nun die Vollständigkeit von R für $(P^{M_t})_{t \in [0, \infty)}$. Sei $f \in B(\mathbb{N}_{\leq n})$ mit $\int f dP^{R(M_t)} = 0$ für alle $t \in [0, \infty)$. Nach (2.3.7) gilt wegen der Beschränktheit von f

$$\begin{aligned} \int f dP^{R(M_t)} &= \sum_{r=1}^n \sum_{m \geq 0} e^{-t} \frac{t^m}{m!} \cdot \frac{\mathcal{A}_{n,r}}{a^{mn}} \binom{n + a^m - r}{n} f(r) \\ &= e^{-t} \sum_{m \geq 0} \frac{t^m}{m!} \left(\sum_{r=1}^n \frac{\mathcal{A}_{n,r}}{a^{mn}} \binom{n + a^m - r}{n} f(r) \right) = 0 \end{aligned}$$

für alle $t \in [0, \infty)$. Aus dem Identitätssatz für Potenzreihen folgt dann für alle $m \geq 0$

$$\sum_{r=1}^n \frac{\mathcal{A}_{n,r}}{a^{mn}} \binom{n+a^m-r}{n} f(r) = 0$$

und somit

$$\sum_{r=1}^n \mathcal{A}_{n,r} \binom{n+a^m-r}{n} f(r) = 0. \quad (2.3.11)$$

Wir definieren die Funktion $g : \mathbb{R} \longrightarrow \mathbb{R}$ für $x \in \mathbb{R}$ durch

$$\begin{aligned} g(x) &\stackrel{\text{def}}{=} \sum_{r=1}^n \mathcal{A}_{n,r} \binom{n+x-r}{n} n! f(r) \\ &= \sum_{r=1}^n \mathcal{A}_{n,r} ((x-r+1) \cdot \dots \cdot (x-r+n)) f(r) \\ &= \mathcal{A}_{n,1} x(x+1) \cdot \dots \cdot (x-1+n) f(1) \\ &\quad + \mathcal{A}_{n,2} (x-1)x \cdot \dots \cdot (x-2+n) f(2) \\ &\quad \vdots \\ &\quad + \mathcal{A}_{n,n} (x-n+1) \cdot \dots \cdot (x-1) x f(n), \end{aligned} \quad (2.3.12)$$

wobei die oben auftretenden Binomialkoeffizienten an dieser Stelle für alle $x \in \mathbb{R}$ definiert seien durch

$$\binom{x}{n} \stackrel{\text{def}}{=} \frac{x(x-1) \cdot \dots \cdot (x-n+1)}{1 \cdot 2 \cdot \dots \cdot n}.$$

Dann ist g ein reelles Polynom vom Grad $\leq n$. Nach (2.3.11) gilt $g(a^m) = 0$ für alle $m \geq 0$. Wegen $a \geq 2$ hat g also unendlich viele Nullstellen und es folgt $g = 0$. Mit $x = 1, \dots, n$ erhalten wir dann aus (2.3.12)

$$f(1) = f(2) = \dots = f(n) = 0.$$

Dies zeigt die Vollständigkeit von R für $(P^{M_t})_{t \in [0, \infty)}$. Nach Lemma 2.3.2 ist $R(M)$ also ein Markov-Prozess mit Anfangsverteilung $P^{R(M_0)}$, wobei nach (2.3.8) $P^{R(M_0)} = \delta_1$ gilt. Da X zeitlich homogen ist, gilt dasselbe für M und somit für

$R(M)$. Nach Kapitel 8 in [5] ist dann auch die in den Markov-Prozess $R(M)$ „eingebettete“ stochastische Folge $R(X) = (R(X_m))_{m \geq 0}$ eine zeitlich homogene Markov-Kette. $R(X)$ hat so wie $R(M)$ den Zustandsraum $\mathbb{N}_{\leq n}$ und wegen $P^{R(X_0)} = P^{R(M_0)} = \delta_1$ die Anfangsverteilung δ_1 . \square

2.4 Konvergenz gegen die Gleichverteilung: Der Cut-Off-Effekt

Nach Satz 2.2.8 konvergiert ein $Q_{a,n}$ -Random Walk auf \mathcal{S}_n in Totalvariation gegen die Gleichverteilung $U_{\mathcal{S}_n}$. Wir untersuchen im Folgenden den GSR-(2, n)-Shuffle und leiten Aussagen über die Konvergenzrate her. In diesem Abschnitt sei $a = 2$, $Q_n \stackrel{\text{def}}{=} Q_{2,n}$ und $X = (X_m)_{m \geq 0}$ ein Q_n -Random Walk.

Die computergestützte Berechnung des Variationabstands $\|Q_n^{*(m)} - U_{\mathcal{S}_n}\|$, also der „Entfernung“ eines m -mal nach dem GSR-(2, n)-Shuffle gemischten Kartenstapels von der Gleichverteilung, mit Hilfe der Formel

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| = \frac{1}{2} \sum_{\pi \in \mathcal{S}_n} \left| Q_n^{*(m)}(\{\pi\}) - \frac{1}{n!} \right|$$

ist schon bei üblichen Kartenstapelgrößen auf Grund der Anzahl von $n!$ zu berechnenden Termen nicht möglich. So sind beispielsweise für einen Skat-Kartenstapel $32!$ Terme zu berechnen, für einen Blackjack-, Bridge- oder Schafkopf-Kartenstapel $52!$ und Rommé-Spieler sehen sich der Anzahl von $104!$ Termen ausgeliefert.

Mit Hilfe der aufsteigenden Sequenzen lässt sich der Rechenaufwand jedoch drastisch auf die Berechnung von n Termen reduzieren, denn für $m \geq 0$ erhalten wir mit Korollar 2.2.6 (ii)

$$\begin{aligned} \|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &= \frac{1}{2} \sum_{\pi \in \mathcal{S}_n} \left| Q_n^{*(m)}(\{\pi\}) - \frac{1}{n!} \right| \\ &= \frac{1}{2} \sum_{\pi \in \mathcal{S}_n} \left| \binom{n + 2^m - R(\pi)}{n} \frac{1}{2^{mn}} - \frac{1}{n!} \right| \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \sum_{r=1}^n \sum_{\pi \in \mathcal{R}_{n,r}} \left| \binom{n+2^m-r}{n} \frac{1}{2^{mn}} - \frac{1}{n!} \right| \\
&= \frac{1}{2} \sum_{r=1}^n \mathcal{A}_{n,r} \left| \binom{n+2^m-r}{n} \frac{1}{2^{mn}} - \frac{1}{n!} \right|. \tag{2.4.1}
\end{aligned}$$

Dabei bezeichnen $\mathcal{A}_{n,1}, \dots, \mathcal{A}_{n,n}$ die Eulerschen Zahlen, die durch

$$\mathcal{A}_{n,r} = \sum_{i=0}^{r-1} (-1)^i \binom{n+1}{i} (r-i)^n$$

für $r = 1, \dots, n$ gegeben sind.

In Tabelle 2.2 ist der Variationsabstand von $Q_n^{*(m)}$ und der Gleichverteilung U_{S_n} für die gebräuchlichen Kartenstapelgrößen $n = 24, 32, 48, 52, 104$ und $m = 1, \dots, 10$ Mischvorgänge aufgelistet. Der Variationsabstand wurde mit Hilfe von (2.4.1) berechnet.

	$m=1$	2	3	4	5	6	7	8	9	10
n=24	1.000	0.999	0.998	0.723	0.418	0.211	0.109	0.055	0.028	0.014
32	1.000	1.000	0.999	0.929	0.597	0.322	0.164	0.084	0.042	0.021
48	1.000	1.000	1.000	0.999	0.888	0.546	0.297	0.149	0.076	0.038
52	1.000	1.000	1.000	0.999	0.924	0.614	0.334	0.167	0.085	0.043
104	1.000	1.000	1.000	1.000	1.000	0.988	0.772	0.454	0.237	0.119

Tabelle 2.2: $\|Q_n^{*(m)} - U_{S_n}\|$ für $n = 24, 32, 48, 52, 104$ und $m = 1, \dots, 10$.

Für $n = 52$ ist der Variationsabstand von $Q_n^{*(m)}$ und U_{S_n} für $m = 1, \dots, 20$ Mischvorgänge in Abbildung 2.6 in einem Graphen abgetragen. Es wird deutlich, dass der Variationsabstand bis zu einer Anzahl von $m = 5$ Mischvorgängen nahe bei seinem maximalen Wert von 1 liegt. Innerhalb der nächsten vier Mischvorgänge nähert er sich jedoch abrupt seinem Minimum von 0 an. Diese abrupte Annäherung innerhalb eines relativ kurzen Zeitraums wird als *Cut-Off-Effekt* oder *Cut-Off-Phänomen* bezeichnet. Ein Blick auf den aktuellen Stand der Forschung offenbart, dass es sich beim Cut-Off-Effekt tatsächlich um ein Phänomen handelt: Es ist bislang nicht gelungen, anhand der Struktur des Zustandsraums und des Übergangskerns einer ergodischen Markov-Kette Bedingungen abzuleiten, unter denen

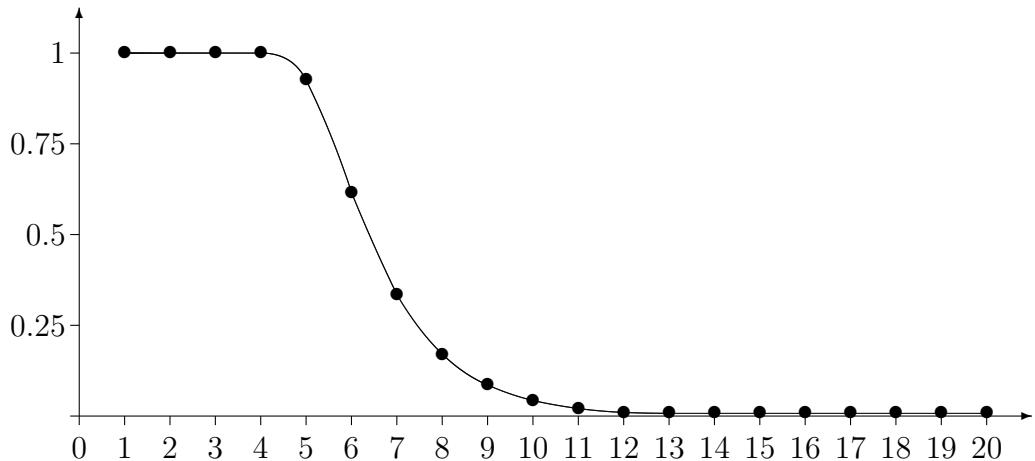


Abbildung 2.6: $m \mapsto \|Q_n^{*(m)} - U_{S_n}\|$ für $n = 52$ und $m = 1, \dots, 20$.

bei der Konvergenz gegen die stationäre Verteilung ein Cut-Off-Effekt auftritt. Vielmehr ist in jedem einzelnen Fall eine explizite Berechnung oder Approximation des Variationsabstands nötig. Diaconis [10] äußert jedoch die Vermutung, dass das Cut-Off-Phänomen vor allem bei reversiblen Markov-Ketten auftritt, bei denen der zweitgrößte Eigenwert der zugehörigen Übergangsmatrix eine hohe Vielfachheit aufweist. Wilson [24] konnte im Jahre 2004 den Cut-Off-Effekt für zahlreiche Markov-Ketten nachweisen, deren Zustandsraum als *Lozenge* interpretiert werden kann. Eine Lozenge ist ein Rhombus mit 60° und 120° Winkeln und Kanten der Länge 1. Er verwendete die Fourier-Theorie und Kopplungsargumente, um obere und untere Schranken für die Totalvariation herzuleiten. Für einige Kartenmischmodelle und Markov-Ketten konnte so der Cut-Off-Effekt nachgewiesen werden.

Wir geben nun eine präzisere Definition des Cut-Off-Effekts (siehe Abschnitt 2.4.2 in [21]). Ist G eine endliche Gruppe, so sei U_G wie im Vorangegangenen die Gleichverteilung auf G . Für $n \geq 1$ sei G_n eine endliche Gruppe. Wir bezeichnen eine Familie von Wahrscheinlichkeitsmaßen $(P_n)_{n \geq 1}$ als *ergodische Familie auf $((G_n, \mathfrak{P}(G_n)))_{n \geq 1}$* , falls für alle $n \geq 1$ P_n ein Wahrscheinlichkeitsmaß auf $(G_n, \mathfrak{P}(G_n))$ ist und der P_n -Random Walk auf G_n ergodisch, das heißt irreduzibel und aperiodisch mit stationärer Verteilung U_{G_n} ist.

2.4.1 Definition. Sei $(G_n)_{n \geq 1}$ eine Familie endlicher Gruppen und $(P_n)_{n \geq 1}$ eine

ergodische Familie von Wahrscheinlichkeitsmaßen auf $((G_n, \mathfrak{P}(G_n)))_{n \geq 1}$. Die Familie $(P_n)_{n \geq 1}$ besitzt einen *Cut-Off-Effekt (in Totalvariation) mit Cut-Off-Zeiten* $(t_n)_{n \geq 1}$, falls folgende Bedingungen erfüllt sind:

- (i) Es gilt $t_n \geq 0$ für alle $n \geq 1$ und $\lim_{n \rightarrow \infty} t_n = \infty$.
- (ii) Für alle $\varepsilon \in (0, 1)$ und $m_n = \lfloor (1 + \varepsilon)t_n \rfloor$ gilt $\lim_{n \rightarrow \infty} \|P_n^{*(m_n)} - U_{G_n}\| = 0$.
- (iii) Für alle $\varepsilon \in (0, 1)$ und $m_n = \lfloor (1 - \varepsilon)t_n \rfloor$ gilt $\lim_{n \rightarrow \infty} \|P_n^{*(m_n)} - U_{G_n}\| = 1$.

Besitzt eine Familie $(P_n)_{n \geq 1}$ einen Cut-Off-Effekt mit Cut-Off-Zeiten $(t_n)_{n \geq 1}$, so bedeutet dies, dass die Funktion $m \mapsto \|P_n^{*(m)} - U_{G_n}\|$ bis zu einem Zeitpunkt „vor“ $\lfloor t_n \rfloor$ nah bei ihrem maximalen Wert von 1 bleibt. Nach dem Zeitpunkt $\lfloor t_n \rfloor$ fällt sie rasch auf einen Wert nahe 0. Dieser „Phasen-Übergang“ vollzieht sich dabei innerhalb eines im Vergleich zu $\lfloor t_n \rfloor$ relativ kurzen Zeitraums der Größenordnung $o(t_n)$.

