

WESTFÄLISCHE WILHELMS-UNIVERSITÄT MÜNSTER
FACHBEREICH MATHEMATIK UND INFORMATIK

Diplomarbeit

**Die Eindeutigkeit der stationären
Verteilung des gleichförmigen
Split-Merge-Prozesses**

Julian Hofrichter

April 2007

Betreut durch Prof. Dr. G. Alsmeyer

Diese Arbeit wurde mithilfe von \LaTeX und KOMA-Script gesetzt.

Danksagung

Mein herzlicher Dank gilt allen, die zum Gelingen dieser Arbeit beigetragen haben. Insbesondere danke ich Herrn Professor Dr. G. Alsmeyer für die Vergabe dieses interessanten und vielseitigen Themas sowie für die umfassende Betreuung. Mit wertvollen Hinweisen half er mir über manche Hürde hinweg.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Koagulations-Fragmentierungs-Prozesse	1
1.2	Split-Merge-Prozesse auf Intervallpartitionen	2
1.3	Invariante Verteilungen von CCF und $\text{DCF}^{(n)}$	4
2	Kontinuierliche und diskrete Koagulation – Fragmentierung	9
2.1	Präliminarien	9
2.2	Der CCF-Prozess	9
2.3	A-priori-Momentenabschätzung	10
2.4	Der $\text{DCF}^{(n)}$ -Prozess	10
2.5	Der Übergangskern von $\text{DCF}^{(n)}$	11
2.6	Die invariante Verteilung von $\text{DCF}^{(n)}$	13
3	Die Poisson-Dirichlet-Verteilung	15
3.1	Vorbemerkungen	15
3.2	Poisson-Prozesse auf allgemeinen Räumen	16
3.3	Definition der Poisson-Dirichlet-Verteilung	18
3.4	Der Poisson-Dirichlet-Grenzfall	21
3.5	Die <i>GEM</i> -Verteilung	25
3.6	Zusammenhang zwischen der <i>GEM</i> - und der Poisson-Dirichlet-Verteilung	26
3.7	Invarianz unter CCF	31
4	Kopplung von DCF und CCF	39
4.1	Der gemeinsame Zustandsraum $\Omega_{cde}^{(n)}$	40
4.2	Approximation von kontinuierlichen Partitionen durch diskrete	41
4.3	Markov-Kette auf $\Omega_{cde}^{(n)}$	43
4.4	Beweis des Theorems	50

5	Gruppendarstellungen und Charaktere	59
5.1	Darstellungen endlicher Gruppen	59
5.2	Charakter einer Darstellung	65
5.3	Anzahl der irreduziblen Darstellungen einer Gruppe	71
5.4	Darstellungen der S_n	74
5.5	$\text{DCF}^{(n)}$ als Markov-Kette auf S_n	83
6	$\text{DCF}^{(n)}$-Konvergenz	91
6.1	Vorbereitungen	91
6.2	Das Konvergenz-Theorem und Bezüge zur Darstellungstheorie	93
6.3	Beweis des Theorems	100
7	Beweis der Vermutung	111
A	Anhang	125
A.1	Beweis der A-priori-Momentenabschätzung	126
A.2	Charaktere und Darstellungen der S_{10}	133
	Literaturverzeichnis	135

1 Einleitung

In der vorliegenden Arbeit werden wir zeigen, dass die Poisson-Dirichlet-Verteilung die eindeutig bestimmte stationäre Verteilung des gleichförmigen Split-Merge-Prozesses ist. Bei diesem Prozess handelt es sich um eine spezielle Variante eines Koagulations-Fragmentierungs-Prozesses. Bevor wir das verwendete Modell beschreiben und das weitere Vorgehen skizzieren, erläutern wir zunächst, was allgemein unter einem Koagulations-Fragmentierungs-Prozess zu verstehen ist (vgl. [4]).

1.1 Koagulations-Fragmentierungs-Prozesse

Koagulations-Fragmentierungs-Prozesse beschreiben die zeitliche Entwicklung eines Systems von Partikeln, in dem sich die einzelnen Partikel aufspalten bzw. miteinander vereinigen können. Die Annahme, dass diese Übergänge zufällig sind, machen sie zu einem stochastischen Prozess.

In der Natur finden sich zahlreiche Beispiele für Fragmentation (Abbrechen von Eisschollen von einem Festeisblock, Brechen von Kristallen in der Kristallografie, Spalten von Atomkernen in der Kernphysik, Aufteilung von Grundstückspartzellen durch fortschreitende Zersiedelung, etc.) wie auch für Koagulation (Bildung von Polymerketten in der Chemie, Verschmelzung von Tröpfchen in Wolken, etc.). Insofern es sich dabei um Gleichgewichtsprozesse handelt, das heißt, falls die beschriebenen Übergänge in beide Richtungen ablaufen können, findet sowohl Fragmentierung wie auch Koagulation statt. Dieses Verhalten wird durch einen Koagulations-Fragmentierungs-Prozess beschrieben.

Um einen solchen Prozess modellieren zu können, sind geeignete Annahmen über den Prozess notwendig. Diese sind üblicherweise:

- (i) Die Evolution ist *gedächtnislos*, hängt also nur von dem Jetzt-Zustand des Prozesses ab. Es handelt sich somit um einen Markov-Prozess.

- (ii) Die Wahrscheinlichkeiten für die Übergänge hängen nur von *einem* Parameter, üblicherweise der Größe der Partikel, ab – nicht etwa von Form oder Struktur.
- (iii) Die Übergangswahrscheinlichkeiten sind unabhängig von der Position der Partikel.
- (iv) Die Partikel vereinigen bzw. spalten sich *unabhängig* von anderen Partikeln.

1.2 Split-Merge-Prozesse auf Intervallpartitionen

Nun stellen wir den in dieser Arbeit betrachteten Spezialfall eines Koagulations-Fragmentierungs-Prozesses vor: einen gleichförmigen Split-Merge-Prozess auf Partitionen des Einheitsintervalls. Diesen Prozess werden wir auch als *stetigen Koagulations-Fragmentierungs-Prozess* (*continuous coagulation-fragmentation, CCF*) bezeichnen (vgl. [7], S. 915).

Zunächst zerlegen wir das Einheitsintervall $[0, 1]$ in eine abzählbare Menge von disjunkten Teilintervallen derart, dass die Vereinigung aller Teilintervalle wieder den gesamten Raum ergibt. Als Partition bezeichnen wir sodann die Folge der Längen aller Teilintervalle. Den Raum der abzählbaren, nichtaufsteigend geordneten Partitionen von $[0, 1]$ bezeichnen wir mit Ω_1 , d. h.

$$\Omega_1 := \{p \in [0, 1]^{\mathbb{N}} : p_1 \geq p_2 \geq \dots \geq 0, |p|_1 = 1\}. \quad (1.1)$$

$|x|_1 = \sum_{i=1}^{\infty} |x_i|$ bezeichnet hierbei die l_1 -Norm einer Folge $(x_i)_{i \geq 1}$. Die einzelnen Teilintervalle p_j von p werden wir als *Abschnitte* bezeichnen.

Diese Abschnitte stellen nunmehr die Partikel dar, die koagulieren und fragmentieren: Der CCF-Prozess operiert auf einer Intervallpartition, indem er Abschnitte aufteilt (Split) bzw. verschiedene Abschnitte zu einem neuen Abschnitt vereinigt (Merge). Die Übergangswahrscheinlichkeiten hängen dabei nur der Größe des aufzuspaltenden Abschnitts bzw. der zu vereinigenden Abschnitte, d. h. der Länge, ab. Die Bezeichnung *gleichförmiger* Split-Merge-Prozess besagt, dass es sich dabei um eine lineare Abhängigkeit handelt. Diesem Umstand trägt die Definition des längenverzerrten Auswählens Rechnung: Ein zufälliges Auswählen von Abschnitten aus

$p \in \Omega_1$ heißt *längenverzerrt*, wenn der j -te Abschnitt genau mit Wahrscheinlichkeit p_j ausgewählt wird.

Somit definieren wir die *Transition des CCF-Prozesses* wie folgt: Wir wählen längenverzerrt zwei Abschnitte von $p \in \Omega_1$ (mit Zurücklegen); wird zweimal derselbe Abschnitt gewählt, so wird dieser aufgeteilt und der Teilungspunkt gleichmäßig aus dem betreffenden Abschnitt gewählt (Split). Werden zwei verschiedene Abschnitte gewählt, so werden sie miteinander verbunden (Merge). In beiden Fällen wird die Partition anschließend neu geordnet.

Offenbar entspricht dieser so definierte Prozess allen Modellannahmen, insbesondere handelt es sich um einen Markov-Prozess.

Um Aussagen über den CCF-Prozess zu erhalten, werden wir auf sein diskretes Analogon zurückgreifen: einen Split-Merge-Prozess auf Partitionen der Menge $\{1, \dots, n\}$, den wir im Folgenden auch als *diskreten Koagulations-Fragmentierungs-Prozess* (*discrete coagulation-fragmentation*, $DCF^{(n)}$) bezeichnen (mit $n \in \mathbb{N}$; vgl. [7], S. 915).

Dazu zerlegen wir $\{1, \dots, n\}$ in disjunkte Teilmengen. Die Folge der Elementanzahlen dieser Teilmengen verstehen wir als diskrete Partition und bezeichnen den Raum der nichtaufsteigend geordneten diskreten Partitionen von $\{1, \dots, n\}$ mit \mathcal{P}_n , d. h.

$$\mathcal{P}_n := \{\ell = (\ell_i)_{i \geq 1} \in \{0, 1, \dots, n\}^{\mathbb{N}} : \ell_1 \geq \ell_2 \geq \dots \geq 0, |\ell|_1 = n\} \subset n\Omega_1. \quad (1.2)$$

Hierbei bezeichnet $|x|_1 = \sum_{i=1}^{\infty} |x_i|$ wiederum die l_1 -Norm. Elemente aus \mathcal{P}_n lassen sich als Folge natürlicher Zahlen wie bspw. $(7, 5, 3, 3, 1, 1, 0, \dots)$ darstellen, wobei die restlichen Einträge notwendigerweise verschwinden. Insgesamt addieren sich nach Konstruktion die Längen der einzelnen Abschnitte gerade zu n .

Offenbar können wir Partitionen aus \mathcal{P}_n auch als Partitionen von $[0, n]$ bzw. $[0, 1]$ interpretieren: Jedes *Atom* $x \in \{1, \dots, n\}$ repräsentiert ein Teilintervall der Länge 1 von $[0, n]$. Gehören k Atome zu einem Abschnitt, so entspricht dieser Abschnitt einem Teilintervall der Länge k von $[0, n]$. Aus einer solchen Partition erhalten wir wiederum eine Partition von $[0, 1]$, wenn wir mit dem Faktor $1/n$ skalieren. \mathcal{P}_n ist somit äquivalent zu einer diskreten Version von Ω_1 , in der die Partitionen auf einem endlichen äquidistanten Gitter in $[0, 1]$ liegen.

Um eine Partition $\ell \in \mathcal{P}_n$ zu charakterisieren, definieren wir weiter ihren *Typ*

$$N_\ell(k) := \#\{i : \ell_i = k\} \quad \text{für } k = 1, \dots, n, \quad (1.3)$$

der die Anzahl der jeweiligen Abschnitte mit Länge k angibt. Durch $(N_\ell(k))_{k=1, \dots, n}$ ist eine Partition bereits vollständig festgelegt. Die *Gesamtzahl der Abschnitte* von ℓ bezeichnen wir mit

$$N_\ell = \sum_{k=1}^n N_\ell(k). \quad (1.4)$$

Die *Transition des DCF⁽ⁿ⁾-Prozesses* definieren wir nun wie folgt: Wir wählen gleichmäßig unabhängig und ohne Zurücklegen zwei Atome x, y aus $\{1, \dots, n\}$. Liegen x und y in demselben Abschnitt, so wird dieser gleichmäßig aufgespaltet. Liegen x und y in zwei verschiedenen Abschnitten, so werden beide Abschnitte zu einem neuen Abschnitt vereinigt. Anschließend wird die Partition in beiden Fällen neu sortiert.

Offenbar liefert dies wiederum einen Markov-Prozess, der den oben aufgeführten Modellannahmen genügt.

1.3 Invariante Verteilungen von CCF und DCF⁽ⁿ⁾

Es ist bekannt, dass der CCF-Prozess eine invariante Verteilung besitzt, nämlich die Poisson-Dirichlet-Verteilung (mit Parameter $\theta = 1$) \mathcal{PD}_1 . Dies werden wir in Kapitel 3, wo wir auch eine Definition der Poisson-Dirichlet-Verteilung geben, noch einmal nachweisen. Allerdings ist nicht a priori klar, dass diese schon die einzige ist: Da der Zustandsraum überabzählbar ist, folgt keinesfalls von selbst, dass eine invariante Verteilung schon eindeutig bestimmt ist.

Im Falle von DCF⁽ⁿ⁾ hingegen ist die Sachlage einfacher: DCF⁽ⁿ⁾ besitzt eine eindeutige bestimmte invariante Verteilung, wie wir in Kapitel 5 zeigen werden. Diese Verteilung lässt sich vergleichsweise einfach bestimmen: Wir können \mathcal{P}_n auch als Menge der Konjugationsklassen der symmetrischen Gruppe S_n interpretieren; DCF⁽ⁿ⁾ beschreibt dann die Evolution des Zyklentyps einer Permutation $\pi \in S_n$ unter der Aktion von zufälligen Transpositionen. So können wir zeigen, dass die durch die

Ewens'sche Stichproben-Formel

$$\pi_S^{(n)}(\ell) = \left(\prod_{k=1}^n k^{N_\ell(k)} N_\ell(k)! \right)^{-1}, \quad \ell \in \mathcal{P}_n \quad (1.5)$$

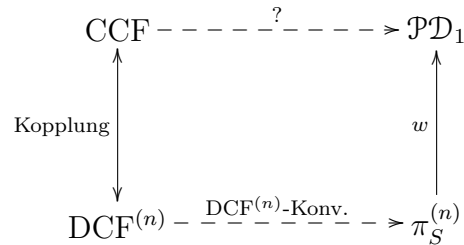
beschriebene Verteilung invariant unter dem DCF⁽ⁿ⁾-Prozess ist. Da der Zustandsraum \mathcal{P}_n endlich ist, folgt aus der Theorie diskreter Markov-Ketten direkt die Eindeutigkeit der stationären Verteilung.

Da nun CCF offenbar der Grenzprozess von DCF⁽ⁿ⁾ für $n \rightarrow \infty$ ist (bei Skalierung der diskreten Partitionen mit einem Faktor $1/n$) und die (eindeutigen) invarianten Verteilungen der skalierten DCF⁽ⁿ⁾-Prozesse $\pi_S^{(n)}(n \cdot)$ für $n \rightarrow \infty$ schwach gegen \mathcal{PD}_1 konvergieren (vgl. etwa [16]), liegt für CCF die folgende Vermutung nahe (vgl. [7], S. 916):

1.1 Vermutung (Vershik). \mathcal{PD}_1 ist die eindeutig bestimmte invariante Verteilung des CCF-Prozesses.

Sie wurde erstmalig von A. M. Vershik im Jahre 1977 formuliert (vgl. [20]). Das Ziel der vorliegenden Arbeit wird nun sein, einen Beweis dafür anzugeben.

Die Grundidee ist, die Verbindung zwischen DCF⁽ⁿ⁾ und CCF auszunutzen und so zu zeigen, dass sich die Eindeutigkeit der stationären Verteilung von DCF⁽ⁿ⁾ auf CCF überträgt. Dazu konstruieren wir einen CCF- und einen DCF⁽ⁿ⁾-Prozess derart, dass (bei Skalierung der diskreten Partitionen um den Faktor $1/n$) eine geeignete Ähnlichkeitsbeziehung zwischen beiden Prozessen besteht. Dies bezeichnen wir als *Kopplung*. Die Verteilung des (skalierten) DCF⁽ⁿ⁾-Prozesses konvergiert für jede Anfangsverteilung gegen $\pi_S^{(n)}(n \cdot)$, was wiederum aufgrund der Konvergenz $\pi_S^{(n)}(n \cdot) \xrightarrow{w} \mathcal{PD}_1$ für große n nahe an der Poisson-Dirichlet-Verteilung liegt. So können wir mit $n \rightarrow \infty$ zeigen, dass die Verteilung des CCF-Prozesses für beliebige Anfangsverteilungen gegen \mathcal{PD}_1 konvergiert, woraus unmittelbar folgt, dass jede invariante Verteilung schon mit der Poisson-Dirichlet-Verteilung übereinstimmt (vgl. etwa [1], S. 20 f.).



Um auf diese Weise ein Resultat zu erzielen, sind allerdings weitere technische Einschränkungen nötig, da sich die Konvergenzrate von $\text{DCF}^{(n)}$ mit wachsendem n verschlechtert. Die Verteilung von $\text{DCF}^{(n)}$ befindet sich für große n also erst ab einem relativ späten Zeitpunkt hinreichend nahe an $\pi_S^{(n)}$. Andererseits aber wird die Kopplung $\text{CCF} \leftrightarrow \text{DCF}^{(n)}$ mit weiterem Fortschreiten der Prozesse immer schwächer, so dass sich zu späten Zeitpunkten CCF und $\text{DCF}^{(n)}$ unter Umständen schon deutlich voneinander entfernt haben. Somit würde das beschriebene Verfahren nicht funktionieren.

Um diese Schwierigkeit zu überwinden, ist es nötig, die Prozesse zu einem Zeitpunkt zu betrachten, wo *sowohl* ausreichend gute Kopplung zwischen $\text{DCF}^{(n)}$ und CCF vorliegt *als auch* $\text{DCF}^{(n)}$ sich hinreichend nahe an seiner Gleichverteilung befindet. Dazu benötigen wir neben einer guten Kopplung vor allem eine schnelle $\text{DCF}^{(n)}$ -Konvergenz. Diese beiden Resultate bereitzustellen, wird jeweils ein Hauptaspekt der vorliegenden Arbeit sein. Der Schlüssel ist in beiden Fällen, als Startverteilung nur solche Verteilungen zuzulassen, die wenig Masse auf Partitionen mit vielen kleinen Abschnitten legen. Wir werden zeigen, dass durch diese Einschränkung die Zahl der in Frage kommenden Verteilungen nicht verringert wird, da jede CCF-invariante Verteilung von vornherein diese Eigenschaft hat.

Im Detail werden wir folgendermaßen vorgehen:

- Wir stellen eine A-priori-Momentenabschätzung bereit, die zeigt, dass jede CCF-invariante Verteilung zu einer guten Kontrolle der Anzahl der „kleinen Abschnitte“ führt.
- Wir koppeln CCF mit $\text{DCF}^{(n)}$, d. h. wir approximieren kontinuierliche Partitionen durch diskrete. Starten wir mit CCF mit einer Anfangsverteilung, die über

eine obige Kontrolle des Partitionsendes verfügt, lässt sich zeigen, dass bis zu einem Zeitpunkt in der Größenordnung von \sqrt{n} Kopplung vorliegt.

3. Für Anfangsbedingungen wie oben werden wir zeigen können, dass auf einem geeigneten Mengensystem ausreichend schnelle $DCF^{(n)}$ -Konvergenz vorliegt.

Im nächsten Kapitel werden wir zunächst die beiden Prozesse CCF und $DCF^{(n)}$ genau definieren und die A-priori-Abschätzung formulieren (Abschnitt 2.3). Es folgt Kapitel 3, in dem wir die Poisson-Dirichlet-Verteilung definieren sowie ihre Invarianz unter CCF zeigen. Im sich anschließenden Kapitel 4 stellen wir sodann die Kopplung von CCF und $DCF^{(n)}$ dar. Da die schnelle $DCF^{(n)}$ -Konvergenz in Kapitel 6 Resultate aus der Darstellungstheorie endlicher Gruppen verwendet, werden wir zuvor in Kapitel 5 den benötigten theoretischen Hintergrund bereitstellen. In Kapitel 7 findet sich schließlich der eigentliche Beweis der Vermutung unter Verwendung aller Hilfsmittel.

2 Kontinuierliche und diskrete Koagulation – Fragmentierung

In diesem Kapitel wollen wir die Prozesse CCF und $\text{DCF}^{(n)}$ exakt definieren und einige ihrer Eigenschaften notieren (vgl. [7], S. 917 f.). Zunächst sind einige Vorbereitungen notwendig.

2.1 Präliminarien

Zu einem gegebenen topologischen Raum (W, \mathfrak{T}) bezeichnen wir die zugehörige Borel'sche σ -Algebra, d. h. die kleinste σ -Algebra, die alle Mengen aus \mathfrak{T} umfasst, mit \mathcal{B}_W . Den Raum der Wahrscheinlichkeitsmaße auf (W, \mathcal{B}_W) bezeichnen wir mit $\mathcal{M}_1(W)$. In nicht ganz korrekter Notation werden wir den Teilraum von $\mathcal{M}_1(W)$, der alle Wahrscheinlichkeitsmaße enthält, deren Träger in einer abgeschlossenen Teilmenge V von W liegt, ebenso mit $\mathcal{M}_1(V)$ bezeichnen. Die Totalvariation eines Maßes $\nu \in \mathcal{M}_1(W)$ bezeichnen wir mit $\|\nu\| := \sup_{A \in \mathcal{B}_W} |\nu(A)|$.

Das in Kapitel 1.2 definierte Ω_1 stattdessen wir mit der durch die l_1 -Norm $|\cdot|_1$ induzierten Topologie aus und bezeichnen diese mit \mathfrak{T}_1 . Die Topologie stimmt auf Ω_1 mit der Topologie der koordinatenweisen Konvergenz, d. h. der schwachen Topologie, überein.

Schließlich bezeichnen wir die Gleichverteilung auf einem Intervall $[a, b] \subset \mathbb{R}$ mit $U[a, b]$.

2.2 Der CCF-Prozess

Auf Ω_1 definieren wir nun die CCF-Markov-Kette $(p(k))_{k \geq 0}$, deren Konstruktion in der Einleitung bereits kurz umrissen wurde. Als Anfangsverteilung wählen wir ein

beliebiges $\nu \in \mathcal{M}_1(\Omega_1)$. Um aus dem Zustand $p = p(k)$ zu einem gegebenen Zeitpunkt k in den nächsten Zustand $p(k+1)$ zu springen, werden aus p zwei Abschnitte p_i und p_j durch längenverzerrtes Auswählen mit Zurücklegen bestimmt. Für $i \neq j$ werden p_i und p_j zu einem neuen Abschnitt der Länge $p_i + p_j$ zusammengeführt (Koagulation/Merge); für $i = j$ wird p_i in zwei neue Abschnitte mit den Längen up_i und $(1-u)p_i$ zerlegt mit $u \sim U[0, 1]$ und unabhängig von allen anderen Zufallsvariablen (Fragmentierung/Split). In beiden Fällen wird die Partition anschließend in nichtaufsteigender Reihenfolge neu angeordnet.

2.3 A-priori-Momentenabschätzung

Wie schon in der Einleitung erwähnt, wird für unsere Argumentation entscheidend sein, bereits im Vorfeld zeigen zu können, dass jede CCF-invariante Verteilung nicht zu viel Gewicht auf Partitionen mit vielen kleinen Abschnitten legt (vgl. [7], S. 918):

2.1 Proposition. *Sei $\mu \in \mathcal{M}_1(\Omega_1)$ CCF-invariant. Dann gilt für alle $\alpha > 2/5$*

$$\int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) < \infty. \quad (2.1)$$

Für $\alpha = 1$ gilt $\sum_{i \geq 1} p_i^\alpha = \sum_{i \geq 1} p_i \equiv 1$, und damit nimmt auch das Integral den Wert 1 an, ist also insbesondere endlich. Verkleinern wir α nun, so werden die einzelnen Abschnitte skaliert, wobei kleine Abschnitte deutlich stärker vergrößert werden als Abschnitte, deren Länge in der Nähe von 1 liegt. Bleibt der Wert des Integrals selbst für kleine α endlich, so muss μ auf Partitionen mit vielen kleinen Abschnitten entsprechend wenig Masse aufweisen.

Ein Beweis der A-priori-Momentenabschätzung findet sich im Anhang A.1.