Aus Abschnitt 2.4.2 in [21] bzw. Abschnitt 1 in [10] erhalten wir eine weitere und zugleich striktere Definition des Cut-Off-Effekts.

2.4.2 Definition. Sei $(G_n)_{n \geq 1}$ eine Familie endlicher Gruppen und $(P_n)_{n \geq 1}$ eine ergodische Familie von Wahrscheinlichkeitsmaßen auf $((G_n, \mathfrak{P}(G_n)))_{n \geq 1}$. Die Familie $(P_n)_{n \geq 1}$ besitzt einen *starken Cut-Off-Effekt (in Totalvariation) mit Cut-Off-Zeiten* $((t_n, b_n))_{n \geq 1}$, falls folgende Bedingungen erfüllt sind:

- (i) Es gilt $t_n \geq 0$, $b_n > 0$ für alle $n \geq 1$, $\lim_{n \rightarrow \infty} t_n = \infty$ und $\lim_{n \rightarrow \infty} b_n/t_n = 0$.
- (ii) Für alle $c \in \mathbb{R}$ und $m_n \stackrel{\text{def}}{=} \lfloor t_n + cb_n \rfloor$ gilt $\lim_{n \rightarrow \infty} \|P_n^{*(m_n)} - U_{G_n}\| = f(c)$ für eine reelle Funktion f mit $\lim_{c \rightarrow \infty} f(c) = 0$ und $\lim_{c \rightarrow -\infty} f(c) = 1$.

In der obigen Definition kann $c \in \mathbb{R}$ auch durch $j \in \mathbb{Z}$ ersetzt werden. Aus der Definition von starkem Cut-Off-Effekt und Cut-Off-Effekt folgt unmittelbar, da $\|P_n^{*(m)} - U_{G_n}\|$ monoton fallend in m ist (siehe Lemma 3.5 in [1]), dass ein starker Cut-Off-Effekt mit Cut-Off-Zeiten $((t_n, b_n))_{n \geq 1}$ einen Cut-Off-Effekt mit Cut-Off-Zeiten $(t_n)_{n \geq 1}$ impliziert:

2.4.3 Bemerkung. Sei $(G_n)_{n \geq 1}$ eine Familie endlicher Gruppen und $(P_n)_{n \geq 1}$ eine ergodische Familie von Wahrscheinlichkeitsmaßen auf $((G_n, \mathfrak{P}(G_n)))_{n \geq 1}$. Falls $(P_n)_{n \geq 1}$ einen starken Cut-Off-Effekt mit Cut-Off-Zeiten $((t_n, b_n))_{n \geq 1}$ besitzt, so besitzt $(P_n)_{n \geq 1}$ einen Cut-Off-Effekt mit Cut-Off-Zeiten $(t_n)_{n \geq 1}$.

Beweis. Sei $((t_n, b_n))_{n \geq 1}$ wie in Definition 2.4.2 und $\varepsilon \in (0, 1)$. Wir zeigen

$$\lim_{n \rightarrow \infty} \|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_G\| = 0.$$

Sei $c > 0$ beliebig. Wegen $\lim_{n \rightarrow \infty} b_n/t_n = 0$ und $\lim_{n \rightarrow \infty} t_n = \infty$ existiert ein $N = N_{\varepsilon, c} \in \mathbb{N}$ mit $\varepsilon t_n > cb_n$ für alle $n \geq N$. Da $\|P_n^{*(m)} - U_{G_n}\|$ monoton fallend in m ist (siehe Lemma 3.5 in [1]), gilt für alle $n \geq N$ wegen $\lfloor(1+\varepsilon)t_n\rfloor = \lfloor t_n + \varepsilon t_n \rfloor \geq \lfloor t_n + cb_n \rfloor$

$$\|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_{G_n}\| \leq \|P_n^{*(\lfloor t_n + cb_n \rfloor)} - U_{G_n}\|$$

und somit

$$\limsup_{n \rightarrow \infty} \|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_{G_n}\| \leq \lim_{n \rightarrow \infty} \|P_n^{*(\lfloor t_n + cb_n \rfloor)} - U_{G_n}\| = f(c).$$

Da $c > 0$ beliebig gewählt war, erhalten wir vermöge $\lim_{c \rightarrow \infty} f(c) = 0$ und $0 \leq f \leq 1$

$$\limsup_{n \rightarrow \infty} \|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_{G_n}\| \leq \inf_{c \in \mathbb{R}} \lim_{n \rightarrow \infty} \|P_n^{*(\lfloor t_n + cb_n \rfloor)} - U_{G_n}\| = \inf_{c \in \mathbb{R}} f(c) = 0.$$

Mit $\|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_{G_n}\| \geq 0$ folgt insgesamt

$$\lim_{n \rightarrow \infty} \|P_n^{*(\lfloor(1+\varepsilon)t_n\rfloor)} - U_{G_n}\| = 0.$$

Analog erhalten wir

$$\lim_{n \rightarrow \infty} \|P_n^{*(\lfloor(1-\varepsilon)t_n\rfloor)} - U_{G_n}\| = 1. \quad \square$$

2.5 Obere und untere Schranke für die Totalvariation

Wir werden im nächsten Abschnitt zeigen, dass die Familie $(Q_n)_{n \geq 1}$ der GSR- $(2, n)$ -Mischmethoden, $n \geq 1$, einen starken Cut-Off-Effekt mit Cut-Off-Zeiten

$((t_n, b_n)) = ((\frac{3}{2} \log_2 n, 1))_{n \geq 1}$ besitzt, wobei \log_2 den Logarithmus zur Basis 2 bezeichne. Mit Blick auf Definition 2.4.2 bedeutet dies, dass für große Kartenstapelgrößen n eine Anzahl von $\lfloor \frac{3}{2} \log_2 n \rfloor$ Mischvorgängen hinreichend und notwendig ist, um den Kartenstapel „gut durchzumischen“.

Die folgenden oberen und unteren Schranken für die Totalvariation $\|Q_n^{*(m)} - U_{\mathcal{S}_n}\|$ nach m -maligem Mischen nach dem GSR-(2, n)-Shuffle gehen auf Reeds [17] zurück und finden sich, abgesehen von der Asymptotik für die untere Schranke, auch in [2], [9] und [16]. Sie gehören zu den zeitlich frühesten Ergebnissen zum Riffle Shuffle und sind deswegen an dieser Stelle aufgeführt. Im Vergleich zu den Resultaten von Bayer und Diaconis [6] besitzen sie jedoch nur geringe Aussagekraft.

2.5.1 Satz (Untere Schranke). *Seien $m, n \in \mathbb{N}$. Dann gilt*

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| \geq 1 - \frac{1}{n!} \sum_{r=1}^{2^m \wedge n} \mathcal{A}_{n,r} \underset{n \rightarrow \infty}{\sim} 1 - \Phi \left(\frac{(2^m \wedge n) - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \right),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung sei.

Beweis. Seien $m, n \in \mathbb{N}$. Für alle $A \subset \mathcal{S}_n$ gilt

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| \geq |Q_n^{*(m)}(A) - U_{\mathcal{S}_n}(A)|.$$

Wählen wir $A = \text{supp}(Q_n^{*(m)})$, so folgt mit Korollar 2.2.7

$$\begin{aligned} \|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &\geq 1 - \frac{|\text{supp}(Q_n^{*(m)})|}{n!} \\ &= 1 - \frac{|\{\pi \in \mathcal{S}_n \mid R(\pi) \leq 2^m \wedge m\}|}{n!} \\ &= 1 - \sum_{r=1}^{2^m \wedge n} \frac{|\{\pi \in \mathcal{S}_n \mid R(\pi) = r\}|}{n!} \\ &= 1 - \frac{1}{n!} \sum_{r=1}^{2^m \wedge n} \mathcal{A}_{n,r}. \end{aligned}$$

Den Nachweis für die asymptotische Aussage $1 - \frac{1}{n!} \sum_{r=1}^{2^m \wedge n} \mathcal{A}_{n,r} \sim 1 - \Phi \left(\frac{(2^m \wedge n) - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \right)$ für $n \rightarrow \infty$ erbringen wir an späterer Stelle und zwar im Beweis von Lemma

2.6.6. Ersetzen wir dort in (2.6.20) h^* durch $(2^m \wedge n) - n/2$, so erhalten wir die Behauptung. \square

Für den Beweis der oberen Schranke wird eine *stark stationäre Stoppzeit* für den Q_n -Random Walk konstruiert. Für die allgemeine Definition der Stoppzeit und der stark stationären Stoppzeit verweisen wir auf die Abschnitte 3 und 18 in [4]. In der Situation des Q_n -Random Walks auf \mathcal{S}_n reicht die folgende Definition aus.

2.5.2 Definition. Sei $X = (X_m)_{m \geq 0}$ ein Q_n -Random Walk auf einem messbaren Raum $(\Omega, \mathfrak{A}, P)$ (mit eindeutig bestimmter stationärer Verteilung $U_{\mathcal{S}_n}$). Eine messbare Abbildung $T : \Omega \longrightarrow \mathbb{N}_0 \cup \{\infty\}$ mit $\{T = m\} \in \sigma(X_0, \dots, X_m)$ für $m \geq 0$ heißt *stark stationäre Stoppzeit*, wenn sie die folgenden zwei Bedingungen erfüllt:

- (i) $P(T < \infty) = 1$,
- (ii) $P(X_m = \pi \mid T = m) = U_{\mathcal{S}_n}(\{\pi\})$ für alle $m \geq 0$ und $\pi \in \mathcal{S}_n$.

2.5.3 Bemerkung. Sei T eine stark stationäre Stoppzeit für den Q_n -Random Walk X . Dann gilt für alle $m \geq 0$

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| \leq P(T > m).$$

Beweis. Sei $m \geq 0$ und $A \subset \mathcal{S}_n$ beliebig. Dann gilt nach Definition 2.5.2 (ii)

$$\begin{aligned} P(X_m \in A) &= P(X_m \in A, T \leq m) + P(X_m \in A, T > m) \\ &= P(X_m \in A \mid T \leq m)P(T \leq m) + P(X_m \in A \mid T > m)P(T > m) \\ &= U_{\mathcal{S}_n}(A)(1 - P(T > m)) + P(X_m \in A \mid T > m)P(T > m) \\ &= U_{\mathcal{S}_n}(A) + (P(X_m \in A \mid T > m) - U_{\mathcal{S}_n}(A))P(T > m). \end{aligned}$$

Daraus folgt unmittelbar

$$\begin{aligned} \|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &= \|P^{X_m} - U_{\mathcal{S}_n}\| \\ &= \sup_{A \subset \mathcal{S}_n} \|P^{X_m}(A) - U_{\mathcal{S}_n}(A)\| \\ &= \sup_{A \subset \mathcal{S}_n} P(T > m)|P(X_m \in A \mid T > m) - U_{\mathcal{S}_n}(A)| \\ &\leq P(T > m). \end{aligned}$$

\square

Die folgende obere Schranke für die Totalvariation $\|Q_n^{*(m)} - U_{\mathcal{S}_n}\|$ ist aus dem „Geburtstagsproblem“ bekannt. Aufgrund des geringen Nutzens der oberen Schranke im Vergleich zu unseren späteren Ergebnissen, verzichten wir auf einen detaillierten Beweis.

2.5.4 Satz (Obere Schranke). *Seien $m, n \in \mathbb{N}$. Dann gilt*

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| \leq 1 - \prod_{i=1}^n \left(1 - \frac{i}{2^m}\right).$$

Beweis. Wir werden für den Beweis das inverse Modell benutzt, jedoch mit einer leichten Modifikation: Im inversen Modell (siehe (2.1.4)) haben wir die Mischmethode $Q_{a,n}^I$ für $B \subset \mathcal{S}_n$ durch

$$Q_{a,n}^I(B) = U_{\mathbb{N}_{0,a-1}^n}(X_{a,n}^I = \pi^{-1}, \pi \in B)$$

definiert. Wir definieren wir nun die Mischmethode $\overline{Q}_{a,n}^I$ für $B \subset \mathcal{S}_n$ durch

$$\overline{Q}_{a,n}^I(B) \stackrel{\text{def}}{=} U_{\mathbb{N}_{0,a-1}^n}(X_{a,n}^I = \pi, \pi \in B).$$

und $\overline{Q}_n \stackrel{\text{def}}{=} \overline{Q}_{2,n}^I$. Dann gilt mit $\mathcal{I} : \mathcal{S}_n \longrightarrow \mathcal{S}_n$ definiert durch $\mathcal{I}(\pi) = \pi^{-1}$, $\pi \in \mathcal{S}_n$, wegen $(\overline{Q}_n^I)^{\mathcal{I}} = Q_n^I$ unter Beachtung von (1.1.1) aus Symmetriegründen

$$\begin{aligned} \|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &= \|(Q_{2,n}^I)^{*(m)} - U_{\mathcal{S}_n}\| \\ &= \|(\overline{Q}_{2,n}^I)^{*(m)} - U_{\mathcal{S}_n}\| \\ &= \|\overline{Q}_n^{*(m)} - U_{\mathcal{S}_n}\| \end{aligned}$$

für alle $m \geq 0$. Wir können also für die Abschätzung der Totalvariation auch das inverse Modell in obiger Form heranziehen. Wir konstruieren nun die stark stationäre Stoppzeit T und erinnern daran, dass sich die Karten vor dem Mischen in der Reihenfolge $(1, \dots, n)$ befinden. Im inversen Modell in obiger Form können wir nun die sukzessiven Mischvorgänge wie folgt abbilden. Für den m -ten Mischvorgang wählen wir ein n -Tupel $x^{(m)} = (x_1^{(m)}, \dots, x_n^{(m)})$ als Realisation einer auf $\{0, 1\}^n$ Laplace-verteilten Zufallsgröße $X^{(m)}$, markieren Karte i mit dem Wert $x_i^{(m)}$ und mischen die Karten gemäß der in (2.1.3) definierten Permutation $\pi_{x^{(m)}-}$. Wir markieren dabei die Werte in den einzelnen Mischvorgängen von links nach rechts auf jeder

einzelnen Karte. Die Zufallsgrößen $X^{(1)}, X^{(2)}, \dots$ seien stochastisch unabhängig. Die Werte auf Karte i nach dem m -ten Mischvorgang sind dann gegeben durch das m -Tupel $y_i^{(m)} = (x_i^{(1)}, \dots, x_i^{(m)})$. Solange die Werte $y_i^{(k)}$ und $y_j^{(k)}$ für $k \leq m$ auf zwei verschiedenen Karten i und j identisch sind, befinden sich die Karten in derselben relativen Reihenfolge wie zu Beginn des ersten Mischvorgangs, da sie bei jedem einzelnen Mischvorgang im selben Päckchen liegen. Wir wählen T als den ersten Zeitpunkt m , zu dem die Tupel $y_1^{(m)}, \dots, y_n^{(m)} \in \{0, 1\}^m$ bzw. die stochastisch unabhängigen und auf $\{0, 1\}^m$ Laplace-verteilten Zufallsgrößen $Y_1^{(m)}, \dots, Y_n^{(m)}$ paarweise verschieden sind. Dann ist T f.s. endlich und eine streng stationäre Stoppzeit: Gegeben $T = m$ sind alle Reihenfolgen $X_m = (\pi(1), \dots, \pi(n))$, $\pi \in \mathcal{S}_n$ des Kartentapels nach m Mischvorgängen gleich wahrscheinlich, da für zwei beliebige Karten i, j aus Symmetriegründen $\pi(i) > \pi(j)$ und $\pi(i) < \pi(j)$ gleich wahrscheinlich sind. Die Wahrscheinlichkeit $P(T > m)$, dass die Zufallsgrößen $Y_1^{(m)}, \dots, Y_n^{(m)}$ nicht paarweise verschieden sind, kann dann analog zum „Geburtagsproblem“ berechnet werden: Es gibt 2^m mögliche $\{0, 1\}^m$ -wertige m -Tupel (Geburtstage), von diesen müssen bei n zufällig ausgewählten Tupeln (Personen) mindestens zwei übereinstimmen. Für die gesuchte Wahrscheinlichkeit gilt dann

$$P(T > m) = 1 - \prod_{i=1}^{n-1} \left(1 - \frac{i}{2^m}\right). \quad \square$$

Aldous und Diaconis zeigen in [2], dass mit $k_n \stackrel{\text{def}}{=} \lfloor 2 \log_2 n \rfloor$

$$\|Q_n^{*(k_n)} - U_{\mathcal{S}_n}\| \leq P(T > k_n) \xrightarrow{n \rightarrow \infty} 1 - e^{-1/2} \approx 0,3935$$

gilt. Sei $\lceil x \rceil$ die obere Gauss-Klammer von $x \in \mathbb{R}$. Für die untere Schranke der Totalvariation aus Satz 2.5.1 gilt bereits mit $m_n \stackrel{\text{def}}{=} \lceil \log_2 n \rceil$ wegen $2^{m_n} \geq n$ und $\sum_{r=1}^n \mathcal{A}_{n,r} = n!$

$$\|Q_n^{*(m_n)} - U_{\mathcal{S}_n}\| \geq 1 - \frac{1}{n!} \sum_{r=1}^{2^{m_n} \wedge n} \mathcal{A}_{n,r} = 0.$$

Es zeigt sich, dass die obere und untere Schranke für die Feststellung des Cut-Off-Effekts ungeeignet sind. In Tabelle 2.3 sind die Werte der Totalvariation $\|Q_n^{*(m)} - U_{\mathcal{S}_n}\|$ und der oberen und unteren Schranke für $n = 52$ und $m = 1, \dots, 10$ aufgelistet.

m	1	2	3	4	5	6	7	8	9	10
obere	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.931	0.732
52	1.000	1.000	1.000	1.000	0.924	0.614	0.334	0.167	0.085	0.043
untere	1.000	1.000	1.000	0.999	0.002	0.000	0.000	0.000	0.000	0.000

Tabelle 2.3: Obere und untere Schranke für $\|Q_{52}^{*(m)} - U_{S_{52}}\|$ und $m = 1, \dots, 10$.

2.6 Nachweis des Cut-Off-Effekts für den Riffle Shuffle

Wir zeigen in diesem Abschnitt den starken Cut-Off-Effekt mit Cut-Off-Zeiten $((\frac{3}{2} \log_2 n, 1))_{n \geq 1}$ für die Familie $(Q_n)_{n \geq 1}$ der GSR-(2, n)-Mischmethoden, wobei \log_2 den Logarithmus zur Basis 2 bezeichne. Bevor wir den Cut-Off-Effekt am Ende dieses Abschnitts nachweisen können, benötigen wir einige Hilfsmittel. Zunächst geben wir eine Approximation für die in (2.2.6) bestimmte Wahrscheinlichkeit an, dass m sukzessive GSR-(2, n)-Shuffle in der Permutation π resultieren.

Ist m eine beliebige Anzahl von Mischvorgängen eines Kartenstapels mit n Karten, so können wir m darstellen als

$$m = \lfloor \log_2(n^{3/2}) \rfloor + j$$

mit einer ganzen Zahl $j \geq -\lfloor \log_2(n^{3/2}) \rfloor$. j beschreibt also die Anzahl der Mischvorgänge vor bzw. nach $\lfloor \log_2(n^{3/2}) \rfloor$. Wählen wir nun $\theta_n \in [0, 1)$ mit

$$\lfloor \log_2(n^{3/2}) \rfloor + \theta_n = \log_2(n^{3/2})$$

und $c \stackrel{\text{def}}{=} 2^j$, so erhalten wir

$$\begin{aligned} m &= \lfloor \log_2(n^{3/2}) \rfloor + j \\ &= \log_2(n^{3/2}) - \theta_n + j \\ &= \log_2(n^{3/2}2^j) - \theta_n \\ &= \log_2(n^{3/2}c) - \theta_n. \end{aligned}$$

Wir können nun eine Umparametrisierung der Anzahl der Mischvorgänge vornehmen, indem wir sie beschreiben durch die Menge

$$\mathfrak{Z} \stackrel{\text{def}}{=} \{c = 2^j \mid j \in \mathbb{Z}\}.$$

Für $m = \log_2(n^{3/2}c) - \theta_n < 0$ setzen wir $Q_n^{*(m)} = Q_n^{*(0)} = \delta_{\text{id}}$.

2.6.1 Satz. Sei $n \geq 1$, $c \in \mathfrak{Z}$, $m = m_{n,c} = \lfloor \log_2(n^{3/2}c) \rfloor$, $\pi \in \mathcal{S}_n$ mit $R(\pi) = r \leq 2^m \wedge n$ und $h_{n,r} \in \mathbb{R}$ gegeben durch $r = \frac{n}{2} + h_{n,r}$. Dann gilt

$$Q_n^{*(m)}(\{\pi\}) = \frac{1}{n!} \exp \left(\frac{1}{c\sqrt{n}} \left(-h_{n,r} + \frac{1}{2} + \mathbf{f}_n \right) - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 + \mathbf{g}_n \right), \quad (2.6.1)$$

wobei $\mathbf{f}_n \in O_c\left(\frac{h_{n,r}}{n}\right)$ und $\mathbf{g}_n \in O_c\left(\frac{1}{n}\right)$.

Beweis. Seien zunächst $n \in \mathbb{N}$ und $c \in \mathfrak{Z}$ fest und $r \leq 2^m \wedge n$. Wir wählen $\theta_n \in [0, 1)$ wie oben mit $m = \log_2(n^{3/2}c) - \theta_n$. Nach (2.2.6) erhalten wir zusammen mit $r = \frac{n}{2} + h_{n,r}$ und $2^m = cn^{3/2}2^{-\theta_n}$

$$\begin{aligned} Q_n^{*(m)}(\{\pi\}) &= \binom{n+2^m-R(\pi)}{n} \frac{1}{2^{mn}} \\ &= \frac{1}{n!} \left(\frac{2^m+n-r}{2^m} \dots \frac{2^m+1-r}{2^m} \right) \\ &= \frac{1}{n!} \exp \left(\sum_{i=1}^n \log \left(1 + \frac{i-r}{2^m} \right) \right) \\ &= \frac{1}{n!} \exp \left(\sum_{i=0}^{n-1} \log \left(1 + \frac{n-i-r}{2^m} \right) \right) \\ &= \frac{1}{n!} \exp \left(\sum_{i=0}^{n-1} \log \left(1 + \frac{(n/2) - h_{n,r} - i}{cn^{3/2}2^{-\theta_n}} \right) \right). \end{aligned} \quad (2.6.2)$$

Mittels einer Taylor-Entwicklung erhalten wir für alle $x \in (-\frac{1}{2}, 1)$

$$x - \frac{x^2}{2} + \frac{x^3}{3} - x^4 \leq \log(1+x) \leq x - \frac{x^2}{2} + \frac{x^3}{3}. \quad (2.6.3)$$

Um die obige Abschätzung auf (2.6.2) anwenden zu können, müssen wir nun $c \in \mathfrak{Z}$ so groß wählen, dass

$$-\frac{1}{2} < \frac{(n/2) - h_{n,r} - i}{cn^{3/2}2^{-\theta_n}} < 1 \quad (2.6.4)$$

für $r = 1, \dots, n$ und $i = 0, \dots, n-1$ gilt. Wegen $\theta_n \in [0, 1)$ gilt

$$\left| \frac{(n/2) - h_{n,r} - i}{cn^{3/2}2^{-\theta_n}} \right| = \left| \frac{r-i}{cn^{3/2}2^{-\theta_n}} \right| \leq \frac{2^{\theta_n}}{cn^{1/2}} < \frac{2}{cn^{1/2}}$$

für $r = 1, \dots, n$ und $i = 0, \dots, n - 1$. Wählen wir nun $c \geq 4n^{-1/2}$, so erhalten wir (2.6.4). Wir können also (2.6.3) auf (2.6.2) anwenden und erhalten wegen $\left| \frac{(n/2) - h_{n,r} - i}{cn^{3/2} 2^{-\theta_n}} \right| > \left| \frac{(n/2) - h_{n,r} - i}{cn^{3/2}} \right|$ für alle $c \geq 4n^{-1/2}$

$$\begin{aligned} Q_n^{*(m)}(\{\pi\}) &\leq \frac{1}{n!} \exp \left(\frac{1}{cn^{3/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right) - \frac{1}{2c^2 n^3} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^2 \right. \\ &\quad \left. + \frac{1}{3c^3 n^{9/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^3 \right), \end{aligned} \tag{2.6.5}$$

$$\begin{aligned} Q_n^{*(m)}(\{\pi\}) &\geq \frac{1}{n!} \exp \left(\frac{1}{cn^{3/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right) - \frac{1}{2c^2 n^3} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^2 \right. \\ &\quad \left. + \frac{1}{3c^3 n^{9/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^3 - \frac{1}{c^4 n^6} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^4 \right). \end{aligned} \tag{2.6.6}$$

Zusammen mit $\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}$, $\sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}$, $\sum_{i=1}^{n-1} i^3 = \frac{n^2(n-1)^2}{4}$ und $\sum_{i=1}^{n-1} i^4 = \frac{n(n-1)(2n-1)(3n(n-1)-1)}{30}$ ergibt sich

$$\begin{aligned} \frac{1}{cn^{3/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right) &= \frac{-h_{n,r} + \frac{1}{2}}{c\sqrt{n}}, \\ \frac{1}{2c^2 n^3} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^2 &= \frac{1}{24c^2} + \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 + \frac{1 - 6h_{n,r}}{12c^2 n^2}, \\ \frac{1}{3c^3 n^{9/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^3 &= \frac{-\frac{1}{4}n^3 h_{n,r} - nh_{n,r}^3 + \frac{1}{8}n^3 + \frac{3}{2}nh_{n,r}^2 - \frac{1}{2}nh_{n,r}}{3c^3 n^{9/2}} \\ &= \frac{1 - 2h_{n,r}}{24c^3 n^{3/2}} - \frac{2h_{n,r}^3 - 3h_{n,r}^2 + h_{n,r}}{6c^3 n^{7/2}}, \\ \frac{1}{c^4 n^6} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^4 &= \frac{3n^4 + 20n^2(6h_{n,r}^2 - 6h_{n,r} + 1) - 8}{240c^4 n^5} \\ &\quad + \frac{h_{n,r}^4 - 2h_{n,r}^3 + h_{n,r}^2}{c^4 n^5}. \end{aligned} \tag{2.6.7}$$

Definieren wir

$$\begin{aligned}\mathbf{g}_{1,n} &\stackrel{\text{def}}{=} \frac{1 - 6h_{n,r}}{12c^2n^2}, \\ \mathbf{f}_{1,n} &\stackrel{\text{def}}{=} \frac{1}{3c^3n^{9/2}} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^3, \\ \mathbf{g}_{2,n} &\stackrel{\text{def}}{=} \frac{1}{c^4n^6} \sum_{i=0}^{n-1} \left(\frac{n}{2} - h_{n,r} - i \right)^4 \text{ und} \\ \mathbf{g}_{3,n} &\stackrel{\text{def}}{=} \mathbf{g}_{1,n} + \mathbf{g}_{2,n},\end{aligned}$$

so gilt $\mathbf{f}_{1,n} \in O_c\left(\frac{h_{n,r}}{n^{3/2}}\right)$ und $\mathbf{g}_{1,n}, \mathbf{g}_{2,n}, \mathbf{g}_{3,n} \in O_c\left(\frac{1}{n}\right)$. Unter Benutzung der obigen Gleichungen folgt aus (2.6.5) und (2.6.6) mit $\mathbf{f}_n \stackrel{\text{def}}{=} c\sqrt{n}\mathbf{f}_{1,n} \in O_c\left(\frac{h_{n,r}}{n}\right)$ für alle $c \geq 4n^{-1/2}$

$$\begin{aligned}Q_n^{*(m)}(\{\pi\}) &\leq \frac{1}{n!} \exp \left(\frac{-h_{n,r} + \frac{1}{2}}{c\sqrt{n}} - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 - \mathbf{g}_{1,n} + \mathbf{f}_{1,n} \right) \\ &= \frac{1}{n!} \exp \left(\frac{1}{c\sqrt{n}} \left(-h_{n,r} + \frac{1}{2} + \mathbf{f}_n \right) - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 - \mathbf{g}_{1,n} \right), \\ Q_n^{*(m)}(\{\pi\}) &\geq \frac{1}{n!} \exp \left(\frac{-h_{n,r} + \frac{1}{2}}{c\sqrt{n}} - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 - \mathbf{g}_{3,n} + \mathbf{f}_{1,n} \right) \\ &= \frac{1}{n!} \exp \left(\frac{1}{c\sqrt{n}} \left(-h_{n,r} + \frac{1}{2} + \mathbf{f}_n \right) - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 - \mathbf{g}_{3,n} \right).\end{aligned}$$