2.4 Der DCF⁽ⁿ⁾-Prozess

Für $n \in \mathbb{N}$ konstruieren wir nun eine DCF⁽ⁿ⁾-Markov-Kette $(\ell(k))_{k \geq 0}$ auf \mathcal{P}_n , die in ähnlicher Weise wie die CCF-Kette definiert wird. Als Anfangsverteilung wählen wir wiederum ein beliebiges $\mu \in \mathcal{M}_1(\mathcal{P}_n)$. Für den Übergang von einem Zustand $\ell = \ell(k)$ in den folgenden $\ell(k+1)$ identifizieren wir jedes $\ell \in \mathcal{P}_n$ wieder mit

einer Zerlegung $\bigcup_i A_i$ von $\{1, \dots, n\}$ in disjunkte Teilmengen A_i . Dabei sind die A_i so zu wählen, dass A_i für alle i genau ℓ_i Elemente enthält. Wir wählen nun unabhängig voneinander zwei Elemente x und y gleichmäßig aus $\{1, \dots, n\}$ ohne Zurücklegen; es sei $x \in A_i$ und $y \in A_j$. Falls $i \neq j$, ersetze A_i und A_j durch $A_i \cup A_j$. Falls $i = j$, so gilt $\ell_i \geq 2$ wegen $x \neq y$, und wir können A_i in zwei Teilmengen aufteilen. Die Teilmengen werden dabei so gewählt, dass die eine die k kleinsten Elemente von A_i enthält und die andere die $\ell_i - k$ übrigen Elemente, wobei k gleichmäßig aus $\{1, \dots, \ell_i - 1\}$ und unabhängig von x und y gewählt wird. Alle anderen Mengen bleiben jeweils unverändert. Die nichtaufsteigende Folge der (entsprechend geänderten) Elementanzahlen aller Teilmengen von $\{1, \dots, n\}$ ergibt schließlich $\ell(k + 1)$.

Obschon die Ähnlichkeit zu CCF unverkennbar ist, bestehen dennoch einige Unterschiede: $DCF^{(n)}$ ist kein *gleichförmiger* Split-Merge-Prozess, da bei einem Übergang das zweite Atom y ohne Zurücklegen gewählt wird. Damit wird verhindert, dass ein Abschnitt mit nur einem Atom aufgespalten wird: Im Gegensatz zu CCF, wo Abschnitte beliebiger Länge aufgespalten werden können, kann beim $DCF^{(n)}$ -Prozess der Split-Übergang nur ausgeführt werden, wenn der entsprechende Abschnitt aus mindestens zwei Atomen besteht.

2.5 Der Übergangskern von $DCF^{(n)}$

Wir wollen nun den $DCF^{(n)}$ -Übergangskern mit seinen Werten auf $\mathcal{P}_n \times \mathcal{P}_n$ angeben (vgl. [7], S. 918 f.):

2.2 Satz. Für den $DCF^{(n)}$ -Übergangskern $\mathbb{M}^{(n)}: \mathcal{P}_n \times \mathcal{B}_{\mathcal{P}_n} \longrightarrow [0, 1]$ gilt

(1) für die Vereinigung verschieden großer Abschnitte bzw. das Aufspalten in verschieden große Abschnitte: Seien $\ell, \ell' \in \mathcal{P}_n$ mit $N_{\ell'}(j) = N_{\ell}(j) - 1$, $N_{\ell'}(k) = N_{\ell}(k) - 1$, $N_{\ell'}(j + k) = N_{\ell}(j + k) + 1$ für $j, k \in \{1, \dots, n\}$ mit $j + k \leq n$ und $N_{\ell'}(q) = N_{\ell}(q)$ für alle $q \notin \{j, k, j + k\}$, so ist

$$\mathbb{M}^{(n)}(\ell, \ell') = \frac{2jk}{n(n-1)} N_{\ell}(j) N_{\ell}(k) \quad (\text{Merge-Übergang}), \quad (2.2)$$

$$\mathbb{M}^{(n)}(\ell', \ell) = \frac{2(j+k)}{n(n-1)} N_{\ell'}(j+k) \quad (\text{Split-Übergang}). \quad (2.3)$$

(2) für die Vereinigung gleich großer Abschnitte bzw. das Aufspalten in gleich große Abschnitte: Seien $\ell, \ell' \in \mathcal{P}_n$ mit $0 \leq N_{\ell'}(j) = N_{\ell}(j) - 2$, $N_{\ell'}(2j) = N_{\ell}(2j) + 1$ für ein $j \in \{1, \dots, \lfloor \frac{n}{2} \rfloor\}$ und $N_{\ell'}(q) = N_{\ell}(q)$ für alle $q \notin \{j, 2j\}$, so ist

$$\mathbb{M}^{(n)}(\ell, \ell') = \frac{j^2}{n(n-1)} N_{\ell}(j) (N_{\ell}(j) - 1) \quad (\text{Merge-Übergang}), \quad (2.4)$$

$$\mathbb{M}^{(n)}(\ell', \ell) = \frac{2j}{n(n-1)} N_{\ell'}(2j) \quad (\text{Split-Übergang}). \quad (2.5)$$

Alle anderen Werte des Übergangskerns verschwinden.

Beweis. (1) Unter den gegebenen Voraussetzungen ist die Wahrscheinlichkeit für den *Merge*-Übergang von ℓ zu ℓ' gerade die Wahrscheinlichkeit, dass durch x ein Abschnitt der Länge j und durch y ein Abschnitt der Länge k ausgewählt wird. Dies ist jeweils gleich der Anzahl der entsprechenden Abschnitte $N_{\ell}(j)$ bzw. $N_{\ell}(k)$ mal Länge derselben, geteilt durch die Gesamtzahl an Möglichkeiten für x und y . Aufgrund der Symmetrie $x \leftrightarrow y$ erhalten wir noch einen Faktor 2. Es ergibt sich

$$\begin{aligned} \mathbb{M}^{(n)}(\ell, \ell') &= 2 \cdot \frac{j \cdot N_{\ell}(j)}{n} \cdot \frac{k \cdot N_{\ell}(k)}{n-1} \\ &= \frac{2jk}{n(n-1)} N_{\ell}(j) N_{\ell}(k). \end{aligned}$$

Der *Split*-Übergang von ℓ' zu ℓ erfolgt genau dann, wenn x und y beide in demselben Abschnitt der Länge $j+k$ liegen. Für dieses Ereignis gibt es genau $(j+k)N_{\ell'}(j+k) \cdot (j+k-1)$ Möglichkeiten. Weiter muss der gewählte Abschnitt passend in Abschnitte mit den Längen j und k aufgeteilt werden. Nach Konstruktion hat dieses Ereignis die Wahrscheinlichkeit $\frac{2}{j+k-1}$, wobei der Faktor 2 von der Symmetrie $j \leftrightarrow k$ herrührt. Es ergibt sich also

$$\begin{aligned} \mathbb{M}^{(n)}(\ell', \ell) &= 2 \cdot \frac{(j+k) \cdot N_{\ell'}(j+k)}{n} \cdot \frac{j+k-1}{n-1} \cdot \frac{1}{j+k-1} \\ &= \frac{2(j+k)}{n(n-1)} N_{\ell'}(j+k). \end{aligned}$$

(2) Unter den gegebenen Voraussetzungen ist die Wahrscheinlichkeit für den *Merge*-Übergang von ℓ zu ℓ' gerade die Wahrscheinlichkeit, dass x einen Abschnitt mit

der Länge j auswählt und y einen anderen Abschnitt derselben Länge, wofür es $jN_\ell(j) \cdot j(N_\ell(j) - 1)$ Möglichkeiten gibt. Wir erhalten

$$\begin{aligned}\mathbb{M}^{(n)}(\ell, \ell') &= \frac{j \cdot N_\ell(j)}{n} \cdot \frac{j \cdot (N_\ell(j) - 1)}{n - 1} \\ &= \frac{j^2}{n(n - 1)} N_\ell(j) (N_\ell(j) - 1).\end{aligned}$$

Der *Split*-Übergang von ℓ' zu ℓ erfolgt genau dann, wenn x und y beide in demselben Abschnitt der Länge $2j$ liegen und die Aufspaltung so erfolgt, dass zwei Abschnitte mit der Länge j entstehen. Für das erste Ereignis gibt es $2jN_{\ell'}(2j) \cdot (2j - 1)$ Möglichkeiten; das zweite hat die Wahrscheinlichkeit $\frac{1}{2j - 1}$. Wir erhalten also

$$\begin{aligned}\mathbb{M}^{(n)}(\ell', \ell) &= \frac{2j \cdot N_{\ell'}(2j)}{n} \cdot \frac{2j - 1}{n - 1} \cdot \frac{1}{2j - 1} \\ &= \frac{2j}{n(n - 1)} N_{\ell'}(2j).\end{aligned}$$

Nach Konstruktion von $DCF^{(n)}$ sind nur diese Übergänge möglich; die restlichen Einträge des Übergangskerns verschwinden daher. \square

2.6 Die invariante Verteilung von $DCF^{(n)}$

Eine Verteilung $\mu \in \mathcal{M}_1(\mathcal{P}_n)$ heißt *invariante* oder *stationäre* Verteilung des $DCF^{(n)}$ -Prozesses, wenn

$$\mu \mathbb{M}^{(n)} = \mu \tag{2.6}$$

gilt (vgl. [1], S. 20). Dies bedeutet, dass die Verteilung μ durch eine $DCF^{(n)}$ -Transition wieder in sich selbst überführt wird. Für CCF gilt Analoges.

Da der Zustandsraum \mathcal{P}_n endlich ist und der $DCF^{(n)}$ -Prozess offenbar irreduzibel, folgt aus der Theorie der diskreten Markov-Ketten schon, dass $DCF^{(n)}$ positiv rekurrent ist und eine eindeutig bestimmte invariante Verteilung besitzt (vgl. [1], S. 65).

Um diese zu bestimmen, nutzen wir aus, dass der $DCF^{(n)}$ -Prozess auch verstanden werden kann als die Evolution des Zyklentyps einer Permutation $\pi \in S_n$ unter der

Aktion von zufälligen Transpositionen, was wir in Kapitel 5.5 näher erläutern. Dort werden wir zeigen, dass die Ewens'sche Stichproben-Formel mit Parameter $\theta = 1$ (vgl. etwa [13], S. 11 f.)

$$\pi_S^{(n)}(\ell) = \left(\prod_{k=1}^n k^{N_\ell(k)} N_\ell(k)! \right)^{-1} = \left(\prod_{i=1}^n \ell_i \prod_{k=1}^n N_\ell(k)! \right)^{-1}, \quad \ell \in \mathcal{P}_n \quad (2.7)$$

die invariante Verteilung von $\text{DCF}^{(n)}$ definiert (vgl. Satz 5.37). In der Tat lässt sich nachrechnen, dass für den $\text{DCF}^{(n)}$ -Übergangskern die detaillierten Gleichgewichtsgleichungen (vgl. [1], S. 96)

$$\pi_S^{(n)}(\ell) \mathbb{M}^{(n)}(\ell, \ell') = \pi_S^{(n)}(\ell') \mathbb{M}^{(n)}(\ell', \ell) \quad (2.8)$$

für alle $\ell, \ell' \in \mathcal{P}_n$ erfüllt sind, woraus ebenfalls schon die Stationarität von $\pi_S^{(n)}$ folgt.

Weiter gilt, dass für alle $n \in \mathbb{N}$ die Verteilung eines $\text{DCF}^{(n)}$ -Prozesses $(\ell^{(n)}(k))_{k \geq 0}$ für jede Anfangsverteilung $\mu \in \mathcal{M}_1(\mathcal{P}_n)$ im zeitlichen Mittel gegen $\pi_S^{(n)}$ konvergiert (vgl. [1], S. 63)

$$\lim_{m \rightarrow \infty} \left\| \frac{1}{k} \sum_{k=0}^m P_\mu^{\ell^{(n)}(k)} - \pi_S^{(n)} \right\| = 0,$$

wobei $P_\mu^{\ell^{(n)}(k)}$ die Verteilung von $(\ell^{(n)}(k))_{k \geq 0}$ mit Anfangsverteilung μ zum Zeitpunkt k bezeichne. Aus diesem Grunde werden wir $\pi_S^{(n)}$ gelegentlich auch als die *Gleichgewichtsverteilung* von $\text{DCF}^{(n)}$ bezeichnen.

Obiges Resultat lässt sich sogar noch verschärfen, wenn zusätzlich Aperiodizität vorliegt (vgl. [1], S. 40). Dennoch ist offensichtlich, dass die Konvergenzrate im Allgemeinen noch von n abhängt, denn je größer die Partition ist, umso mehr Transitionen sind nötig, um sich der Gleichgewichtsverteilung anzunähern. Insbesondere bedeutet dies, dass sich – wie bereits in der Einleitung erwähnt – die $\text{DCF}^{(n)}$ -Konvergenz und die Kopplung $\text{CCF} \leftrightarrow \text{DCF}^{(n)}$ nicht ohne Weiteres miteinander kombinieren lassen. Für genauere Aussagen über die Konvergenzgeschwindigkeit verweisen wir auf [6], Kap. 3A und 3D. Dort wird gezeigt, dass in einem dem $\text{DCF}^{(n)}$ -Prozess ähnlichen Modell die Gleichgewichtsverteilung ungefähr nach $\frac{1}{2}n \log n + cn$ mit $c > 0$ Schritten erreicht ist. Wie wir in Kapitel 4 zeigen werden, können wir die Kopplung mit CCF jedoch nur in etwa bis zum Zeitpunkt \sqrt{n} aufrechterhalten.

3 Die Poisson-Dirichlet-Verteilung

Wie bereits erwähnt, stellt die Poisson-Dirichlet-Verteilung mit Parameter $\theta = 1$ eine invariante Verteilung des CCF-Prozesses dar. In diesem Kapitel werden wir diese Verteilung vorstellen sowie anhand einer weiteren Verteilung näher charakterisieren. Schließlich zeigen wir das benannte Invarianzresultat.

3.1 Vorbemerkungen

Bei der Poisson-Dirichlet-Verteilung handelt es sich um eine wenig intuitive Verteilung, deren Einführung einiger Vorarbeit bedarf. Wir stützen uns bei der folgenden Darstellung im Wesentlichen auf das Buch von Kingman [17].

Zur Motivation betrachten wir den Wahrscheinlichkeitsvektor einer diskreten Wahrscheinlichkeitsverteilung, d. h.

$$p = (p_1, \dots, p_n) \tag{3.1}$$

mit

$$p_j \geq 0 \quad \text{für } j = 1, \dots, n \quad \text{und} \quad \sum_{j=1}^n p_j = 1. \tag{3.2}$$

Der Raum aller Vektoren mit nichtnegativen, sich zu eins addierenden Komponenten wird auch der *n-dimensionale Simplex* Δ_n genannt. Nun besteht die Möglichkeit, dass ein solcher Vektor selbst zufällig ist. Ein einfaches, nichttriviales Beispiel für eine Wahrscheinlichkeitsverteilung auf Δ_n ist die *Dirichlet-Verteilung* $\mathcal{D}(\alpha_1, \dots, \alpha_n)$ mit $\alpha_j \geq 0$ für $j = 1, \dots, n$. Sie hat die Dichte

$$f(p_1, \dots, p_n) = \frac{\Gamma(\alpha_1 + \dots + \alpha_n)}{\Gamma(\alpha_1) \dots \Gamma(\alpha_n)} p_1^{\alpha_1-1} \dots p_n^{\alpha_n-1} \tag{3.3}$$

bezüglich des $(n-1)$ -dimensionalen Lebesgue-Maßes auf Δ_n .

Lässt man nun n gegen unendlich gehen, so geht Δ_n in den *unendlichdimensionalen Simplex* $\Delta_\infty := \{p \in [0, 1]^\mathbb{N} : |p|_1 = 1\}$ über. Hierbei stellt sich die Frage, ob und inwiefern eine Folge von Dirichlet-Verteilungen auf Δ_n gegen eine Grenzverteilung auf Δ_∞ konvergiert. In der Tat ist dies unter bestimmten Umständen der Fall; die Grenzverteilung ist dann gerade die Poisson-Dirichlet-Verteilung. Dies gilt es im Folgenden zu präzisieren.

Zunächst führen wir dazu Poisson-Prozesse auf allgemeinen Räumen sowie den Moran'schen Gammaprozess ein, mithilfe derer wir eine erste Definition der Poisson-Dirichlet-Verteilung bringen können. Sodann gehen wir näher auf das benannte Konvergenzresultat ein. Weiter zeigen wir noch die Verwandtschaft der Poisson-Dirichlet-Verteilung mit einer weiteren Wahrscheinlichkeitsverteilung auf Δ_∞ , unter deren Verwendung wir schließlich das gewünschte Invarianzresultat zeigen können.

3.2 Poisson-Prozesse auf allgemeinen Räumen

Poisson-Prozesse sind vor allem als Prozesse auf \mathbb{R} bekannt. Es ist jedoch ohne Schwierigkeiten möglich, einen Poisson-Prozess auf einem allgemeinen Raum S zu definieren. Von den Eigenschaften eines euklidischen Raumes wird hierfür lediglich die Messbarkeit von S benötigt. Ein Poisson-Prozess ist dann eine zufällige Punktmenge in S dergestalt, dass für Teilmengen A von S , die wir gewissermaßen als „Test“-Mengen verstehen können,

$$N(A) := \#\{\Pi \cap A\} \quad (3.4)$$

wohldefinierte, Poisson-verteilte Zufallsvariablen sind. Die genaue Definition lautet:

3.1 Definition. Sei (S, \mathcal{S}) ein messbarer Raum derart, dass $D := \{(x, y) : x = y\}$ bezüglich der Produkt- σ -Algebra $\mathcal{S} \otimes \mathcal{S}$ auf $S \times S$ messbar ist¹, und sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum. Weiter bezeichne S^∞ die Menge aller abzählbaren Teilmengen von S . Ein *Poisson-Prozess* Π ist eine Abbildung

$$\Pi: \Omega \longrightarrow S^\infty, \quad (3.5)$$

¹Dies stellt sicher, dass \mathcal{S} ausreichend viele Punkte enthält. Aus der Messbarkeit der Diagonalen D folgt nämlich bereits $\{x\} \in \mathcal{S}$ für alle $x \in S$.

die folgenden Bedingungen genügt:

- (i) Für beliebige, disjunkte Teilmengen A_1, \dots, A_n von S gilt:
 $N(A_1), \dots, N(A_n)$ sind unabhängige Zufallsvariable.
- (ii) Für jedes $A \in \mathcal{S}$ ist $N(A)$ Poisson-verteilt mit Parameter $\mu = \mu(A)$ und
 $0 \leq \mu \leq \infty$.

Dies bedeutet: Ist $\mu(A)$ endlich, so ist $\Pi \cap A$ mit Sicherheit eine endliche Menge und mit Sicherheit leer, falls $\mu(A) = 0$. Gilt $\mu(A) = \infty$, so ist $\Pi \cap A$ mit Sicherheit abzählbar unendlich. Aus den Eigenschaften von μ ergibt sich sofort das folgende Korollar:

3.2 Korollar. μ ist ein Maß auf (S, \mathcal{S}) .

Beweis. Offensichtlich ist $\mu(\emptyset) = 0$. Nach Eigenschaften der Poisson-Verteilung gilt für $A \in \mathcal{S}$

$$\mu(A) = EN(A).$$

Sind $A_1, A_2, \dots \in \mathcal{S}$ paarweise disjunkt, so gilt

$$N\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} N(A_i).$$

Der Übergang zum Erwartungswert liefert

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i),$$

womit die σ -Additivität gezeigt ist. □

Das Maß μ bezeichnen wir auch als das *Intensitätsmaß* des Poisson-Prozesses Π . Beachte, dass umgekehrt nicht jedes Maß auch ein Intensitätsmaß eines Poisson-Prozesses ist, denn nehmen wir an, es gebe ein $x \in S$ mit $m = \mu(\{x\}) > 0$, so folgt

$$P(N(\{x\}) \geq 2) = 1 - e^{-m} - me^{-m} > 0,$$

was mit

$$\#\{\Pi \cap \{x\}\} \leq 1 \quad \text{fast sicher}$$

unverträglich ist. Für das Intensitätsmaß eines Poisson-Prozesses muss also gelten

$$\mu(\{x\}) = 0 \quad \text{für alle } x \in S.$$

Für $S = \mathbb{R}^d$ bedeutet dies vor allem, dass μ \mathbb{X}_d -stetig ist und es eine messbare Funktion $\lambda: S \rightarrow \mathbb{R}^+$ gibt mit

$$\mu(A) = \int_A \lambda d\mathbb{X}_d \quad \text{für alle } A \in \mathcal{S}. \quad (3.6)$$

λ heißt auch die *Intensität* des Poisson-Prozesses. Ist λ konstant, sprechen wir von einem *homogenen* Poisson-Prozess. Beachte, dass das Integral in (3.6) nicht notwendigerweise konvergieren muss, da nach Definition auch $\mu(A) = \infty$ zugelassen ist, was insbesondere weiter unten bei dem von uns betrachteten Prozess der Fall sein wird.

Für unsere weitere Argumentation stellen wir an dieser Stelle noch das Abbildungstheorem vor. Es macht eine Aussage darüber, wann ein Poisson-Prozess Π mit Zustandsraum S derart in einen anderen Raum T abgebildet werden kann, dass der Bildprozess wiederum ein Poisson-Prozess ist. Es leuchtet sofort ein, dass dafür entscheidend ist, dass nicht zwei Punkte von Π auf denselben Punkt in T abgebildet werden. Eine notwendige Bedingung hierfür ist, dass das Bildmaß des Intensitätsmaßes keine Atome hat. Es zeigt sich, dass diese Bedingung unter der zusätzlichen Annahme der σ -Endlichkeit von μ bereits hinreichend ist. Hierauf wollen wir jedoch nicht weiter eingehen und notieren ohne Beweis (vgl. [17], S. 18):

3.3 Theorem (Abbildungstheorem). *Sei Π ein Poisson-Prozess mit Zustandsraum S und σ -endlichem Intensitätsmaß μ , und sei $f: S \rightarrow T$ eine messbare Funktion, so dass $\mu(f^{-1}(\{x\})) = 0$ für alle $x \in T$ gilt. Dann ist $f(\Pi)$ ein Poisson-Prozess auf T mit Intensitätsmaß $\mu^* := \mu^f = \mu(f^{-1}(\cdot))$.*

3.3 Definition der Poisson-Dirichlet-Verteilung

Nun sind alle Grundlagen geschaffen, um einen stochastischen Prozess zu definieren, der auch als der Moran'sche Gammaprozess bekannt ist. Mithilfe dieses Prozesses

können wir dann die Poisson-Dirichlet-Verteilung definieren. Wir folgen hier im Wesentlichen der Darstellung in [13].

Sei also Π^* ein Poisson-Prozess auf $S^* := \{(x, z) : x, z \geq 0\} \subset \mathbb{R}^2$ mit der Intensität $\lambda^*(x, z) = \frac{e^{-x}}{x}$. Für die Anzahl der Punkte von Π^* in der Menge

$$A_{\theta, z_0} := \{(x, z) : 0 \leq x \leq \theta, z > z_0\} \quad (3.7)$$

mit $\theta, z_0 > 0$ gilt

$$\#A_{\theta, z_0} \sim Poi(\theta E_1(z_0)),$$

wobei $E_1(z_0)$ die exponentielle Integralfunktion

$$E_1(z_0) := \int_{z_0}^{\infty} \frac{e^{-z}}{z} dz = \int_1^{\infty} \frac{e^{-z_0 z}}{z} dz \quad (3.8)$$

bezeichnet. Beachte, dass das Integral $E_1(z_0)$ für $z_0 = 0$ divergiert; dies steht jedoch nicht im Widerspruch zur Definition des Poisson-Prozesses.

3.4 Satz. *Unter Voraussetzungen wie oben gilt fast sicher*

$$\Pi^* \cap A_{\theta, 0} = \{(X_j(\theta), Z_{(j)}(\theta)) : j = 1, 2, \dots\} \quad (3.9)$$

mit:

(i) $Z_{(1)}(\theta) > Z_{(2)}(\theta) > \dots > 0$ sind die absteigend angeordneten Punkte eines nichthomogenen Poisson-Prozesses Π_θ auf \mathbb{R}^+ mit Intensität

$$\lambda_\theta(z) = \frac{\theta e^{-z}}{z}. \quad (3.10)$$

(ii) $X_1(\theta), X_2(\theta), \dots$ sind davon unabhängige, auf $[0, \theta]$ gleichverteilte Zufallsvariable.

Beweis. Der Beweis ist eine Anwendung des Abbildungstheorems. Π_θ ist die Projektion von Π^* auf die z -Achse; dies liefert die Aussage über die $Z_{(j)}(\theta)$. Die Intensität von Π_θ ergibt sich aus derjenigen von Π^* durch Integration über x . Ebenso ergibt die Projektion von Π^* auf die x -Achse einen homogenen Poisson-Prozess; die $X_j(\theta)$

sind daher gleichverteilt auf $[0, \theta]$. Die Unabhängigkeit resultiert schließlich aus der Wahl von λ^* als Produktmaß. \square

Beachte, dass wegen

$$\int_0^\infty \lambda_\theta(z) dz = \infty$$

Π_θ fast sicher unendlich viele Punkte enthält, die sich bei 0 häufen. Daraus folgt weiter, dass die X_j dicht liegen in $(0, \theta)$. Wir notieren ohne Beweis (vgl. [13], S. 2):

3.5 Korollar. *Für die Dichte von $Z_{(j)}(\theta)$ mit $\theta > 0$ gilt*

$$f_{Z_{(j)}(\theta)}(z) = \frac{\theta e^{-z} (\theta E_1(z))^{j-1}}{z (j-1)!} e^{-\theta E_1(z)}, \quad z > 0. \quad (3.11)$$

Schließlich haben wir (ohne Beweis, vgl. [13], S. 2):

3.6 Satz. *Die Zufallsvariable*

$$S(\theta) := \sum_{j=1}^{\infty} Z_{(j)}(\theta) < \infty \quad (3.12)$$

mit $\theta > 0$ ist $\Gamma(\theta, 1)$ -verteilt.