Insgesamt ergibt sich schließlich wegen $\mathbf{g}_{1,n}, \mathbf{g}_{3,n} \in O\left(\frac{1}{n}\right)$

$$Q_n^{*(m)}(\{\pi\}) = \frac{1}{n!} \exp \left(\frac{1}{c\sqrt{n}} \left(-h_{n,r} + \frac{1}{2} + \mathbf{f}_n \right) - \frac{1}{24c^2} - \frac{1}{2} \left(\frac{h_{n,r}}{cn} \right)^2 + \mathbf{g}_n \right)$$

mit $\mathbf{f}_n \in O_c\left(\frac{h_{n,r}}{n}\right)$ und $\mathbf{g}_n \in O_c\left(\frac{1}{n}\right)$. □

Sei $m \in \mathbb{N}$. Für den Variationsabstand von $Q_n^{*(m)}$ und $U_{\mathcal{S}_n}$ gilt

$$\begin{aligned}\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &= \frac{1}{2} \sum_{\pi \in \mathcal{S}_n} \left| Q_n^{*(m)}(\{\pi\}) - \frac{1}{n!} \right| \\ &= \sup_{A \subset \mathcal{S}_n} |Q_n^{*(m)}(A) - U_{\mathcal{S}_n}(A)|.\end{aligned}$$

Dabei wird das Supremum für die Menge

$$A_{n,m} \stackrel{\text{def}}{=} \left\{ \pi \in \mathcal{S}_n \mid Q_n^{*(m)}(\{\pi\}) \geq \frac{1}{n!} \right\}$$

angenommen. Da $Q_n^{*(m)}(\{\pi\})$ nach (2.2.6) jedoch nur über $R(\pi)$ von π abhängt und monoton fallend in $R(\pi)$ ist, gilt

$$A_{n,m} = \left\{ \pi \in \mathcal{S}_n \mid R(\pi) \leq r^* \right\},$$

wobei r^* definiert sei durch

$$Q_n^{*(m)}(\{\pi\}) \geq \frac{1}{n!} \iff R(\pi) \leq r^*.$$

Sei $h_{n,R(\pi)}$ wie in Satz 2.6.1 gegeben durch $R(\pi) = \frac{n}{2} + h_{n,R(\pi)}$ und $h^* \stackrel{\text{def}}{=} h_{n,r^*} = r^* - \frac{n}{2}$. Dann erhalten wir

$$Q_n^{*(m)}(\{\pi\}) \geq \frac{1}{n!} \iff h_{n,R(\pi)} \leq h^* \text{ und} \quad (2.6.8)$$

$$A_{n,m} = \left\{ \pi \in \mathcal{S}_n \mid h_{n,R(\pi)} \leq h^* \right\}.$$

Der Variationsabstand von $Q_n^{*(m)}$ und $U_{\mathcal{S}_n}$ kann somit, unter Beachtung von $-\frac{n}{2} + 1 \leq h_{n,r} \leq \frac{n}{2}$ für $r = 1, \dots, n$, wie folgt berechnet werden

$$\begin{aligned} \|Q_n^{*(m)} - U_{\mathcal{S}_n}\| &= Q_n^{*(m)}(A_{n,m}) - U_{\mathcal{S}_n}(A_{n,m}) \\ &= \sum_{\pi \in A_{n,m}} (Q_n^{*(m)}(\{\pi\}) - U_{\mathcal{S}_n}(\{\pi\})) \\ &= \sum_{1 \leq r \leq r^*} \sum_{\pi \in \mathcal{R}_{n,r}} \left(Q_n^{*(m)}(\{\pi\}) - \frac{1}{n!} \right) \\ &= \sum_{1 \leq r \leq r^*} \mathcal{A}_{n,r} \left(q_n^{*(m)}(r) - \frac{1}{n!} \right) \\ &= \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \mathcal{A}_{n,h+\frac{n}{2}} \left(q_n^{*(m)}\left(h + \frac{n}{2}\right) - \frac{1}{n!} \right), \end{aligned} \quad (2.6.9)$$

wobei $\mathcal{A}_{n,h+\frac{n}{2}}$ die Eulerschen Zahlen sind mit $\mathcal{A}_{n,h+\frac{n}{2}} = 0$ für alle $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ mit $h + \frac{n}{2} \notin \{1, \dots, n\}$ und $q_n^{*(m)}(r)$ für $r = 1, \dots, n$ definiert sei durch

$$q_n^{*(m)}(r) \stackrel{\text{def}}{=} \binom{n+2^m-r}{n} \frac{1}{2^{mn}}.$$

Der nächste Satz gibt Auskunft über die Gestalt von h^* . Zuvor notieren wir allerdings noch das folgende Lemma.

2.6.2 Lemma. *Gegeben sei die Situation von Satz 2.6.1. Dann gilt*

$$\mathbf{f}_n = \frac{1}{12c^2} + \mathbf{G}_{1,n} + \mathbf{G}_{2,n}, \quad (2.6.10)$$

wobei $\mathbf{G}_{1,n} \in O_c\left(\frac{1}{n}\right)$ und $\mathbf{G}_{2,n} \in O_c(1)$.

Beweis. In der Situation von Satz 2.6.1 gilt nach (2.6.7) wegen $\mathbf{f}_n = c\sqrt{n}\mathbf{f}_{1,n}$ mit $h = h_{n,r}$

$$\mathbf{f}_n = \frac{1-2h}{24c^2n} - \frac{2h^3 - 3h^2 + h}{6c^2n^3}.$$

Wegen $h = r - \frac{n}{2}$ erhalten wir daraus

$$\begin{aligned} \mathbf{f}_n &= \frac{1-2(r-\frac{n}{2})}{24c^2n} - \frac{2(r-\frac{n}{2})^3 - 3(r-\frac{n}{2})^2 + (r-\frac{n}{2})}{6c^2n^3} \\ &= \frac{1}{24c^2} + \frac{1-2r}{24c^2n} - \frac{(r-\frac{n}{2})^3}{3c^2n^3} + \frac{3(r-\frac{n}{2})^2 - (r-\frac{n}{2})}{6c^2n^3}. \end{aligned} \quad (2.6.11)$$

Wir betrachten nun die einzelnen Terme in (2.6.11). Zunächst setzen wir

$$\begin{aligned} \mathbf{G}_{1,n} &\stackrel{\text{def}}{=} \frac{3(r-\frac{n}{2})^2 - (r-\frac{n}{2})}{6c^2n^3} \quad \text{und} \\ \mathbf{G}_{3,n} &\stackrel{\text{def}}{=} \frac{1-2r}{24c^2n}. \end{aligned}$$

Wegen $1 \leq r \leq n$ erhalten wir $\mathbf{G}_{1,n} \in O_c\left(\frac{1}{n}\right)$ und $\mathbf{G}_{3,n} \in O_c(1)$. Ferner gilt

$$\begin{aligned} -\frac{(r-\frac{n}{2})^3}{3c^2n^3} &= -\frac{1}{3c^2} \left(\frac{r}{n} - \frac{1}{2}\right)^3 \\ &= -\frac{1}{3c^2} \left(\frac{r^3}{n^3} - \frac{3r^2}{2n^2} + \frac{3r}{4n} - \frac{1}{8}\right) \\ &= \frac{1}{24c^2} - \frac{1}{3c^2} \left(\frac{r^3}{n^3} - \frac{3r^2}{2n^2} + \frac{3r}{4n}\right) \\ &= \frac{1}{24c^2} + \mathbf{G}_{4,n}, \end{aligned}$$

wobei $\mathbf{G}_{4,n} \stackrel{\text{def}}{=} -\frac{1}{3c^2} \left(\frac{r^3}{n^3} - \frac{3r^2}{2n^2} + \frac{3r}{4n}\right) \in O_c(1)$. Setzen wir nun noch

$$\mathbf{G}_{2,n} \stackrel{\text{def}}{=} \mathbf{G}_{3,n} + \mathbf{G}_{4,n},$$

so folgt insgesamt

$$\mathbf{f}_n = \frac{1}{12c^2} + \mathbf{G}_{1,n} + \mathbf{G}_{2,n}$$

mit $\mathbf{G}_{1,n} \in O_c\left(\frac{1}{n}\right)$ und $\mathbf{G}_{2,n} \in O_c(1)$. □

Wir kommen nun zu dem bereits angekündigten Resultat über h^* . Es besagt, dass h^* von der Form $h^* = -\frac{\sqrt{n}}{24c} + O_c(1)$ ist.

2.6.3 Satz. Sei $n \geq 1$, $c \in \mathfrak{Z}$, $m = m_{n,c} = \log_2(n^{3/2}c) - \theta_n$, $\pi \in \mathcal{S}_n$ und h^* definiert durch (2.6.8), das heißtt

$$Q_n^{*(m)}(\{\pi\}) \geq \frac{1}{n!} \iff R(\pi) - \frac{n}{2} \leq h^*.$$

Dann gilt

$$h^* = -\frac{\sqrt{n}}{24c} + \frac{1}{12c^2} + \frac{1}{2} + \mathbf{F}_n + \mathbf{G}_n, \quad (2.6.12)$$

wobei $\mathbf{F}_n \in O_c\left(\frac{1}{\sqrt{n}}\right)$ und $\mathbf{G}_n \in O_c(1)$.

Beweis. Sei $c \in \mathfrak{Z}$ fest und $\pi \in \mathcal{S}_n$ mit $R(\pi) = r \leq 2^m \wedge n$. In der Situation von Satz 2.6.1 setzen wir wieder $h = h_{n,r}$. Dann ist $Q_n^{*(m)}(\{\pi\}) \geq \frac{1}{n!}$ nach (2.6.1) äquivalent zu

$$\frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + \mathbf{f}_n\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + \mathbf{g}_n \geq 0,$$

wobei $\mathbf{f}_n \in O_c\left(\frac{h_{n,r}}{n}\right)$ und $\mathbf{g}_n \in O_c\left(\frac{1}{n}\right)$. Setzen wir den Exponenten in (2.6.1) gleich 0 so erhalten wir

$$\begin{aligned} 0 &= \frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + \mathbf{f}_n\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + \mathbf{g}_n \\ \iff h &= -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n - \frac{h^2}{2cn^{3/2}} + c\sqrt{n}\mathbf{g}_n. \end{aligned} \quad (2.6.13)$$

Wir können nun h iterativ in zwei Schritten ermitteln, indem wir h bzw. den Term $-\frac{h^2}{2cn^{3/2}}$ in (2.6.13) wiederholt nach oben und unten abschätzen. Als obere Schranke von $-\frac{h^2}{2cn^{3/2}}$ werden wir dabei 0 wählen.

Wegen $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ gilt $-h^2 \geq -\frac{n^2}{4}$ und somit nach (2.6.13) unter Beachtung von $c > 0$

$$\begin{aligned} h &\geq -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n - \frac{n^2}{8cn^{3/2}} + c\sqrt{n}\mathbf{g}_n \\ &= -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n - \frac{\sqrt{n}}{8c} + c\sqrt{n}\mathbf{g}_n \\ &= -\frac{\sqrt{n}}{6c} + \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n. \end{aligned}$$

Zusammen mit $h \leq -\frac{\sqrt{n}}{24c^2} + \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n$ erhalten wir dann den ersten Iterationsschritt

$$-\frac{\sqrt{n}}{6c} + \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n \leq h \leq -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n. \quad (2.6.14)$$

Wir setzen $\mathbf{F}_{1,n} \stackrel{\text{def}}{=} \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n$. Dann ist $\mathbf{F}_{1,n} \in O_c\left(\frac{h}{n}\right)$ und nach (2.6.14) gilt

$$h^2 \leq \max \left\{ \left(\mathbf{F}_{1,n} - \frac{\sqrt{n}}{24c} \right)^2, \left(\mathbf{F}_{1,n} - \frac{\sqrt{n}}{6c} \right)^2 \right\} \stackrel{\text{def}}{=} \mathbf{F}_{n,\max},$$

wobei $\mathbf{F}_{n,\max} \in O_c(n)$. Daraus folgt

$$\begin{aligned} 0 &\geq -\frac{h^2}{2cn^{3/2}} \\ &\geq -\frac{1}{2cn^{3/2}} \mathbf{F}_{n,\max} \stackrel{\text{def}}{=} -\mathbf{F}_{2,n} \end{aligned}$$

mit $\mathbf{F}_{2,n} \in O_c\left(\frac{1}{\sqrt{n}}\right)$. Dann erhalten wir wiederum zusammen mit (2.6.13) den zweiten Iterationsschritt

$$-\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n - \mathbf{F}_{2,n} + c\sqrt{n}\mathbf{g}_n \leq h \leq -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \mathbf{f}_n + c\sqrt{n}\mathbf{g}_n. \quad (2.6.15)$$

Vermöge Lemma 2.6.2 gilt $\mathbf{f}_n = \frac{1}{12c^2} + \mathbf{G}_n$ mit $\mathbf{G}_n \in O_c(1)$ und somit

$$-\frac{\sqrt{n}}{24c} + \frac{1}{2} + \frac{1}{12c^2} + \mathbf{G}_n - \mathbf{F}_{2,n} + c\sqrt{n}\mathbf{g}_n \leq h \leq -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \frac{1}{12c^2} + \mathbf{G}_n + c\sqrt{n}\mathbf{g}_n.$$

Insgesamt gilt somit wegen $c\sqrt{n}\mathbf{g}_n \in O_c\left(\frac{1}{\sqrt{n}}\right)$

$$h = -\frac{\sqrt{n}}{24c} + \frac{1}{2} + \frac{1}{12c^2} + \mathbf{G}_n + \mathbf{F}_n$$

für ein $\mathbf{F}_n \in O_c\left(\frac{1}{\sqrt{n}}\right)$ und $\mathbf{G}_n \in O_c(1)$. Dies bedeutet, dass h^* von der gewünschten Gestalt in (2.6.12) ist. \square

Wir werden nun Approximationen für die in der Totalvariation

$$\|Q_n^{*(m)} - U_{\mathcal{S}_n}\| = \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \mathcal{A}_{n,h+\frac{n}{2}} \left(q_n^{*(m)} \left(h + \frac{n}{2} \right) - \frac{1}{n!} \right)$$

auftrtenden Terme $\frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}}$ für $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ und $\sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}}$ für $n \rightarrow \infty$ angeben.

Wir beginnen mit dem Term $\frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}}$. Das folgende Lemma entstammt [23] (Proposition 2, Teil (ii) und (iii)). Der zweite Teil ergibt sich allerdings aus einer Modifikation des dortigen Beweises von Proposition 2 (iii). Zur Vollständigkeit geben wir die Beweise für beide Teile an.

2.6.4 Lemma. *Sei $n \in \mathbb{N}$.*

(i) *Für $x > 0$ und $x_n \stackrel{\text{def}}{=} x\sqrt{\frac{n}{12}} + \frac{n}{2}$ gilt*

$$\lim_{n \rightarrow \infty} \sqrt{\frac{n}{12}} \frac{1}{n!} \mathcal{A}_{n,\lfloor x_n \rfloor} = \phi(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

(ii) *Für $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ mit $h + \frac{n}{2} \in \mathbb{N}$ gilt*

$$\frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} = \sqrt{\frac{6}{\pi n}} \exp\left(-\frac{1}{2} \left(\frac{h}{\sqrt{\frac{n}{12}}}\right)^2\right) + \mathbf{R}_n,$$

wobei $\mathbf{R}_n \in o\left(\frac{1}{n}\right)$.

Beweis. Seien X_1, \dots, X_n stochastisch unabhängige und identisch $R(0, 1)$ -verteilte Zufallsgrößen, $S_n \stackrel{\text{def}}{=} \sum_{i=1}^n X_i$ und F_n die Verteilungsfunktion von S_n . Es gilt $ES_n = n/2$ und $Var S_n = n/12$. Wir definieren die Standardisierung von S_n durch

$$T_n \stackrel{\text{def}}{=} \frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}.$$

Dann gilt für die Verteilungsfunktion G_n von T_n und alle $x \in \mathbb{R}$ nach dem zentralen Grenzwertsatz

$$G_n(x) = F_n\left(x\sqrt{\frac{n}{12}} + \frac{n}{2}\right) \xrightarrow{n \rightarrow \infty} \Phi(x) \stackrel{\text{def}}{=} \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt.$$

zu (i): Sei $x > 0$ beliebig, $x_n \stackrel{\text{def}}{=} x\sqrt{\frac{n}{12}} + \frac{n}{2}$, $a_n = a_n(x) \stackrel{\text{def}}{=} x_n - \lfloor x_n \rfloor$ und $n \geq x^2/3$. Dann gilt $x_n \leq n$. Nach Satz 2.3.1 folgt vermöge

$$F_n(x) = G_n\left(\frac{x - \frac{n}{2}}{\sqrt{\frac{n}{12}}}\right)$$

für alle $x \in \mathbb{R}$ und dem Mittelwertsatz der Differentialrechnung

$$\begin{aligned} \frac{1}{n!} \mathcal{A}_{n,\lfloor x_n \rfloor} &= F_n(\lfloor x_n \rfloor) - F_n(\lfloor x_n \rfloor - 1) \\ &= F_n(x_n - a_n) - F_n(x_n - (1 + a_n)) \\ &= G_n\left(x - a_n \sqrt{\frac{12}{n}}\right) - G_n\left(x - (1 + a_n) \sqrt{\frac{12}{n}}\right) \\ &= \sqrt{\frac{12}{n}} G'_n(y_n) \end{aligned}$$

mit $x - (1 + a_n) \sqrt{\frac{12}{n}} < y_n < x - a_n \sqrt{\frac{12}{n}}$. Sei g_n die Dichte der standardisierten Zufallsgröße T_n . Dann gilt $G'_n(x) = g_n(x)$ für alle $x \in \mathbb{R}$. Mit Hilfe einer Edgeworth-Entwicklung (Theorem 1, Abschnitt XVI.2 in [11]) erhalten wir für $n \rightarrow \infty$

$$g_n(x) - \phi(x) = o\left(\frac{1}{\sqrt{n}}\right) \quad (2.6.16)$$

gleichmäßig in x . Hierbei beachten wir $ET_1^3 = 0$ und die quadratische Integrierbarkeit der charakteristischen Funktion von T_1 . Daraus erhalten wir

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt{\frac{n}{12}} \frac{1}{n!} \mathcal{A}_{n,\lfloor x_n \rfloor} &= \lim_{n \rightarrow \infty} G'_n(y_n) \\ &= \lim_{n \rightarrow \infty} g_n(y_n) \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}. \end{aligned}$$

zu (ii): Nach Theorem 1, Abschnitt I.9 in [11] besitzt S_{n+1} die Lebesgue-Dichte (λ -Dichte) f_{n+1} mit

$$f_{n+1}(x) = \frac{1}{n!} \sum_{i=0}^{\lfloor x \rfloor} (-1)^i \binom{n+1}{i} (x-i)^n 1_{(0,n+1)}(x) \quad (2.6.17)$$

für $x \in \mathbb{R}$. Sei $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ mit $h + \frac{n}{2} \in \mathbb{N}$. Dann gilt $x_n \stackrel{\text{def}}{=} h + \frac{n}{2} \leq n$ und mit (2.3.5) und (2.6.17) folgt

$$\begin{aligned} \frac{1}{n!} \mathcal{A}_{n,x_n} &= F_n(x_n) - F_n(x_n - 1) \\ &= \frac{1}{n!} \sum_{i=0}^{x_n} (-1)^i \binom{n+1}{i} (x_n - i)^n \\ &= f_{n+1}(x_n). \end{aligned} \quad (2.6.18)$$

Für die Dichten f_{n+1} von S_{n+1} und g_{n+1} von $T_{n+1} = (S_{n+1} - \frac{n+1}{2}) / \sqrt{\frac{n+1}{12}}$ gilt nach dem Transformationssatz die Beziehung

$$f_{n+1}(x) = \sqrt{\frac{12}{n+1}} g_{n+1} \left(\frac{x_n - \frac{n+1}{2}}{\sqrt{\frac{n+1}{12}}} \right),$$

woraus mit (2.6.18)

$$\frac{1}{n!} \mathcal{A}_{n,x_n} = \sqrt{\frac{12}{n+1}} g_{n+1} \left(\frac{x_n - \frac{n+1}{2}}{\sqrt{\frac{n+1}{12}}} \right)$$

folgt. Mit Hilfe von (2.6.16) erhalten wir für $n \rightarrow \infty$

$$g_{n+1}(x) - \phi(x) = o\left(\frac{1}{\sqrt{n+1}}\right) = o\left(\frac{1}{\sqrt{n}}\right) \quad (2.6.19)$$

gleichmäßig in x . Aus (2.6.19) erhalten wir dann unter Beachtung von $\sqrt{\frac{12}{n+1}} o\left(\frac{1}{\sqrt{n}}\right) = o\left(\frac{1}{n}\right)$ und

$$\begin{aligned} \sqrt{\frac{12}{n+1}} \phi \left(\frac{x_n - \frac{n+1}{2}}{\sqrt{\frac{n+1}{12}}} \right) &= \sqrt{\frac{12}{n+1}} \phi \left(\frac{h - \frac{1}{2}}{\sqrt{\frac{n+1}{12}}} \right) \\ &= \sqrt{\frac{12}{n}} \phi \left(\frac{h}{\sqrt{\frac{n}{12}}} \right) + o\left(\frac{1}{n}\right) \end{aligned}$$

insgesamt

$$\begin{aligned} \frac{1}{n!} \mathcal{A}_{n,\lfloor x_n \rfloor} &= \sqrt{\frac{12}{n+1}} \phi \left(\frac{x_n - \frac{n+1}{2}}{\sqrt{\frac{n+1}{12}}} \right) + \sqrt{\frac{12}{n+1}} o\left(\frac{1}{\sqrt{n}}\right) \\ &= \sqrt{\frac{12}{n}} \phi \left(\frac{h}{\sqrt{\frac{n}{12}}} \right) + o\left(\frac{1}{n}\right) \\ &= \sqrt{\frac{12}{n}} \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{h}{\sqrt{\frac{n}{12}}} \right)^2 \right) + o\left(\frac{1}{n}\right) \\ &= \sqrt{\frac{6}{\pi n}} \exp \left(-\frac{1}{2} \left(\frac{h}{\sqrt{\frac{n}{12}}} \right)^2 \right) + o\left(\frac{1}{n}\right). \end{aligned} \quad \square$$

Bayer und Diaconis behaupten in [6], dass sich die Asymptotik in Lemma 2.6.4 (ii) unter Benutzung derselben Quellen noch verschärfen lässt:

2.6.5 Lemma. *Sei $n \in \mathbb{N}$. Für $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ mit $h + \frac{n}{2} \in \mathbb{N}$ gilt*

$$\frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} = \sqrt{\frac{6}{\pi n}} \exp\left(-\frac{1}{2} \left(\frac{h}{\sqrt{\frac{n}{12}}}\right)^2\right) (1 + \mathbf{R}_n),$$

wobei $\mathbf{R}_n \in o\left(\frac{1}{\sqrt{n}}\right)$.

Mit Hilfe des obigen Lemmas zeigen Bayer und Diaconis [6] Theorem 2.6.8. Wir werden jedoch einen anderen Weg einschlagen (vergleiche den Beweis von Theorem 4, Abschnitt 4 in [6] und den Beweis von Gleichung (2.6.22) im Beweis von Theorem 2.6.8). Zunächst benötigen wir das folgende Lemma, das Auskunft über das Verhalten der Summe $\sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}}$ für $n \rightarrow \infty$ gibt, wobei nach Satz 2.6.3 $h^* = -\frac{\sqrt{n}}{24c} + O_c(1)$ gilt.

2.6.6 Lemma. *Sei $n \in \mathbb{N}$ und Φ die Verteilungsfunktion der Standardnormalverteilung. Dann gilt*

$$\sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} = \Phi\left(-\frac{1}{4c\sqrt{3}}\right) + \mathbf{S}_n,$$

wobei $\mathbf{S}_n \in O_c\left(\frac{1}{\sqrt{n}}\right)$.

Beweis. Wie im Beweis von Lemma 2.6.4 sei wiederum S_n die Summe von n stochastisch unabhängigen und identisch $R(0, 1)$ -verteilten Zufallsgrößen, F_n die Verteilungsfunktion von S_n ,

$$T_n = \frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

die Standardisierung von S_n und G_n die Verteilungsfunktion von T_n . Durch Integration der Edgeworth-Entwicklung (2.6.19) der Dichte g_n von T_n (siehe auch (4.1), Abschnitt XVI.4 in [11]) erhalten wir für $n \rightarrow \infty$

$$G_n(x) - \Phi(x) = o\left(\frac{1}{\sqrt{n}}\right)$$

gleichmäßig in x . Dann gilt vermöge $F_n(x) = G_n\left(\left(x - \frac{n}{2}\right) / \sqrt{\frac{n}{12}}\right)$ für alle $x \in \mathbb{R}$ und (2.3.4) für $n \rightarrow \infty$

$$\begin{aligned} \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} &= \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \left(F_n\left(h + \frac{n}{2}\right) - F_n\left(h + \frac{n}{2} - 1\right) \right) \\ &= F_n\left(h^* + \frac{n}{2}\right) - F_n(0) \\ &= G_n\left(\frac{h^*}{\sqrt{\frac{n}{12}}}\right) \\ &= \Phi\left(\frac{h^*}{\sqrt{\frac{n}{12}}}\right) + o\left(\frac{1}{\sqrt{n}}\right). \end{aligned} \quad (2.6.20)$$

Wegen $h^* = -\frac{\sqrt{n}}{24c} + O_c(1)$ gilt

$$\frac{h^*}{\sqrt{\frac{n}{12}}} = -\frac{1}{2c\sqrt{12}} + O_c\left(\frac{1}{\sqrt{n}}\right) = -\frac{1}{4c\sqrt{3}} + O_c\left(\frac{1}{\sqrt{n}}\right)$$

und somit für $n \rightarrow \infty$

$$\begin{aligned} \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} &= \Phi\left(-\frac{1}{4c\sqrt{3}} + O_c\left(\frac{1}{\sqrt{n}}\right)\right) + o\left(\frac{1}{\sqrt{n}}\right) \\ &= \Phi\left(-\frac{1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{\sqrt{n}}\right) + o\left(\frac{1}{\sqrt{n}}\right) \\ &= \Phi\left(-\frac{1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

dabei geht in der zweiten Zeile der Mittelwertsatz der Differentialrechnung ein. \square

Das nächste Lemma ist wiederum technischer Natur.

2.6.7 Lemma. Für $n \in \mathbb{N}$ sei S_n die Summe von n stochastisch unabhängigen und identisch $R(0, 1)$ -verteilten Zufallsgrößen X_1, \dots, X_n und

$$T_n = \frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

die Standardisierung von S_n . Dann gelten die folgenden Aussagen:

(a) Die Familie $(\exp(aT_n))_{n \in \mathbb{N}}$ ist für jedes $a \in \mathbb{R}$ gleichgradig integrierbar.

(b) Sei G_n die Verteilungsfunktion von T_n und $(x_n)_{n \geq 1}$ eine reelle Folge mit $\lim_{n \rightarrow \infty} x_n / \sqrt{n} = 0$, das heißt $x_n \in o(\sqrt{n})$. Dann gilt

$$\lim_{n \rightarrow \infty} \frac{1 - G_n(x_n)}{1 - \Phi(x_n)} = \lim_{n \rightarrow \infty} \frac{G_n(-x_n)}{\Phi(-x_n)} = 1,$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung sei.