Hierbei bezeichnet $\Gamma(\alpha, \beta)$ die Gammaverteilung mit Parametern α und β ($\alpha, \beta > 0$; vgl. [2], S. 143).

Der so definierte stochastische Prozess $(S(\theta))_{\theta>0}$ wird auch als der *Moran'sche Gammaprozess* bezeichnet. Da der zugrunde liegende Poisson-Prozess invariant ist unter Translationen parallel zur x -Achse, hängen die Zuwächse

$$S(\theta) - S(\sigma) \quad \text{mit } \theta > \sigma$$

nur von der Differenz $\theta - \sigma$ ab. Es gilt also mit Satz 3.6

$$S(\theta) - S(\sigma) \sim S(\theta - \sigma) \sim \Gamma(\theta - \sigma, 1). \quad (3.13)$$

Einen stochastischen Prozess mit positiven, unabhängigen Zuwächsen, die nur von der Differenz abhängen, nennt man auch einen *Subordinator*. In der Literatur wird

der Moran'sche Gammaprozess daher gelegentlich auch als *Moran'scher Subordinator* bezeichnet.

Die Verteilung der nichtaufsteigend angeordneten, normalisierten (fast sicher verschiedenen) Sprünge des Moran'schen Gammaprozesses definieren wir nun als die Poisson-Dirichlet-Verteilung:

3.7 Definition (Poisson-Dirichlet-Verteilung). Sei für alle $j \geq 1$ und $\theta > 0$

$$\xi_j := \frac{Z_{(j)}(\theta)}{S(\theta)}. \quad (3.14)$$

Dann heißt die Verteilung von $(\xi_j)_{j \geq 1}$ die *Poisson-Dirichlet-Verteilung* mit Parameter θ , kurz $\mathcal{PD}(\theta)$. Falls $\theta = 1$, so schreiben wir auch \mathcal{PD}_1 .

Die Poisson-Dirichlet-Verteilung tritt in einer Vielzahl von Anwendungen auf. Darüber hinaus kommt sie als stationäre Verteilung bei verschiedenen Prozessen vor, so auch bei dem hier betrachteten CCF-Prozess. Dies werden wir in Abschnitt 3.7 zeigen.

3.4 Der Poisson-Dirichlet-Grenzfall

Nun haben wir das nötige Handwerkszeug beisammen, um auf die Konvergenz der Dirichlet-Verteilungen auf Δ_n einzugehen. Unsere Ausführungen stützen sich auf [15], Kapitel 5.

Wir starten mit einem zufälligen Vektor $p \in \Delta_n$ mit Verteilung $\mathcal{D}(\alpha_1, \dots, \alpha_n)$, wobei wir zunächst der Einfachheit halber annehmen, alle α_j haben denselben Wert, etwa $\alpha_j = \alpha$ für $j = 1, \dots, n$. Aufgrund von Symmetrie, und da $E \sum_j p_j = 1$ ist, erhalten wir

$$E p_j = \frac{1}{n} \quad \text{für alle } j.$$

Dies bedeutet, dass bei großem n die einzelnen p_j mit hoher Wahrscheinlichkeit kleine Werte annehmen. Dennoch können die einzelnen Werte sehr verschieden ausfallen, und es ist sinnvoll, den Vektor nach nichtaufsteigender Komponentengröße anzuordnen;

$$p_{(1)} \geq p_{(2)} \geq \dots \geq p_{(n)} \quad (3.15)$$

bezeichne die nichtaufsteigend geordnete Sequenz. Die gesuchte Konvergenz besteht nun in dem Sinne, dass wir zeigen können, dass für jedes $k \in \mathbb{N}$ die Verteilung des Vektors $(p_{(1)}, \dots, p_{(k)})$ für $n \rightarrow \infty$ konvergiert, allerdings vorausgesetzt, dass α dabei derart gegen null geht, dass $n\alpha$ gegen einen strikt positiven Wert $\theta \in \mathbb{R}^+$ konvergiert.

Dieser Sachverhalt wird in etwas allgemeinerer Form durch das folgende Theorem präzisiert (vgl. [15], S. 6):

3.8 Theorem. Für jedes $n \in \mathbb{N}$ sei $(p_{(1)}^{(n)}, \dots, p_{(n)}^{(n)})$ die nichtaufsteigende Umordnung eines Zufallsvektors $p^{(n)}$ mit Verteilung $\mathcal{D}(\alpha_1^{(n)}, \dots, \alpha_n^{(n)})$, $\alpha_j^{(n)} \geq 0$. Für $n \rightarrow \infty$ gelte

$$\max(\alpha_1^{(n)}, \dots, \alpha_n^{(n)}) \rightarrow 0 \quad (3.16)$$

und

$$\theta^{(n)} := \alpha_1^{(n)} + \dots + \alpha_n^{(n)} \rightarrow \theta \quad (3.17)$$

für ein θ mit $0 < \theta < \infty$. Dann konvergiert für jedes $k \in \mathbb{N}$ die Verteilung des Vektors $(p_{(1)}^{(n)}, \dots, p_{(k)}^{(n)})$ für $n \rightarrow \infty$ gegen die Randverteilung der k ersten Komponenten von $\mathcal{PD}(\theta)$.

Für den Beweis benötigen wir etwas Vorarbeit. Zunächst nutzen wir aus, dass sich ein Dirichlet-verteilter Vektor (p_1, \dots, p_n) ebenfalls durch den Moran'schen Gammaprozess darstellen lässt, und zwar in der Weise, dass wir die p_j als normalisierte Zuwächse des Gammaprozesses auffassen. Dies ist möglich, da zwischen der Dirichlet-Verteilung und der Gammaverteilung folgende enge Beziehung besteht (vgl. [17], S. 91):

3.9 Lemma. Seien Y_1, \dots, Y_n unabhängige, $\Gamma(\alpha_j, 1)$ -verteilte Zufallsvariablen mit $\alpha_j \geq 0$ für $j = 1, \dots, n$ und sei $Y = \sum_{j=1}^n Y_j$. Dann ist der Vektor $p = (p_1, \dots, p_n)$ mit

$$p_j := \frac{Y_j}{Y}, \quad j = 1, \dots, n \quad (3.18)$$

$\mathcal{D}(\alpha_1, \dots, \alpha_n)$ -verteilt und unabhängig von Y . Ferner ist Y $\Gamma(\alpha, 1)$ -verteilt mit $\alpha = \sum_{j=1}^n \alpha_j$.

Beweis. Y_j hat die Dichte

$$f_{\alpha_j}(y_j) = \frac{1}{\Gamma(\alpha_j)} y_j^{\alpha_j-1} e^{-y_j}, \quad y_j > 0$$

für $j = 1, \dots, n$. Damit ergibt sich als Dichte der gemeinsamen Verteilung von (Y_1, \dots, Y_n)

$$f_{(\alpha_1, \dots, \alpha_n)}(y_1, \dots, y_n) = \frac{1}{\prod_{j=1}^n \Gamma(\alpha_j)} \cdot \prod_{j=1}^n y_j^{\alpha_j-1} \cdot e^{-\sum_{j=1}^n y_j}, \quad y_1, \dots, y_n > 0.$$

Ersetzen wir nun die Variablen (Y_1, \dots, Y_n) durch (Y, p_1, \dots, p_{n-1}) , transformiert sich die Dichte zu

$$f_{(\alpha_1, \dots, \alpha_n)}(y, p_1, \dots, p_{n-1}) = \frac{1}{\prod_{j=1}^n \Gamma(\alpha_j)} \cdot \prod_{j=1}^{n-1} p_j^{\alpha_j-1} \cdot \left(1 - \sum_{j=1}^{n-1} p_j\right)^{\alpha_n-1} \cdot y^{\alpha-n} e^{-y} \cdot \left| \frac{\partial(y_1, \dots, y_n)}{\partial(y, p_1, \dots, p_{n-1})} \right|.$$

Die Funktionaldeterminante $\left| \frac{\partial(y_1, \dots, y_n)}{\partial(y, p_1, \dots, p_{n-1})} \right|$ ergibt sich gerade zu y^{n-1} , so dass wir insgesamt erhalten:

$$f_{(\alpha_1, \dots, \alpha_n)}(y, p_1, \dots, p_{n-1}) = \frac{\Gamma(\alpha_1 + \dots + \alpha_n)}{\prod_{j=1}^n \Gamma(\alpha_j)} \prod_{j=1}^{n-1} p_j^{\alpha_j-1} \cdot \left(1 - \sum_{j=1}^{n-1} p_j\right)^{\alpha_n-1} \cdot \frac{y^{\alpha-1} e^{-y}}{\Gamma(\alpha)}.$$

Daraus ergeben sich die Aussagen über die Verteilung sowie die Unabhängigkeit. \square

Für den Moran'schen Gammaprozess gilt nach Satz 3.6

$$S(\theta) \sim \Gamma(\theta, 1)$$

und vermöge Gleichung (3.13)

$$S\left(\sum_{i=1}^j \alpha_i^{(n)}\right) - S\left(\sum_{i=1}^{j-1} \alpha_i^{(n)}\right) \sim S(a_j^{(n)}) \sim \Gamma(\alpha_j^{(n)}, 1) \quad \text{für } j = 1, \dots, n. \quad (3.19)$$

Damit ergibt sich insgesamt in Kombination mit dem vorhergehenden Lemma:

3.10 Lemma. Der Vektor $\bar{p}^{(n)} = (\bar{p}_1^{(n)}, \dots, \bar{p}_n^{(n)})$ mit

$$\bar{p}_j^{(n)} := \frac{S\left(\sum_{i=1}^j \alpha_i^{(n)}\right) - S\left(\sum_{i=1}^{j-1} \alpha_i^{(n)}\right)}{S(\theta^{(n)})} \quad (3.20)$$

ist $\mathcal{D}(\alpha_1^{(n)}, \dots, \alpha_n^{(n)})$ -verteilt.

Wir können also für den Beweis des Theorems $p^{(n)}$ durch den identisch verteilten Vektor $\bar{p}^{(n)}$ ersetzen. Das folgende Konvergenzresultat liefert den Schlüssel:

3.11 Lemma. *Sei $\bar{p}^{(n)}$ wie im vorhergehenden Lemma definiert. Dann gilt für die nichtaufsteigend angeordneten Sprunghöhen $Z_{(1)} \geq Z_{(2)} \geq \dots \geq 0$ von S auf $(0, \theta)$ und alle $j \geq 1$*

$$\lim_{n \rightarrow \infty} \bar{p}_{(j)}^{(n)} = \frac{Z_{(j)}}{S(\theta)}. \quad (3.21)$$

Beweis. Wir betrachten eine beliebige Realisierung des Prozesses S . Für beliebiges $k \in \mathbb{N}$ wähle $m \in \mathbb{N}$ so groß, dass für alle $n \geq m$ die zu den Sprunghöhen $Z_{(1)}, \dots, Z_{(k)}$ gehörenden Unstetigkeitsstellen von S alle in verschiedenen Intervallen $(\alpha_{j-1}^{(n)}, \alpha_j^{(n)}]$ liegen. Dann gilt (für passend gewähltes j')

$$\bar{p}_{(j)}^{(n)} S(\theta^{(n)}) = S\left(\sum_{i=1}^{j'} \alpha_i^{(n)}\right) - S\left(\sum_{i=1}^{j'-1} \alpha_i^{(n)}\right) \geq Z_{(j)}$$

für $1 \leq j \leq k$ und $n \geq m$, und es folgt

$$\liminf_{n \rightarrow \infty} \bar{p}_{(j)}^{(n)} S(\theta^{(n)}) \geq Z_{(j)} \quad \text{für } j = 1, \dots, k.$$