Beweis. zu (a): Für $a = 0$ ist nichts zu zeigen. Sei daher $a \neq 0$ beliebig. Nach Korollar 50.3. (e) in [3] ist $(\exp(aT_n))_{n \in \mathbb{N}}$ gleichgradig integrierbar, falls ein $p > 1$ existiert mit

$$\sup_{n \in \mathbb{N}} E |\exp(aT_n)|^p < \infty.$$

Wegen $|\exp(aT_n)|^p = \exp(paT_n)$ für alle $p \in \mathbb{R}$ reicht es daher zu zeigen

$$\sup_{n \in \mathbb{N}} E \exp(aT_n) < \infty.$$

Sei ψ die momenterzeugende Funktion von X_1 , das heißt $\psi(t) = E \exp(tX_1)$, $t \in \mathbb{R}$. Es gilt $\psi(0) = 1$ und für $t \neq 0$

$$\psi(t) = \int_0^1 e^{tx} dx = \frac{e^t - 1}{t}.$$

Dann erhalten wir für $n \geq 1$, da X_1, \dots, X_n stochastisch unabhängig und identisch verteilt sind, mit $b \stackrel{\text{def}}{=} a\sqrt{12}$ für $n \geq 1$

$$\begin{aligned} E \exp(aT_n) &= E \exp \left(a \left(\frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}} \right) \right) \\ &= E \exp \left(\sum_{k=1}^n \left(\frac{a\sqrt{12}}{\sqrt{n}} X_k - \frac{a\sqrt{12}}{2\sqrt{n}} \right) \right) \\ &= \left(\exp \left(-\frac{b}{2\sqrt{n}} \right) E \exp \left(\frac{b}{\sqrt{n}} X_1 \right) \right)^n \\ &= \left(\exp \left(-\frac{b}{2\sqrt{n}} \right) \psi \left(\frac{b}{\sqrt{n}} \right) \right)^n \\ &= \left(\exp \left(-\frac{b}{2\sqrt{n}} \right) \frac{\exp \left(\frac{b}{\sqrt{n}} \right) - 1}{\frac{b}{\sqrt{n}}} \right)^n \end{aligned}$$

$$= \left(\frac{\exp\left(\frac{b}{2\sqrt{n}}\right) - \exp\left(-\frac{b}{2\sqrt{n}}\right)}{\frac{b}{\sqrt{n}}} \right)^n.$$

Mittels der Taylor-Entwicklungen

$$\begin{aligned}\exp\left(\frac{b}{2\sqrt{n}}\right) &= 1 + \frac{b}{2\sqrt{n}} + \frac{b^2}{8n} + \frac{b^3}{48n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right), \\ \exp\left(-\frac{b}{2\sqrt{n}}\right) &= 1 - \frac{b}{2\sqrt{n}} + \frac{b^2}{8n} - \frac{b^3}{48n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right)\end{aligned}$$

ergibt sich daraus

$$\begin{aligned}\lim_{n \rightarrow \infty} E \exp(aT_n) &= \lim_{n \rightarrow \infty} \left(\frac{\frac{b}{\sqrt{n}} + \frac{b^3}{24n^{3/2}} + o\left(\frac{1}{n^{3/2}}\right)}{\frac{b}{\sqrt{n}}} \right)^n \\ &= \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \left(\frac{b^2}{24} + o(1) \right) \right)^n \\ &= \exp\left(\frac{b^2}{24}\right) = \exp\left(\frac{a^2}{2}\right) < \infty.\end{aligned}$$

Dann folgt $\sup_{n \in \mathbb{N}} E \exp(aT_n) < \infty$, was wir beweisen mussten.

zu (b): Die Aussage ist ein Spezialfall von Theorem 2, Gleichung (7.28), Abschnitt XVI.7 in [11]. In unserer Situation gilt dort $\lambda_1 = 0$ wegen $\mu_3 = ET_1^3 = 0$. Die Behauptung folgt dann aus der Gültigkeit von $1 - G_n(x) = G_n(-x)$ und $1 - \Phi(x) = \Phi(-x)$ für alle $x \in \mathbb{R}$ und $n \geq 1$. \square

Wir kommen nun zur Hauptaussage dieses Abschnitts, mit Hilfe derer wir nachweisen, dass die Familie $(Q_n)_{n \geq 1}$ der GSR-(2, n)-Mischnmethoden einen starken Cut-Off-Effekt mit Cut-Off-Zeiten $((\frac{3}{2} \log_2 n, 1))_{n \geq 1}$ besitzt. Wir erinnern daran, dass $Q_n^{*(m)} = Q_n^{*(0)} = \delta_{\text{id}}$ für $m < 0$ gesetzt war.

2.6.8 Theorem. Sei $c \in \mathfrak{Z} = \{2^j \mid j \in \mathbb{Z}\}$ und $m_n = m_{n,c} = \lfloor \log_2(n^{3/2}c) \rfloor$ für $n \in \mathbb{N}$. Dann gilt

$$\lim_{n \rightarrow \infty} \|Q_n^{*(m_n)} - U_{S_n}\| = 1 - 2\Phi\left(-\frac{1}{4c\sqrt{3}}\right),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung sei.

Beweis. Sei $n \geq 1$, $c \in \mathfrak{Z}$, $m_n = \lfloor \log_2(n^{3/2}c) \rfloor$. Nach (2.6.9) gilt

$$\begin{aligned}\|Q_n^{*(m_n)} - U_{S_n}\| &= \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \mathcal{A}_{n,h+\frac{n}{2}} \left(q_n^{*(m_n)} \left(h + \frac{n}{2} \right) - \frac{1}{n!} \right) \\ &= \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)} \left(h + \frac{n}{2} \right) - \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}},\end{aligned}\tag{2.6.21}$$

wobei $q_n^{*(m_n)} \left(h + \frac{n}{2} \right)$ gegeben ist durch

$$q_n^{*(m_n)} \left(h + \frac{n}{2} \right) = \binom{n + 2^{m_n} - \left(h + \frac{n}{2} \right)}{n} \frac{1}{2^{m_n n}}.$$

Aufgrund von unserer Definition $\mathcal{A}_{n,r} = 0$ für $r \neq 1, \dots, n$, sei im Folgenden $h + \frac{n}{2}$ stets ganzzahlig. Nach Lemma 2.6.6 gilt für den zweiten Summanden in (2.6.21)

$$\lim_{n \rightarrow \infty} \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} = \Phi \left(-\frac{1}{4c\sqrt{3}} \right).\tag{2.6.22}$$

Wir betrachten nun den ersten Summanden in (2.6.21). Zunächst setzen wir

$$\begin{aligned}I_n^1 &\stackrel{\text{def}}{=} \left\{ h \mid -\frac{10n^{3/4}}{\sqrt{c}} \leq h \leq h^* \right\} \quad \text{und} \\ I_n^2 &\stackrel{\text{def}}{=} \left\{ h \mid -\frac{n}{2} + 1 \leq h < -\frac{10n^{3/4}}{\sqrt{c}} \right\},\end{aligned}$$

dabei sei n so groß, dass $I_n^1, I_n^2 \neq \emptyset$.

Wir betrachten zunächst

$$\sum_{h \in I_n^1} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)} \left(h + \frac{n}{2} \right).$$

Sei S_n die Summe von n stochastisch unabhängigen und identisch $R(0, 1)$ -verteilten Zufallsgrößen und

$$T_n = \frac{S_n - \frac{n}{2}}{\sqrt{\frac{n}{12}}}$$

die Standardisierung von S_n . Nach Satz 2.3.1 gilt für alle h

$$\frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} = P \left(h + \frac{n}{2} - 1 < S_n \leq h + \frac{n}{2} \right).\tag{2.6.23}$$

Für $h \in I_n^1$, also $-\frac{10n^{3/4}}{\sqrt{c}} \leq h \leq h^* = -\frac{\sqrt{n}}{24c} + O_c(1)$, gilt ferner

$$-\frac{10}{n^{1/4}c^{3/2}} \leq \frac{h}{cn} \leq -\frac{1}{24c^2\sqrt{n}} + O_c\left(\frac{1}{n}\right)$$

und damit für alle $h \in I_n^1$

$$\frac{1}{2c\sqrt{n}} + O_c\left(\frac{1}{\sqrt{n}}\right) - \frac{1}{2} \left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right) = O_c\left(\frac{1}{\sqrt{n}}\right).$$

Insgesamt erhalten wir daraus weiter unter Benutzung von Satz 2.6.1 und (2.6.23)

$$\begin{aligned} \sum_{h \in I_n^1} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)}\left(h + \frac{n}{2}\right) &= \sum_{h \in I_n^1} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} n! q_n^{*(m_n)}\left(h + \frac{n}{2}\right) \\ &= \sum_{h \in I_n^1} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} \exp\left(\frac{1}{c\sqrt{n}}\left(-h + \frac{1}{2} + O_c\left(\frac{h}{n}\right)\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right) \\ &= \exp\left(\frac{-1}{24c^2}\right) \sum_{h \in I_n^1} \frac{\mathcal{A}_{n,h+\frac{n}{2}}}{n!} \exp\left(-\frac{h}{c\sqrt{n}} + \frac{1}{2c\sqrt{n}} + O_c\left(\frac{1}{\sqrt{n}}\right) - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right) \\ &= \exp\left(\frac{-1}{24c^2}\right) \sum_{h \in I_n^1} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} \exp\left(-\frac{h}{c\sqrt{n}} + O_c\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \exp\left(\frac{-1}{24c^2}\right) \sum_{h \in I_n^1} P\left(h + \frac{n}{2} - 1 < S_n \leq h + \frac{n}{2}\right) \exp\left(-\frac{h}{c\sqrt{n}} + O_c\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \exp\left(\frac{-1}{24c^2}\right) \sum_{h \in I_n^1} P\left(\frac{h-1}{\sqrt{\frac{n}{12}}} < T_n \leq \frac{h}{\sqrt{\frac{n}{12}}}\right) \exp\left(-\frac{h}{c\sqrt{n}} + O_c\left(\frac{1}{\sqrt{n}}\right)\right) \\ &= \exp\left(\frac{-1}{24c^2}\right) E\left(\sum_{h \in I_n^1} 1_{\left\{\frac{h-1}{\sqrt{n/12}} < T_n \leq \frac{h}{\sqrt{n/12}}\right\}} \exp\left(-\frac{h}{c\sqrt{n}}\right)\right) e^{O_c\left(\frac{1}{\sqrt{n}}\right)} \\ &= \exp\left(\frac{-1}{24c^2}\right) E\left(\sum_{h \in I_n^1} 1_{\left\{\frac{h-1}{\sqrt{n/12}} < T_n \leq \frac{h}{\sqrt{n/12}}\right\}} \exp\left(-\frac{1}{c\sqrt{12}} \frac{h}{\sqrt{\frac{n}{12}}}\right)\right) e^{O_c\left(\frac{1}{\sqrt{n}}\right)}. \end{aligned} \tag{2.6.24}$$

Für $h \in I_n^1$ gilt auf der Menge

$$\left\{ \frac{h-1}{\sqrt{\frac{n}{12}}} < T_n \leq \frac{h}{\sqrt{\frac{n}{12}}} \right\}$$

wegen $c > 0$

$$-\frac{1}{c\sqrt{12}} \left(T_n + \frac{1}{\sqrt{\frac{n}{12}}} \right) < -\frac{1}{c\sqrt{12}} \frac{h}{\sqrt{\frac{n}{12}}} \leq -\frac{1}{c\sqrt{12}} T_n.$$

Sei nun n so groß, dass $h^* = -\frac{\sqrt{n}}{24c} + O_c(1) < 0$ und $\underline{h}_n \stackrel{\text{def}}{=} \min\{h \in I_n^1\}$. Dann gilt wegen $h + \frac{n}{2} \in \mathbb{N}$ für alle $h \in I_n^1$ und $I_n^1 = \left\{ h \mid -\frac{10n^{3/4}}{\sqrt{c}} \leq h \leq h^* \right\}$

$$\underline{h}_n = -\frac{10n^{3/4}}{\sqrt{c}} + a_n$$

für ein $a_n \in [0, 1)$. Wir können den Erwartungswert in (2.6.24) somit unter Benutzung von

$$\frac{\underline{h}_n - 1}{\sqrt{\frac{n}{12}}} = -\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} + \frac{a_n - 1}{\sqrt{\frac{n}{12}}}$$

nach oben abschätzen durch

$$\begin{aligned} E \left(\sum_{h \in I_n^1} 1_{\left\{ \frac{h-1}{\sqrt{n/12}} < T_n \leq \frac{h}{\sqrt{n/12}} \right\}} \exp \left(-\frac{1}{c\sqrt{12}} T_n \right) \right) \\ = \int_{-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} + \frac{a_n - 1}{\sqrt{\frac{n}{12}}} }^{-\frac{1}{4c\sqrt{3}} + O_c\left(\frac{1}{\sqrt{n}}\right)} \exp \left(-\frac{1}{c\sqrt{12}} x \right) dP^{T_n} \end{aligned}$$

und nach unten abschätzen durch

$$\begin{aligned} E \left(\sum_{h \in I_n^1} 1_{\left\{ \frac{h-1}{\sqrt{n/12}} < T_n \leq \frac{h}{\sqrt{n/12}} \right\}} \exp \left(-\frac{1}{c\sqrt{12}} \left(T_n + \frac{1}{\sqrt{\frac{n}{12}}} \right) \right) \right) \\ = \int_{-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} + \frac{a_n - 1}{\sqrt{\frac{n}{12}}} }^{-\frac{1}{4c\sqrt{3}} + O_c\left(\frac{1}{\sqrt{n}}\right)} \exp \left(-\frac{1}{c\sqrt{12}} \left(x + \frac{1}{\sqrt{\frac{n}{12}}} \right) \right) dP^{T_n}. \end{aligned}$$

Nach Lemma 2.6.7 ist die Familie $(\exp(-1/(c\sqrt{12}) T_n))_{n \geq 1}$ gleichgradig integrierbar. Für die obere Schranke gilt dann wegen der gleichgradigen Integrierbarkeit

von $(\exp(-1/(c\sqrt{12}) T_n))_{n \geq 1}$ und dem zentralen Grenzwertsatz

$$\lim_{n \rightarrow \infty} \int_{-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} + \frac{a_n - 1}{\sqrt{\frac{n}{12}}} }^{-\frac{1}{4c\sqrt{3}} + O_c\left(\frac{1}{\sqrt{n}}\right)} \exp\left(-\frac{1}{c\sqrt{12}}x\right) dP^{T_n} = \int_{-\infty}^{-\frac{1}{4c\sqrt{3}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} - \frac{x}{c\sqrt{12}}\right) dx.$$

Dasselbe Ergebnis erhalten wir analog für die untere Schranke. Insgesamt gilt somit zusammen mit (2.6.24)

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{h \in I_n^1} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)}\left(h + \frac{n}{2}\right) &= \exp\left(-\frac{1}{24c^2}\right) \int_{-\infty}^{-\frac{1}{4c\sqrt{3}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2} - \frac{x}{c\sqrt{12}}\right) dx \\ &= \int_{-\infty}^{-\frac{1}{4c\sqrt{3}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(x + \frac{1}{2c\sqrt{3}}\right)^2\right) dx \\ &= \Phi\left(\frac{1}{4c\sqrt{3}}\right). \end{aligned} \quad (2.6.25)$$

Wir betrachten nun

$$\sum_{h \in I_n^2} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)}\left(h + \frac{n}{2}\right).$$

Da die Funktion

$$r \mapsto q_n^{*(m_n)}(r) = \binom{n + 2^{m_n} - r}{n} \frac{1}{2^{m_n n}}$$

monoton fallend in r ist, gilt für alle $-\frac{n}{2} + 1 \leq h \leq \frac{n}{2}$ mit $h + \frac{n}{2} \in \mathbb{N}$

$$q_n^{*(m_n)}\left(h + \frac{n}{2}\right) \leq q_n^{*(m_n)}(1).$$

Für $n \rightarrow \infty$ folgt daraus mit Hilfe von Satz 2.6.1, wenn dort $h_{n,r} = -\frac{n}{2} + 1$ gewählt wird,

$$\begin{aligned} q_n^{*(m_n)}\left(h + \frac{n}{2}\right) \\ \leq q_n^{*(m_n)}(1) \\ = \frac{1}{n!} \exp\left(\frac{1}{c\sqrt{n}}\left(\frac{n}{2} - \frac{1}{2} + O_c(1)\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{-\frac{n}{2} + 1}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n!} \exp\left(\frac{\sqrt{n}}{2c}\right) \exp\left(O_c\left(\frac{1}{\sqrt{n}}\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{1}{4c^2} - \frac{1}{c^2n} + \frac{1}{c^2n^2}\right)\right) \\
&= \frac{1}{n!} \exp\left(\frac{\sqrt{n}}{2c}\right) \exp\left(O_c\left(\frac{1}{\sqrt{n}}\right)\right) \exp\left(-\frac{1}{6c^2}\right) \\
&\leq \frac{1}{n!} \exp\left(\frac{\sqrt{n}}{2c}\right) \exp\left(O_c\left(\frac{1}{\sqrt{n}}\right)\right) \\
&\sim \frac{1}{n!} \exp\left(\frac{\sqrt{n}}{2c}\right). \tag{2.6.26}
\end{aligned}$$

Sei $\bar{h}_n \stackrel{\text{def}}{=} \max\{h \in I_n^2\}$. Dann gilt wegen $h + \frac{n}{2} \in \mathbb{N}$ für alle $h \in I_n^2$ und $I_n^2 = \left\{h \mid -\frac{n}{2} + 1 \leq h < -\frac{10n^{3/4}}{\sqrt{c}}\right\}$

$$\bar{h}_n = -\frac{10n^{3/4}}{\sqrt{c}} - b_n$$

für ein $b_n \in [0, 1)$. Damit erhalten wir nach Satz 2.3.1 mit $S_n > 0$ P -f.s

$$\begin{aligned}
\sum_{h \in I_n^2} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} &= \sum_{h \in I_n^2} P\left(h + \frac{n}{2} - 1 < S_n \leq h + \frac{n}{2}\right) \\
&= P\left(0 < S_n \leq \bar{h}_n + \frac{n}{2}\right) \\
&= F_n\left(\bar{h}_n + \frac{n}{2}\right) \\
&= G_n\left(\frac{\bar{h}_n}{\sqrt{\frac{n}{12}}}\right) \\
&= G_n\left(\left(-\frac{10n^{3/4}}{\sqrt{c}} - b_n\right) \sqrt{\frac{12}{n}}\right) \\
&\sim G_n\left(-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}}\right). \tag{2.6.27}
\end{aligned}$$

Bayer und Diaconis [6] behaupten, dass nach Abschnitt XVI.7 in [11] die folgende Beziehung gilt

$$\sum_{h \in I_n^2} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} \sim \frac{1}{10n^{1/4}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}}\right)^2\right).$$

Mittels (2.6.27) gelangen wir zu einer anderen Beziehung (vergleiche (2.6.28)). Nach Lemma 2.6.7 folgt mit $x_n \stackrel{\text{def}}{=} 10\sqrt{12}n^{1/4}/\sqrt{c}$ und $\Phi(-x) \sim 1/(x\sqrt{2\pi})e^{-x^2/2}$,

$x \rightarrow \infty$, für $n \rightarrow \infty$

$$\begin{aligned} \sum_{h \in I_n^2} \frac{1}{n!} \mathcal{A}_{n,h+\frac{n}{2}} &= G_n \left(-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right) \\ &\sim \Phi \left(-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right) \\ &\sim \frac{\sqrt{c}}{10n^{1/4}\sqrt{24\pi}} \exp \left(-\frac{1}{2} \left(\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right)^2 \right) \end{aligned} \quad (2.6.28)$$

und somit zusammen mit (2.6.26) für $n \rightarrow \infty$

$$\begin{aligned} \sum_{h \in I_n^2} \mathcal{A}_{n,h+\frac{n}{2}} q_n^{*(m_n)} \left(h + \frac{n}{2} \right) &\leq G_n \left(-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right) \exp \left(\frac{\sqrt{n}}{2c} \right) \exp \left(O_c \left(\frac{1}{\sqrt{n}} \right) \right) \\ &\sim \Phi \left(-\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right) \exp \left(\frac{\sqrt{n}}{2c} \right) \\ &\sim \frac{\sqrt{c}}{10n^{1/4}\sqrt{24\pi}} \exp \left(-\frac{1}{2} \left(\frac{10\sqrt{12}n^{1/4}}{\sqrt{c}} \right)^2 + \frac{\sqrt{n}}{2c} \right) \\ &= \frac{\sqrt{c}}{10n^{1/4}\sqrt{24\pi}} \exp \left(-\frac{((10\sqrt{12})^2 - 1)\sqrt{n}}{2c} \right) \\ &= O_c \left(\frac{1}{n^{1/4}} \right). \end{aligned} \quad (2.6.29)$$

Fassen wir nun die Ergebnisse aus (2.6.22), (2.6.25) und (2.6.29) zusammen, so erhalten wir abschließend unter Benutzung von $\Phi(x) = 1 - \Phi(-x)$ für alle $x \in \mathbb{R}$

$$\begin{aligned} \lim_{n \rightarrow \infty} \|Q_n^{*(m_n)} - U_{S_n}\| &= \lim_{n \rightarrow \infty} \sum_{-\frac{n}{2}+1 \leq h \leq h^*} \mathcal{A}_{n,h+\frac{n}{2}} \left(q_n^{*(m_n)} \left(h + \frac{n}{2} \right) - \frac{1}{n!} \right) \\ &= \Phi \left(\frac{1}{4c\sqrt{3}} \right) - \Phi \left(-\frac{1}{4c\sqrt{3}} \right) \\ &= 1 - 2\Phi \left(-\frac{1}{4c\sqrt{3}} \right) \end{aligned}$$

und somit die Behauptung. \square

Ersetzen wir in Theorem 2.6.8 den Parameter c wiederum durch 2^j , $j \in \mathbb{Z}$, so entspricht j der Anzahl der Mischvorgänge vor bzw. nach $\lfloor \frac{3}{2} \log_2 n \rfloor$ und wir erhalten vermöge

$$\lim_{c \rightarrow \infty} \left[1 - 2\Phi \left(-\frac{1}{4c\sqrt{3}} \right) \right] = \lim_{j \rightarrow \infty} \left[1 - 2\Phi \left(-\frac{2^{-j}}{4\sqrt{3}} \right) \right] = 0 \quad \text{und}$$

$$\lim_{c \rightarrow 0} \left[1 - 2\Phi \left(-\frac{1}{4c\sqrt{3}} \right) \right] = \lim_{j \rightarrow -\infty} \left[1 - 2\Phi \left(-\frac{2^{-j}}{4\sqrt{3}} \right) \right] = 1$$

unter Hinweis auf Definition 2.4.2 unmittelbar das folgende Theorem.

2.6.9 Theorem. *Wird ein Stapel mit n Karten $m_n = (\lfloor \frac{3}{2} \log_2 n \rfloor + j)$ -mal, $j \in \mathbb{Z}$, $j \geq -\lfloor \frac{3}{2} \log_2 n \rfloor$, nach der GSR-(2, n)-Mischmethode $Q_{2,n}$ gemischt, so gilt*

$$\lim_{n \rightarrow \infty} \|Q_{2,n}^{*(m_n)} - U_{S_n}\| = 1 - 2\Phi \left(-\frac{2^{-j}}{4\sqrt{3}} \right),$$

wobei Φ die Verteilungsfunktion der Standardnormalverteilung sei.

Ferner besitzt die Familie $(Q_{2,n})_{n \geq 1}$ einen starken Cut-Off-Effekt mit Cut-Off-Zeiten $((\lfloor \frac{3}{2} \log_2 n \rfloor, 1))_{n \geq 1}$.

Für große n sind somit $\lfloor \frac{3}{2} \log_2 n \rfloor$ Riffle Shuffle hinreichend und notwendig, um einen Kartenstapel mit n Karten hinreichend zu mischen. Mit $m_n = \lfloor \frac{3}{2} \log_2 n \rfloor$ gilt für den Variationsabstand $\|Q_{2,n}^{*(m_n)} - U_{S_n}\|$ für große n

$$\|Q_{2,n}^{*(m_n)} - U_{S_n}\| \approx 1 - 2\Phi \left(-\frac{1}{4\sqrt{3}} \right) \approx 0.115.$$

In Tabelle 2.4 und Abbildung 2.7 sind die Werte der Funktion $j \mapsto 1 - 2\Phi \left(-\frac{2^{-j}}{4\sqrt{3}} \right)$ abgetragen.

j	-6	-5	-4	-3	-2	-1	0	1	2	3	4
$f(j)$	1.000	0.999	0.979	0.752	0.436	0.227	0.115	0.058	0.029	0.014	0.007

Tabelle 2.4: $f(j) = 1 - 2\Phi \left(-\frac{2^{-j}}{4\sqrt{3}} \right)$ für $j = -6, \dots, 4$.

Welcher Wert von der Totalvariation tatsächlich unterschritten werden muss, um den Kartenstapel als hinreichend gemischt anzusehen, liegt in der Beurteilung des jeweiligen Betrachters. So findet sich beispielsweise in [20] die Schranke

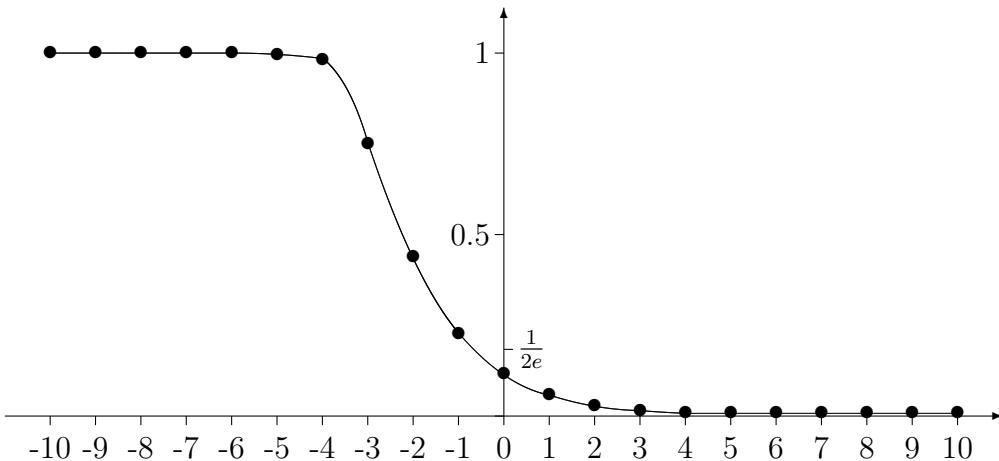


Abbildung 2.7: $j \mapsto 1 - 2\Phi\left(-\frac{2^{-j}}{4\sqrt{3}}\right)$ für $j = -10, \dots, 10$.

$1/(2e) \approx 0.184$, während von Bayer und Diaconis [6] der Kartenstapel schon bei einem Wert von $1/2$ als hinreichend gemischt angesehen wird. Für große n sind nach Theorem 2.6.9 $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ Mischvorgänge nötig um den Wert $1/2$ in Totalvariation zu unterschreiten. Die Schranke $1/2$ wird von Mann [16] durch die folgende Überlegung motiviert:

Wir nehmen an, ein Kartenstapel mit n Karten sei so gemischt, dass jede Reihenfolge der Karten gleich wahrscheinlich ist, das heißt jede Reihenfolge $\pi \in \mathcal{S}_n$ habe die Wahrscheinlichkeit $U_{\mathcal{S}_n} = 1/n!$. n sei als gerade vorausgesetzt. Nach dem Mischen fällt die oberste Karte vom Stapel, so dass wir deren Wert i erkennen. Legen wir diese Karte nun wieder mit der Rückseite oben auf den Kartenstapel, so ist die zufällige Reihenfolge der Karten zerstört, da wir wissen, welchen Wert die oberste Karte trägt. Legen wir die Karte wieder an eine zufällige Position innerhalb des Stapels, so ist die zufällige Reihenfolge der Karten wieder gegeben. Wir platzieren die Karte allerdings an einer zufälligen Position innerhalb der oberen Hälfte des Stapels und zwar so, dass wir nicht wissen, welche Position dies ist. Die möglichen Permutationen nach denen der Kartenstapel dann gemischt ist, entsprechen exakt der Menge der Permutationen in \mathcal{S}_n , die Karte i in die obere Hälfte des Stapels mischen. Dies ist genau die Hälfte aller Permutationen in \mathcal{S}_n . Die Wahrscheinlichkeit für eine solche Permutation π ist dann gegeben durch

$Q(\{\pi\}) = 2/n!$, alle übrigen Permutationen treten mit Wahrscheinlichkeit 0 auf. Wir berechnen nun den Verlust, gemessen in Totalvariation, den wir durch dieses Vorgehen im Vergleich zur vollständigen Zufälligkeit der Kartenreihenfolgen erleiden:

$$\begin{aligned}\|Q - U_{S_n}\| &= \frac{1}{2} \sum_{\pi \in S_n} \left| Q(\{\pi\}) - \frac{1}{n!} \right| \\ &= \frac{1}{2} \left(\frac{n!}{2} \left| \frac{2}{n!} - \frac{1}{n!} \right| + \frac{n!}{2} \left| 0 - \frac{1}{n!} \right| \right) \\ &= \frac{1}{2}.\end{aligned}$$

Die Totalvariation liefert den Wert $1/2$. Nach Mann ist dieser Wert innerhalb der Spanne von 0 bis 1 vergleichsweise groß, obwohl sich Karte i an einer zufälligen Position in der oberen Hälfte des Kartenstapels befindet und der Kartenstapel wegen der zufälligen Reihenfolge aller übrigen Karten von ihm als hinreichend gemischt angesehen werden kann.