Da $\{\theta\}$ eine \mathbb{A} -Nullmenge ist und somit θ fast sicher keine Unstetigkeitsstelle ist, folgt

$$\liminf_{n \rightarrow \infty} S(\theta^{(n)}) = \lim_{n \rightarrow \infty} S(\theta^{(n)}) = S(\theta)$$

und damit insgesamt

$$\liminf_{n \rightarrow \infty} \bar{p}_{(j)}^{(n)} \geq \frac{Z_{(j)}}{S(\theta)} \quad \text{für } j = 1, \dots, k. \quad (3.22)$$

Da k beliebig war, gilt dies sogar für alle $j \in \mathbb{N}$, und mit dem Lemma von Fatou ergibt sich zusammen mit $S(\theta) = \sum_{j=1}^{\infty} Z_{(j)}$

$$\begin{aligned} \limsup_{n \rightarrow \infty} \bar{p}_{(j)}^{(n)} &= \limsup_{n \rightarrow \infty} \left(1 - \sum_{i \neq j} \bar{p}_{(i)}^{(n)}\right) \\ &\leq 1 - \sum_{i \neq j} \liminf_{n \rightarrow \infty} \bar{p}_{(i)}^{(n)} \end{aligned}$$

$$\begin{aligned} &\leq 1 - \sum_{i \neq j} \frac{Z_{(i)}}{S(\theta)} \\ &= \frac{Z_{(j)}}{S(\theta)} \end{aligned} \quad (3.23)$$

für alle $j \in \mathbb{N}$. Damit ist die gewünschte Konvergenz gezeigt. \square

Beweis des Theorems. Wir setzen für alle $j \geq 1$

$$\xi_j := \frac{Z_{(j)}(\theta)}{S(\theta)}. \quad (3.24)$$

Der Vektor $(\xi_j)_{j \geq 1}$ ist damit $\mathcal{PD}(\theta)$ -verteilt (vgl. Definition 3.7), und aus dem obigen Lemma folgt, dass für jedes $k \in \mathbb{N}$ die gemeinsame Verteilung von $(\bar{p}_{(1)}^{(n)}, \dots, \bar{p}_{(k)}^{(n)})$ und damit auch die von $(p_{(1)}^{(n)}, \dots, p_{(k)}^{(n)})$ für $n \rightarrow \infty$ gegen diejenige von (ξ_1, \dots, ξ_k) konvergiert. Damit ist alles gezeigt. \square

Das gezeigte Konvergenzresultat lässt sich wie folgt noch einmal zusammenfassen: Die nichtaufsteigende Ordnungsstatistik der Dirichlet-Verteilung $\mathcal{D}(\alpha_1, \dots, \alpha_n)$ approximiert $\mathcal{PD}(\theta)$, solange n groß ist, die α_j gleichmäßig klein und $\sum_{j=1}^{\infty} \alpha_j$ in der Nähe von θ liegt.

Als Nächstes wollen wir noch eine weitere Charakterisierung der Poisson-Dirichlet-Verteilung geben. Diese erfolgt über die GEM-Verteilung und ist möglicherweise etwas anschaulicher.

3.5 Die GEM-Verteilung

Bei der GEM-Verteilung – benannt nach Griffiths, Engen and McCloskey – handelt es sich um eine Wahrscheinlichkeitsverteilung auf Δ_{∞} (vgl. [18], S. 505).

3.12 Definition (GEM-Verteilung). Eine Sequenz von Zufallsvariablen $(X_n)_{n \geq 1}$ heißt *GEM(θ)-verteilt* mit Parameter $\theta > 0$, wenn gilt

$$X_n := U_n \prod_{i=1}^{n-1} (1 - U_i) \quad \text{für } n \geq 1 \quad (3.25)$$

mit unabhängigen, identisch betaverteilten Zufallsvariablen $(U_i)_{i \geq 1}$ mit $\beta(1, \theta)$ -Dichte

$$\theta(1-x)^{\theta-1}, \quad 0 < x < 1. \quad (3.26)$$

$GEM(1)$ kann aufgefasst werden als die Verteilung der Bruchstücklängen, die durch gleichmäßiges sukzessives Abbrechen von einem Stab mit der Ausgangslänge 1 entstehen: X_n entspricht der Länge des Bruchstücks im n -ten Schritt, und $Y_{n+1} := \prod_{i=1}^n (1 - U_i)$ stellt die Länge des hernach übriggebliebenen Reststücks dar (mit $Y_1 := 1$, vgl. auch [7], S. 918).

3.6 Zusammenhang zwischen der GEM - und der Poisson-Dirichlet-Verteilung

Wir wollen nun auf die enge Beziehung zwischen der GEM - und der Poisson-Dirichlet-Verteilung eingehen. Dazu definieren wir die folgenden zwei Operationen (vgl. [18], Definition 3):

3.13 Definition (Längenverzernte Permutation). Für $(p_n)_{n \geq 1} \in \Delta_\infty$ sei $(A_n)_{n \geq 1}$ eine Zerlegung von $[0, 1]$ in disjunkte Intervalle derart, dass A_n die Länge p_n hat für alle n , und sei $(U_i)_{i \geq 1}$ eine Folge unabhängiger, auf $[0, 1]$ gleichverteilter Zufallsvariablen. Weiter sei $(j_n)_{n \geq 1}$ eine Folge natürlicher Zahlen, die so gewählt ist, dass gilt

$$\begin{aligned} U_{\sigma_1} &\in A_{j_1} & \text{mit } \sigma_1 &:= 1 \\ U_{\sigma_2} &\in A_{j_2} & \text{mit } \sigma_2 &:= \min\{i \geq 2 : U_i \notin A_{j_1}\} \\ U_{\sigma_3} &\in A_{j_3} & \text{mit } \sigma_3 &:= \min\{i \geq 3 : U_i \notin A_{j_1} \cup A_{j_2}\} \\ &\vdots & & \vdots \end{aligned} \quad (3.27)$$

Dann heißt die Sequenz $(q_n)_{n \geq 1} := (p_{j_n})_{n \geq 1} \in \Delta_\infty$ *längenverzernte Permutation* (*size-biased permutation*, *SBP*) von $(p_n)_{n \geq 1}$.

Durch diese Operation wird ein Markov-Übergangskern von Δ_∞ nach Δ_∞ definiert, den wir mit $SBP(\cdot, \cdot)$ bezeichnen. Ebenso können wir Vektoren aus Δ_n abbilden, indem wir diese als Elemente von Δ_∞ auffassen.

Die Verteilung von $\text{SBP}((p_n)_{n \geq 1}, \cdot)$ auf Δ_∞ für ein gegebenes $(p_n)_{n \geq 1}$ hängt allein von der *nichtaufsteigenden Anordnung* der $(p_n)_{n \geq 1}$ ab. Die nichtaufsteigende Anordnung ist eine Abbildung $\Delta_\infty \longrightarrow \Omega_1$, und wir werden sie mit sort bezeichnen. Dies ist die zweite benötigte Operation.

Nun lässt sich die Beziehung zwischen der Poisson-Dirichlet-Verteilung \mathcal{PD} und der GEM-Verteilung wie folgt darstellen (vgl. [18], (3.1)):

3.14 Satz. Für $\theta > 0$ gilt

$$\mathcal{PD}(\theta) \xrightleftharpoons[\text{sort}]{\text{SBP}} \text{GEM}(\theta). \quad (3.28)$$

Dies geht aus dem folgenden Satz hervor (Proposition 4.2 in [13]):

3.15 Satz. Sei $(p_k)_{k \geq 1}$ $\mathcal{PD}(\theta)$ -verteilt mit $\theta > 0$. Dann gilt für die längenverzerzte Permutation von $(p_k)_{k \geq 1}$

$$\text{SBP}((p_k)_{k \geq 1}) = (U_1, (1 - U_1)U_2, (1 - U_1)(1 - U_2)U_3, \dots) \quad (3.29)$$

mit unabhängigen, identisch $\beta(1, \theta)$ -verteilten Zufallsvariablen $(U_k)_{k \geq 1}$.

Für den Beweis benötigen wir folgendes Lemma:

3.16 Lemma. Sei $Z = (Z_1, \dots, Z_n)$ $\mathcal{D}(\alpha, \dots, \alpha)$ -verteilt mit $\alpha > 0$. Dann lässt sich die längenverzerzte Permutation von Z darstellen als

$$\text{SBP}(Z) = (U_1, (1 - U_1)U_2, \dots, (1 - U_1) \cdots (1 - U_{n-1})U_n) \quad (3.30)$$

mit unabhängigen Zufallsvariablen U_j , $j = 1, \dots, n$, wobei U_j die $\beta(\alpha + 1, (n - j)\alpha)$ -Dichte

$$f_{U_j}(x) = \frac{\Gamma((n - j + 1)\alpha + 1)}{\Gamma(\alpha + 1)\Gamma((n - j)\alpha)} x^\alpha (1 - x)^{(n-j)\alpha - 1}, \quad 0 < x < 1 \quad (3.31)$$

hat.

Beweis. Die längenverzerzte Permutation von Z stellen wir zunächst als $(Z_{\nu_i})_{i \geq 1}$ dar; die $(\nu_i)_{i \geq 1}$ seien dementsprechend Zufallsvariablen mit Werten in \mathbb{N} , die so gewählt

sind, dass gilt

$$\begin{aligned} P(\nu_1 = j|Z) &= Z_j && \text{für } j = 1, 2, \dots, \\ P(\nu_2 = k|Z, \nu_1) &= Z_k/(1 - Z_j) && \text{für } k = 1, 2, \dots, k \neq \nu_1, \\ &\vdots && \vdots \end{aligned}$$

Um die Verteilung des Vektors nach dem ersten Umordnungsschritt zu bestimmen, nutzen wir ein Resultat aus der Bayes-Theorie, das besagt (vgl. [11], S. 76): Ist $X = (X_1, \dots, X_n)$ $\mathcal{D}(\alpha_1, \dots, \alpha_n)$ -verteilt und $\beta = (\beta_1, \dots, \beta_n)$ ein beliebiger Zufallsvektor mit

$$\beta|X = (\beta_1, \dots, \beta_n)|X \sim M(k, X),$$

wobei $M(k, X)$ die Multinomialverteilung mit Parametern k und X_1, \dots, X_n bezeichnet, so gilt

$$X|\beta \sim \mathcal{D}(\alpha + \beta).$$

In unserem konkreten Fall ($k = 1$) gilt für $\beta_{\nu_1} := (\delta_{1,\nu_1}, \dots, \delta_{n,\nu_1})$

$$\beta_{\nu_1}|Z \sim M(1, Z),$$

und es folgt, dass der Vektor

$$Z' = (Z_{\nu_1}, Z_1, \dots, Z_{\nu_1-1}, Z_{\nu_1+1}, \dots, Z_n)$$

die Verteilung $\mathcal{D}(\alpha + 1, \alpha, \dots, \alpha)$ hat.

Unter Verwendung des Lemmas 3.9 und der Faltungseigenschaft der Gammaverteilung (vgl. [2], S. 144) erhalten wir: Der Vektor $(Z_{\nu_1}, 1 - Z_{\nu_1})$ hat die Verteilung

$$\mathcal{D}(\alpha + 1, (n - 1)\alpha),$$

und somit hat Z_{ν_1} die Wahrscheinlichkeitsdichte

$$f_{Z_{\nu_1}}(x) = \frac{\Gamma(n\alpha + 1)}{\Gamma(\alpha + 1)\Gamma((n - 1)\alpha)} x^\alpha (1 - x)^{(n-1)\alpha-1}, \quad 0 < x < 1. \quad (3.32)$$

Wir greifen abermals auf das Lemma 3.9 zurück und betrachten unabhängige Zufallsvariablen Y_1, \dots, Y_n mit Verteilungen $\Gamma(\alpha_1, 1), \dots, \Gamma(\alpha_n, 1)$ für $\alpha_j \geq 0$. Dann hat $Y := \sum_{i=1}^n Y_i$ die Verteilung $\Gamma(\alpha_1 + \dots + \alpha_n, 1)$ und ist unabhängig vom Vektor $(Y_1/Y, \dots, Y_n/Y)$, welcher $\mathcal{D}(\alpha_1, \dots, \alpha_n)$ -verteilt ist. Ebenso ist mit $\tilde{Y} := Y_2 + \dots + Y_n$

$$(Y_2/\tilde{Y}, \dots, Y_n/\tilde{Y})$$

$\mathcal{D}(\alpha_2, \dots, \alpha_n)$ -verteilt und unabhängig von \tilde{Y} und damit auch unabhängig von

$$\frac{Y_1}{Y} = \frac{Y_1}{Y_1 + \tilde{Y}}.$$

Die Komponenten des obenstehenden Vektors können wir auch als

$$\frac{Y_j}{\tilde{Y}} = \frac{Y_j}{Y} \cdot \frac{Y}{\tilde{Y}} = \frac{Y_j}{Y} \cdot \frac{1}{1 - Y_1/Y}$$

darstellen ($j = 2, \dots, n$).

Übertragen wir diese Resultate nun vermöge $Z_j \equiv Y_j/Y$ auf unseren Fall, folgt: Der Vektor

$$(Z_1, \dots, Z_{\nu_1-1}, Z_{\nu_1+1}, \dots, Z_n)/(1 - Z_{\nu_1})$$

ist $\mathcal{D}(\alpha, \dots, \alpha)$ -verteilt und unabhängig von Z_{ν_1} . Wir können daher für $Z_{\nu_2}, \dots, Z_{\nu_n}$ analog argumentieren wie für Z_{ν_1} . Damit ergibt sich für $\text{SBP}(Z) = (Z_{\nu_1}, \dots, Z_{\nu_n})$ die Darstellung

$$Z_{\nu_1} = U_1, \quad Z_{\nu_2} = (1 - U_1) \cdot U_2, \quad \dots, \quad Z_{\nu_n} = \prod_{j=1}^{n-1} (1 - U_j) \cdot U_n,$$

wobei U_j , $j = 1, \dots, n$ unabhängige Zufallsvariablen sind; U_r hat die Dichte (3.32) mit n ersetzt durch $n - r + 1$. □

Beweis des Satzes. Es sei $Z = (Z_1, \dots, Z_n) \sim \mathcal{D}(\alpha, \dots, \alpha)$ mit $\alpha := \theta/n$ und $\text{SBP}(Z)$ sei dargestellt wie im vorangehenden Lemma. Lassen wir nun n gegen unendlich

gehen, so folgt $\alpha \rightarrow 0$ und die Dichten der Zufallsvariablen U_j

$$f_{U_j}(x) = \frac{\Gamma(\theta - (j+1)\alpha + 1)}{\Gamma(\alpha + 1)\Gamma(\theta - j\alpha)} x^\alpha (1-x)^{\theta-j\alpha-1}, \quad 0 < x < 1$$

konvergieren für alle $j \in \mathbb{N}$ gegen

$$f_{\beta(1,\theta)}(x) = \frac{\Gamma(\theta + 1)}{\Gamma(1)\Gamma(\theta)} x^0 (1-x)^{\theta-1} = \theta(1-x)^{\theta-1}, \quad 0 < x < 1,$$

die $\beta(1, \theta)$ -Dichte. Somit konvergiert $(\text{SBP}(Z_1, \dots, Z_n), 0, 0, \dots)$ und damit auch $(\text{SBP}(Z_{(1)}, \dots, Z_{(n)}), 0, 0, \dots)$ für $n \rightarrow \infty$ gegen

$$(U_1, (1 - U_1)U_2, (1 - U_1)(1 - U_2)U_3, \dots),$$

wobei $(U_j)_{j \geq 1}$ unabhängige, identisch $\beta(1, \theta)$ -verteilte Zufallsvariable sind.

Andererseits konvergiert, wie oben gezeigt (vgl. Theorem 3.8),

$$(Z_{(1)}, \dots, Z_{(n)}, 0, 0, \dots)$$

für $n \rightarrow \infty$ gegen die Poisson-Dirichlet-verteilte Sequenz (p_1, p_2, \dots) .

Für einen rigorosen Beweis dieser Konvergenzresultate verweisen wir auf [8], Theorem 3. □

Damit erhalten wir sofort auch den umgekehrten Zusammenhang:

3.17 Korollar. *Sei $(X_n)_{n \geq 1}$ $GEM(\theta)$ -verteilt mit $\theta > 0$. Dann ist $\text{sort}((X_n)_{n \geq 1})$ $\mathcal{PD}(\theta)$ -verteilt.*

Beweis. Wir schreiben $(X_n)_{n \geq 1} = \text{SBP}((Y_n)_{n \geq 1})$ für eine $\mathcal{PD}(\theta)$ -verteilte Sequenz $(Y_n)_{n \geq 1}$. Dann ist $\text{sort}((X_n)_{n \geq 1}) = \text{sort}(\text{SBP}((Y_n)_{n \geq 1})) = (Y_n)_{n \geq 1}$ $\mathcal{PD}(\theta)$ -verteilt. □

Für späteren Gebrauch notieren wir noch: Da jedes $(x_i)_{i \geq 1} \in \Delta_\infty$ aufgrund der \mathbb{X} -Stetigkeit der $\beta(1, \theta)$ -Verteilung offensichtlich eine $GEM(\theta)$ -Nullmenge darstellt, folgt insbesondere, dass jedes $p \in \Omega_1$ eine $\mathcal{PD}(\theta)$ -Nullmenge darstellt.

3.7 Invarianz unter CCF

Wir werden nun zeigen, dass die Poisson-Dirichlet-Verteilung eine invariante Verteilung des CCF-Prozesses auf Ω_1 ist. Dazu folgen wir im Wesentlichen den Ausführungen in [12].

Zunächst führen wir einen weiteren Split-Merge-Prozess auf Δ_∞ ein, den wir *W-Split-Merge* nennen wollen, weil er von einer Zufallsvariablen W gesteuert wird. Für diesen Prozess lässt sich zeigen, dass er die *GEM*-Verteilung invariant lässt. Sodann nutzen wir aus, dass die CCF-Transition mittels längenverzerrtem Permutieren und nichtaufsteigendem Anordnen aus der *W*-Split-Merge-Transition gewonnen werden kann. Da die Poisson-Dirichlet-Verteilung und die *GEM*-Verteilung ebenfalls über SBP und sort zusammenhängen, erhalten wir das gewünschte Ergebnis.

Wir definieren nun den *W*-Split-Merge-Prozess:

3.18 Definition (*W*-Split-Merge). Sei $(X_n)_{n \geq 1}$ eine Sequenz von nichtnegativen Zufallsvariablen mit $\sum_n X_n = 1$. Für eine Zufallsvariable W mit Werten in $[0, 1]$ und unabhängig von $(X_n)_{n \geq 1}$ definieren wir die *W-Split-Merge-transformierte Sequenz* $(Y_n)_{n \geq 1}$ durch

$$(Y_1, Y_2, Y_3, \dots) := \begin{cases} (W, X_1 - W, X_2, X_3, \dots) & \text{für } W < X_1, \\ (X_1 + X_2, X_3, X_4, \dots) & \text{für } W \geq X_1. \end{cases} \quad (3.33)$$

Es ist zweckmäßig, sich $(X_n)_{n \geq 1}$ als Partition von $[0, 1]$ zu denken. W markiert dann einen zufälligen Punkt in $[0, 1]$. Liegt der Punkt im ersten Abschnitt, so wird dieser unterteilt in zwei neue Abschnitte W und $X_1 - W$. Liegt W in einem anderen Abschnitt, so werden der erste und der zweite Abschnitt miteinander verbunden.

Offenbar liefert dies eine Markov-Kette auf Δ_∞ . Wir werden später zeigen, wie wir aus diesem Prozess für $W \sim U[0, 1]$ den CCF-Prozess erhalten.

Wir zeigen nun die Invarianz der *GEM*-Verteilung unter *W*-Split-Merge (vgl. [12], Theorem 1):

3.19 Satz. Seien W, U_1, U_2, \dots unabhängige, identisch $\beta(1, \theta)$ -verteilte Zufallsvara-

ben auf $[0, 1]$ mit $\theta > 0$. Dann ist

$$(X_n)_{n \geq 1} := \left(U_n \prod_{i=1}^{n-1} (1 - U_i) \right)_{n \geq 1} \quad (3.34)$$

invariant unter der W -Split-Merge-Transformation.

Für den Beweis formulieren wir zunächst im folgenden Lemma die Invarianz als Integralgleichung. Dazu führen wir die Verteilungsfunktionen F und G der komplementären Zufallsvariablen $1 - W$ und $1 - U_n$, $n \geq 1$

$$F(t) := P(1 - U_n < t) \quad \text{und} \quad G(t) := P(1 - W < t) \quad (3.35)$$

ein. Damit gilt:

3.20 Lemma. $(X_n)_{n \geq 1} = (U_n \prod_{i=1}^{n-1} (1 - U_i))_{n \geq 1}$ ist genau dann invariant unter der W -Split-Merge-Transformation, wenn F und G der Gleichung

$$\begin{aligned} \int_a^1 F(u) - F(bu) dG(u) + (1 - F(b)) \int_a^1 (1 - F(au^{-1})) G(u) dF(u) \\ = (1 - F(a))(1 - F(b)) \end{aligned} \quad (3.36)$$

für $a, b \in [0, 1]$ genügen.

Beweis. Wir berechnen neue Zufallsvariablen $(V_n)_{n \geq 1}$ aus der W -Split-Merge-transformierten Sequenz $(Y_n)_{n \geq 1}$ vermöge

$$V_1 := Y_1; \quad V_n := \frac{Y_n}{1 - \sum_{i=1}^{n-1} Y_i} \quad \text{für } n \geq 2 \quad (3.37)$$

und drücken diese Sequenz anschließend durch die Zufallsvariablen $(U_n)_{n \geq 1}$ aus. Für $W < U_1$ erhalten wir

$$\begin{aligned} V_1 &= Y_1 = W \\ V_2 &= \frac{Y_2}{1 - Y_1} = \frac{X_1 - W}{1 - W} = \frac{U_1 - W}{1 - W} \\ V_3 &= \frac{Y_3}{1 - Y_1 - Y_2} = \frac{X_2}{1 - W - U_1 + W} = U_2 \end{aligned}$$

$$\begin{aligned} V_4 &= U_3 \\ &\vdots \end{aligned}$$

Im Falle $W \geq U_1$ gilt

$$\begin{aligned} V_1 &= Y_1 = X_1 + X_2 = U_1 + (1 - U_1)U_2 \\ V_2 &= \frac{Y_2}{1 - Y_1} = \frac{X_3}{(1 - U_1) - (1 - U_1)U_2} = U_3 \\ V_3 &= U_4 \\ &\vdots \end{aligned}$$

insgesamt also

$$(V_1, V_2, V_3, V_4, \dots) = \begin{cases} (W, \frac{U_1 - W}{1 - W}, U_2, U_3, \dots) & \text{für } W < U_1, \\ (U_1 + (1 - U_1)U_2, U_3, U_4, U_5, \dots) & \text{für } W \geq U_1. \end{cases} \quad (3.38)$$

Invarianz liegt genau dann vor, wenn $(V_n)_{n \geq 1}$ dieselbe Verteilung hat wie $(U_n)_{n \geq 1}$. Offensichtlich sind die Zufallsvariablen V_3, V_4, \dots unabhängig voneinander und haben dieselbe Verteilung wie U_3, U_4, \dots . Weiter sind sie nach Konstruktion auch unabhängig von V_1 und V_2 . Es ist daher lediglich zu zeigen, dass die gemeinsame Verteilung von V_1 und V_2 mit der von U_1 und U_2 übereinstimmt.