Wir geben allerdings zu bedenken, dass die Beschränkung auf nur die Hälfte aller Permutationen in S_n eine hinreichende Mischung der Karten verhindert. Das folgende Beispiel soll diesen Einwand motivieren: Ein fairer Würfel werde einmal geworfen, das Ergebnis des Wurfes sei gerade, aber über diese Information hinaus unbekannt. Jede der Zahlen 2, 4, 6 ist dann gleich wahrscheinlich mit Wahrscheinlichkeit $1/3$. Die richtige Augenzahl kann also mit Wahrscheinlichkeit $1/3$ geraten werden. Sei U die Gleichverteilung auf $\{1, \dots, 6\}$ und \bar{U} die Gleichverteilung auf $\{2, 4, 6\}$. Dann gilt für die Totalvariation wie im obigen Beispiel

$$\begin{aligned}\|\bar{U} - U\| &= \frac{1}{2} \left(3 \left| \frac{1}{3} - \frac{1}{6} \right| + 3 \left| 0 - \frac{1}{6} \right| \right) \\ &= \frac{1}{2}.\end{aligned}$$

Die Werte $\frac{3}{2} \log_2 n$, $\lfloor \frac{3}{2} \log_2 n \rfloor$ und $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ für gebräuchliche Kartenstapelgrößen sind in Tabelle 2.5 angegeben.

Die tatsächlichen Anzahlen an Mischvorgängen, die notwendig sind, um die Werte $1/2$ und $1/(2e)$ in Totalvariation zu unterschreiten, weichen von den in Tabelle 2.5 angegebenen Werten ab. Die Abweichungen resultieren daraus, dass

n	24	32	48	52	104
$\frac{3}{2} \log_2 n$	6.88	7.50	8.38	8.55	10.05
$\lfloor \frac{3}{2} \log_2 n \rfloor$	6	7	8	8	10
$\lfloor \frac{3}{2} \log_2 n \rfloor - 2$	4	5	6	6	8

Tabelle 2.5: $\frac{3}{2} \log_2 n$, $\lfloor \frac{3}{2} \log_2 n \rfloor$ und $\lfloor \frac{3}{2} \log_2 n \rfloor - 2$ für $n = 24, 32, 48, 52, 104$.

Theorem 2.6.9 eine asymptotische Aussage ist, die nur für große n von Bedeutung ist. Die tatsächlichen Werte ergeben sich aus einer exakten Berechnung der Totalvariation (siehe Tabelle 2.2 auf Seite 50) und sind in Tabelle 2.6 angegeben. An dieser Stelle soll nun auch die Aussage „In card shuffling, 7 is winning number“ aus der Einleitung präzisiert werden. Für einen Kartenstapel mit 52 Karten sind 7 Riffle Shuffle notwendig, um die Entfernung 1/2 zur Gleichverteilung auf S_{52} in Totalvariation gerade zu unterschreiten.

n	24	32	48	52	104
$1/(2e)$	7	7	8	8	10
$1/2$	5	6	7	7	8

Tabelle 2.6: Minimale Anzahl an Mischvorgängen m mit $\|Q_n^{*(m)} - U_{S_n}\| < S$, $S = 1/(2e), 1/2$ für $n = 24, 32, 48, 52, 104$.

Symboverzeichnis

\mathbb{N}	$= \{1, 2, 3, \dots\}$
\mathbb{N}_0	$= \mathbb{N} \cup \{0\}$
$\mathbb{N}_{\geq n}$	$= \{k \in \mathbb{N} \mid k \geq n\}, n \in \mathbb{N}$
$\mathbb{N}_{\leq n}$	$= \{1, \dots, n\}, n \in \mathbb{N}$
$\mathbb{N}_{0,n-1}$	$= \{0, \dots, n-1\}, n \in \mathbb{N}$
B_{\neq}^n	$= \{(x_1, \dots, x_n) \in B^n \mid x_i \neq x_j \forall i \neq j\}, B \subset \mathbb{R}, n \in \mathbb{N}$
$ G $	Mächtigkeit der endlichen Menge G
$\mathfrak{P}(G)$	Potenzmenge der Menge G
\sum_n^a	$= \{(j_1, \dots, j_a) \in \mathbb{N}_{0,n}^a \mid \sum_{k=1}^a j_k = n\}, a, n \in \mathbb{N}$
\mathfrak{Z}	$= \{2^j \mid j \in \mathbb{Z}\}$
\mathfrak{B}^n	Borelsche σ -Algebra über \mathbb{R}^n
\mathfrak{B}_C^n	Spur- σ -Algebra von \mathfrak{B}^n unter $C \in \mathfrak{B}^n$
δ_x	Dirac-Maß in x
$M(n, 1/a)$	Multinomialverteilung mit Parametern n und a -mal $1/a$
Φ	Verteilungsfunktion der Standardnormalverteilung
ϕ	Dichte der Standardnormalverteilung
$Poi(t)$	Poisson-Verteilung mit Parameter t
$Q^{*(m)}$	m -fache Faltung der Verteilung Q
$Q_{a,n}$	GSR- (a, n) -Mischnmethode, $n \in \mathbb{N}, a \geq 2$
Q_n	GSR- $(2, n)$ -Mischnmethode, $n \in \mathbb{N}$
$R(0, 1)$	Gleichverteilung auf $[0, 1]$
$\sigma(X_0, \dots, X_m)$	von den Zufallsvariablen X_0, \dots, X_m erzeugte σ -Algebra
$\text{supp}(Q)$	$= \{x \in \Omega \mid Q(x) > 0\}, Q$ diskretes Wahrscheinlichkeitsmaß

	auf $(\Omega, \mathfrak{P}(\Omega))$
U_G	$= \frac{1}{ G } \sum_{x \in G} \delta_x$, G endliche Menge
\mathcal{S}_n	$= \{\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\} \mid \pi \text{ bijektiv}\},$
$\pi = [i_1, \dots, i_r]_Z$	$\pi(i_j) = i_{j+1}$, $j = 1, \dots, r-1$, $\pi(i_r) = i_1$ und $\pi(i) = i$ für $i \neq i_1, \dots, i_r$
\mathcal{A}_n	$= \{\pi \in \mathcal{S}_n \mid \text{sgn}(\pi) = 1\}$
$\mathcal{D}_{n,k}$	$= \{\pi \in \mathcal{S}_n \mid \pi \text{ hat } k \text{ Sprungstellen}\}, k = 1, \dots, n-1$
$\mathcal{R}_{n,k}$	$= \{\pi \in \mathcal{S}_n \mid \pi \text{ hat } k \text{ aufsteigende Sequenzen}\}, k = 1, \dots, n$
\mathcal{T}_n	$= \{[i, j]_Z \in \mathcal{S}_n \mid i, j \in \{1, \dots, n\}, i \neq j\},$
\mathcal{V}_4	$= \{\text{id}, [1, 2]_Z [3, 4]_Z, [1, 3]_Z [2, 4]_Z, [1, 4]_Z [2, 3]_Z\},$
$\mathcal{A}_{n,k}$	$= \mathcal{R}_{n,k} \text{ falls } k = 1, \dots, n \text{ und } = 0 \text{ sonst}$
T^-	$\{x^{-1} \mid x \in T\}, T \subset G, G \text{ Gruppe}$
$\langle T \rangle$	$\bigcup_{n \in \mathbb{N}_0} \{x_1 \circ \dots \circ x_n \mid x_i \in T \cup T^-, i = 1, \dots, n\}, T \subset G,$ $G \text{ Gruppe}$
$\mathbf{f}_n \sim \mathbf{g}_n$	$\lim_{n \rightarrow \infty} \mathbf{f}_n / \mathbf{g}_n = 1$
$\mathbf{f}_n \in o(\mathbf{g}_n)$	$\lim_{n \rightarrow \infty} f_n / g_n = 0$
$\mathbf{f}_n \in O(\mathbf{g}_n)$	$\exists n_0 \in \mathbb{N}, K \in [0, \infty) : \mathbf{f}_n / \mathbf{g}_n \leq K \forall n \geq n_0$
$\mathbf{f}_n \in O_c(\mathbf{g}_n)$	$\forall c \in \mathbb{R} : \mathbf{f}_n = \mathbf{f}_{n,c}, \mathbf{g}_n = \mathbf{g}_{n,c} \text{ und } \exists n_0(c) \in \mathbb{N},$ $K_c \in [0, \infty) : \mathbf{f}_{n,c} / \mathbf{g}_{n,c} \leq K_c \forall n \geq n_0(c)$
\log_2	Logarithmus zur Basis 2
$x \vee y$	$= \max\{x, y\}$
$x \wedge y$	$= \min\{x, y\}$
$\lfloor x \rfloor$	$= \max\{k \in \mathbb{Z} \mid k \leq x\}$, untere Gauss-Klammer
$\lceil x \rceil$	$= \min\{k \in \mathbb{Z} \mid k \geq x\}$, obere Gauss-Klammer

Literaturverzeichnis

- [1] Aldous, D. (1981/82): *Random Walks on finite groups and rapidly mixing Markov Chains*. Séminaire de Probabilités XVII. Lecture Notes in Mathematics 986, 243-297.
- [2] Aldous, D.; Diaconis P. (1986): *Shuffling cards and Stopping Times*. American Mathematical Monthly, Volume 93, No. 5, 333-348.
- [3] Alsmeyer, G. (2003): *Wahrscheinlichkeitstheorie*. Skripten zur Mathematischen Statistik Nr. 30, 3. Auflage, Universität Münster.
- [4] Alsmeyer, G. (2002): *Stochastische Prozesse Teil 1*. Skripten zur Mathematischen Statistik Nr. 33, 2. Auflage, Universität Münster.
- [5] Brémaud, P. (1999): *Markov Chains, Gibbs Fields, Monte Carlo Simulation and Queues*. Springer, New York.
- [6] Bayer, D.; Diaconis, P. (1992): *Trailing the Dovetail Shuffle to its lair*. Annals of Applied Probability, Volume 2, No. 2, 294-313.
- [7] Carlitz, L. (1959): *Eulerian Numbers and Polynomials*. Mathematics Magazine, Volume 32, No. 5, 247-260.
- [8] Carlitz, L.; Kurtz, D. C.; Scoville, R.; Stackelberg, O. P. (1972): *Asymptotic properties of Eulerian Numbers*. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 23, 47-54.

- [9] Diaconis, P. (1988): *Group Representations in probability and statistics*. Lecture Notes-Monograph Series 11, Institute of Mathematical Statistics, Hayward, California.
- [10] Diaconis, P. (1996): *The Cutoff Phenomenon in finite Markov Chains*. Proceedings of the National Academy of Science, Volume 93, 1659-1664.
- [11] Feller, W. (1971): *An introduction to Probability Theory and its applications*. Volume II, 2. Auflage, John Wiley and Sons, Inc.
- [12] Gilbert, E. (1955): *Theory of Shuffling*. Technical Memorandum, Bell Laboratories.
- [13] Kolata, G. (1990): *In shuffling cards, 7 is winning number*. New York Times, Spätausgabe vom 09.01.1990.
- [14] Lalley, S. P. (1999): *Riffle Shuffles and their associated Dynamical Systems*. Journal of Theoretical Probability, Volume 12, No. 4, 903-932.
- [15] Lorenz, F. (1996): *Einführung in die Algebra*. 3. Auflage, Spektrum Verlag, Heidelberg/Berlin.
- [16] Mann, B. (1995): *How many times should you shuffle a deck of cards?* Topics in Contemporary Probability and Its Applications, Ed. J. Laurie Snell, CRC Press Boca Raton, 261-289.
- [17] Reeds, J. (1981): *Theory of Riffle Shuffling*. Unveröffentlichtes Manuskript.
- [18] Rogers, L. C. G.; Pitman, J. W. (1981): *Markov Functions*. Annals of Probability, Volume 9, No. 4, 573-582.
- [19] Salama, I. A.; Kupper L. L. (1986): *A geometric interpretation for the Eulerian Numbers*. American Mathematical Monthly, Volume 93, No. 1, 51-52.
- [20] Saloff-Coste, L. (2004): *Random Walks on finite groups*. Probability on discrete structures, Encyclopaedia of Mathematical Science, Volume 110 (Harry Kesten, Editor), Springer, Berlin, 263-346.

- [21] Saloff-Coste, L. (1997): *Lectures on finite Markov Chains*. Lecture Notes in Mathematics 1665, Lectures on Probability Theory and Statistics, Ecole d'Eté de Probabilités de Saint-Flour XXVI-1996, E. Giné et al., 301-413.
- [22] Stanley, R. P. (1997): *Enumerative combinatorics*. Cambridge studies in advanced mathematics, Volume 49, Cambridge University Press, Cambridge.
- [23] Tanny, S. (1973): *A probabilistic interpretation of the Eulerian Numbers*. Duke Mathematical Journal, Volume 40, 717-722. Corrigenda (1974), Volume 41, 689.
- [24] Wilson, D. B. (2004): *Mixing times of Lozenge Tiling and card shuffling Markov Chains*. Annals of Applied Probability, Volume 14, No. 1, 274-325.
- [25] Woess, W. (1980): *Aperiodische Wahrscheinlichkeitsmaße auf topologischen Gruppen*. Monatshefte Mathematik 90, 339-345.
- [26] Worpitzky, J. (1883): *Studien über die Bernoullischen Zahlen und Eulerschen Zahlen*, Journal für die reine und angewandte Mathematik 94, 203-232.

Ich versichere, dass ich die Diplomarbeit selbständig 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 2005

(Walter Peter Sendfeld)