Um die gemeinsame Verteilung von V_1, V_2 zu bestimmen, schreiben wir diese als Summe von zwei Termen:

$$\begin{aligned} P(V_1 < 1 - a, V_2 < 1 - b) &= P(V_1 < 1 - a, V_2 < 1 - b, W < U_1) \\ &\quad + P(V_1 < 1 - a, V_2 < 1 - b, W \geq U_1). \end{aligned} \quad (3.39)$$

Der erste Term lässt sich unter Verwendung von $V_1 = W$ und $V_2 = \frac{U_1 - W}{1 - W} < 1 - b \Leftrightarrow 1 - U_1 > (1 - W)b$ umschreiben zu

$$\begin{aligned} &P(V_1 < 1 - a, V_2 < 1 - b, W < U_1) \\ &= P(1 - W > a, 1 - W > 1 - U_1 > (1 - W)b), \end{aligned}$$

was sich wiederum als Integral darstellen lässt:

$$\begin{aligned} &= \int P(1 - W > 1 - U_1 > (1 - W)b \mid 1 - W = u) dP^{1-W}(u) \\ &= \int_a^1 F(u) - F(bu) dG(u). \end{aligned} \quad (3.40)$$

Der zweite Term lässt sich analog mit $V_1 = U_1 + (1 - U_1)U_2 < 1 - a \Leftrightarrow (1 - U_1)(1 - U_2) > a$ und $V_2 = U_3$ umformulieren zu

$$\begin{aligned} &P(V_1 < 1 - a, V_2 < 1 - b, W \geq U_1) \\ &= P((1 - U_1)(1 - U_2) > a, 1 - U_3 > b, 1 - W \leq 1 - U_1) \\ &= (1 - F(b))P(1 - U_2 > a/(1 - U_1), 1 - W \leq 1 - U_1). \end{aligned}$$

Dies kann ebenfalls als Integral dargestellt werden:

$$\begin{aligned} &= (1 - F(b)) \int_a^1 P(1 - U_2 > a/(1 - U_1), 1 - W \leq 1 - U_1 \mid 1 - U_1 = u) dP^{1-U_1}(u) \\ &= (1 - F(b)) \int_a^1 (1 - F(au^{-1}))G(u) dF(u). \end{aligned} \quad (3.41)$$

$P(V_1 < 1 - a, V_2 < 1 - b)$ ist also gleich der linken Seite der Integralgleichung. Andererseits gilt

$$P(U_1 < 1 - a, U_2 < 1 - b) = (1 - F(a))(1 - F(b)),$$

und die Aussage des Lemmas folgt. \square

Beweis des Satzes. Da nach Voraussetzung die Zufallsvariablen W, U_1, U_2, \dots alle $\beta(1, \theta)$ -verteilt sind, ergibt sich für die Verteilungsfunktionen der inversen Zufallsvariablen $1 - W$ und $1 - U_n, n \geq 1$

$$\begin{aligned} F(u) = G(u) = P(1 - W < u) &= \int_{(1-u)}^1 F_{\beta(1, \theta)}(x) d\mathbb{X}(x) \\ &= \int_{(1-u)}^1 \theta(1 - x)^{\theta-1} d\mathbb{X}(x) \end{aligned}$$

$$= -[(1-x)^\theta]_{1-u}^1 = u^\theta, \quad u \in [0, 1].$$

Damit folgt

$$F(u) - F(bu) = u^\theta(1 - b^\theta) = F(u)(1 - F(b)),$$

und die Integralgleichung (3.36) nimmt nach Division durch $1 - F(b)$ die Form

$$\int_a^1 F(u) dG(u) + \int_a^1 G(u) dF(u) = 1 - F(a) + \int_a^1 F(au^{-1})G(u) dF(u) \quad (3.42)$$

an. Die linke Seite lässt sich partiell integrieren zu $[GF]_a^1 = 1 - G(a)F(a)$, und wir können die Gleichung weiter umformen zu

$$\int_a^1 F(au^{-1})G(u) dF(u) = F(a)(1 - G(a)). \quad (3.43)$$

Mit $F(au^{-1})G(u) = a^\theta u^{-\theta} u^\theta = a^\theta = F(a)$ folgt

$$\int_a^1 F(au^{-1})G(u) dF(u) = F(a) \int_a^1 dF(u) = F(a)(1 - F(a)) = F(a)(1 - G(a)).$$

Für die genannte Wahl von W und $(U_n)_{n \geq 1}$ ist die Integralgleichung (3.36) also offenbar erfüllt, und mit Lemma 3.20 folgt die Invarianz. \square

Dieses Ergebnis gilt es nun von der W -Split-Merge-Transformation auf die CCF-Transformation zu übertragen. Dies gelingt mittels des folgenden Satzes (vgl. [18], (4.1)):

3.21 Satz. *Seien $(p_n)_{n \geq 1}, (p'_n)_{n \geq 1} \in \Omega_1$ und seien $(q_n)_{n \geq 1}, (q'_n)_{n \geq 1} \in \Delta_\infty$. Dann gilt für die Transitionsmechanismen des CCF-Prozesses und des W -Split-Merge-Prozesses mit $W \sim U[0, 1]$ folgendes Diagramm:*

$$\begin{array}{ccc} (p_n)_{n \geq 1} & \xrightarrow{\text{CCF}} & (p'_n)_{n \geq 1} \\ \text{SBP} \downarrow \uparrow \text{sort} & & \uparrow \text{sort} \\ (q_n)_{n \geq 1} & \xrightarrow{W\text{-Split-Merge}} & (q'_n)_{n \geq 1}. \end{array} \quad (3.44)$$

Beweis. Wir zeigen, dass die Hintereinanderausführung von SBP, W -Split-Merge

und sort der Ausführung von CCF auf Ω_1 entspricht. Den zugehörigen Übergangskern für Ersteres bezeichnen wir mit \mathbb{P}_W , den für Letzteres mit \mathbb{P} .

Wir betrachten zunächst einen *Split*-Übergang $(p_n)_{n \geq 1} \mapsto (p'_n)_{n \geq 1}$. Für CCF erhalten wir nach Konstruktion $\mathbb{P}(p_i; \text{split}) = p_i^2$. Für \mathbb{P}_W gilt: Die Wahrscheinlichkeit, dass p_i durch SBP an den Anfang der Partition gesetzt wird (also $q_1 = p_i$), ist gerade p_i . Zur Aufteilung des ersten Abschnitts unter W -Split-Merge kommt es auf $\{W < p_i\}$, was nach Konstruktion ebenfalls die Wahrscheinlichkeit p_i hat. Die nachfolgende sort-Operation hat keinen Einfluss mehr auf die Zusammensetzung der Partition, lediglich auf die Anordnung. Es ergibt sich also

$$\mathbb{P}_W(p_i; \text{split}) = p_i^2 \equiv \mathbb{P}(p_i; \text{split}).$$

Für die Bruchstücke gilt: Bei CCF wird p_i in zwei neue Abschnitte up_i und $(1-u)p_i$ mit $u \sim U[0, 1]$ aufgeteilt. Bei W -Split-Merge wird p_i in W und $p_i - W$ aufgeteilt. Bedingt auf $\{W < p_i\}$ ist W gleichverteilt auf $[0, p_i]$. Wir erhalten also in beiden Fällen dieselbe Verteilung der Bruchstücklängen.

Beim *Merge*-Übergang gilt $\mathbb{P}(p_i, p_j; \text{merge}) = p_i p_j$. Für \mathbb{P}_W erhalten wir: Die Wahrscheinlichkeit, dass p_i durch SBP an den Anfang und p_j an die zweite Stelle der Partition gesetzt wird (d. h. $(q_1, q_2) = (p_i, p_j)$), beträgt $p_i \frac{p_j}{1-p_i}$. Zur Vereinigung von p_i und p_j unter W -Split-Merge kommt es auf $\{W \geq p_i\}$. Dieses Ereignis hat die Wahrscheinlichkeit $1 - p_i$, und wir erhalten somit

$$\mathbb{P}_W(p_i, p_j; \text{merge}) = p_i p_j \equiv \mathbb{P}(p_i, p_j; \text{merge}).$$

Alle anderen Abschnitte werden nicht verändert, allenfalls umgeordnet. Da sowohl bei CCF als auch bei der Hintereinanderausführung SBP, W -Split-Merge und sort als Letztes die ganze Partition nichtaufsteigend angeordnet wird, ist dies unerheblich. Die Verteilungen nach der Transition stimmen in beiden Fällen überein. \square

Es ist offensichtlich, dass eine invariante Verteilung für die untere Transition nach entsprechender Transformation durch SBP bzw. sort eine invariante Verteilung für die obere Transition liefert. Wie bereits gezeigt, ist die $GEM(\theta)$ -Verteilung invariant unter W -Split-Merge mit $W \sim \beta(1, \theta)$ (mit $\theta > 0$, vgl. Satz 3.19). Wählen wir nun $\theta = 1$, so gilt $\beta(1, \theta) \equiv U[0, 1]$. Aus dem vorhergehenden Satz in Kombination mit

Satz 3.15 und Korollar 3.17 erhalten wir schließlich:

3.22 Korollar. *Die Poisson-Dirichlet-Verteilung mit Parameter $\theta = 1$ \mathcal{PD}_1 ist invariant unter der CCF-Transition auf Ω_1 .*

4 Kopplung von DCF und CCF

Um die Eindeutigkeit der invarianten Verteilung von CCF durch Rückgriff auf die eindeutige invariante Verteilung von $\text{DCF}^{(n)}$ zu zeigen, müssen wir eine geeignete Beziehung zwischen beiden Prozessen herstellen. Wir approximieren dazu eine kontinuierliche Partition durch eine (mit dem Faktor $1/n$ skalierte) diskrete Partition und verwenden diese Partitionen als Startpartitionen für einen CCF- und einen $\text{DCF}^{(n)}$ -Prozess auf einem gemeinsamen Wahrscheinlichkeitsraum. Die Transition dieser Prozesse wird so gewählt, dass sich in der kontinuierlichen und in der diskreten Partition möglichst ähnliche Übergänge ergeben. Haben wir für die Startpartitionen zudem noch eine Kontrolle über die Anzahl der kleinen Abschnitte, so erhalten wir schließlich das folgende Ergebnis (vgl. [7], S. 919):

4.1 Theorem. *Für alle $\mu \in \mathcal{M}_1(\Omega_1)$ und alle $\alpha < 1/2$, die*

$$\int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) < \infty \quad (4.1)$$

erfüllen, lassen sich für alle $n \geq 1$ eine CCF-Markov-Kette $(p(k))_{k \geq 0}$ mit Anfangsverteilung μ und eine $\text{DCF}^{(n)}$ -Markov-Kette $(\ell(k))_{k \geq 0}$ auf einem gemeinsamen Wahrscheinlichkeitsraum mit Wahrscheinlichkeitsmaß $Q_\mu^{(n)}$ definieren, so dass gilt

$$\lim_{n \rightarrow \infty} Q_\mu^{(n)}[N_{\ell(0)} \geq n^\beta] = 0 \quad \text{für alle } 1 > \beta > \alpha \quad (4.2)$$

und

$$\lim_{n \rightarrow \infty} E_\mu^{(n)} \left[\left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 \right] = 0 \quad \text{für alle } 0 < \beta < 1/2, \quad (4.3)$$

wobei $E_\mu^{(n)}$ den Erwartungswert bezüglich $Q_\mu^{(n)}$ bezeichnet.

Die Kontrolle über die Anzahl der kleinen Abschnitte in der Startpartition von CCF in Gleichung (4.1) setzt sich fort zu einer entsprechenden asymptotischen Aussage

über die Startpartition des $\text{DCF}^{(n)}$ -Prozesses (Gleichung (4.2)). In Gleichung (4.3) findet sich das eigentliche Kopplungsergebnis: Die mittlere Abweichung der beiden Prozesse zu einem Zeitpunkt kleiner \sqrt{n} verschwindet für $n \rightarrow \infty$.

Um den Beweis zu erbringen, werden wir zunächst den gemeinsamen Zustandsraum einführen und dann die Konstruktion der Markov-Kette darauf beschreiben, die uns einen CCF- und einen diesen approximierenden $\text{DCF}^{(n)}$ -Prozess liefert.

4.1 Der gemeinsame Zustandsraum $\Omega_{cde}^{(n)}$

Als gemeinsamen Zustandsraum wählen wir

$$\begin{aligned} \Omega_{cde}^{(n)} := \{ (c, d, e) \mid & c: [0, n) \longrightarrow \mathbb{Z} \text{ messbar, } \mathbb{Z} \setminus c([0, n)) \text{ unendlich,} \\ & d: \{1, \dots, n\} \longrightarrow \mathbb{Z}, e \in \{0, 1\} \}. \end{aligned} \quad (4.4)$$

Jedes Element von $\Omega_{cde}^{(n)}$ enthält gleichsam eine um den Faktor n skalierte kontinuierliche Partition des Einheitsintervalls sowie eine diskrete Partition: c ist eine Labelfunktion, die Borel-Mengen von $[0, n)$ ganze Zahlen als Label zuordnet. Eine Partition $p \in \Omega_1$ erhält man daraus durch die Funktion

$$\pi_c: \Omega_{cde}^{(n)} \longrightarrow \Omega_1, \quad (4.5)$$

die jedem Label $k \in \mathbb{Z}$ das Lebesgue-Maß der Urbildmenge in $[0, n)$ zuordnet (skaliert mit dem Faktor $1/n$) und diese Werte anschließend nach nichtaufsteigender Größe sortiert, d. h.

$$\pi_c(c, d, e) := \text{sort} \left(\left(\frac{1}{n} \mathbb{X}(c^{-1}(\{k\})) \right)_{k \in \mathbb{Z}} \right). \quad (4.6)$$

Hierbei bezeichnet \mathbb{X} das Lebesgue-Maß und $\text{sort}((x_i)_{i \geq 1})$ wiederum die Sequenz, die man erhält, wenn man die x_i nach nichtaufsteigender Größe anordnet, wobei alle Einträge gleich null ignoriert werden, falls unendlich viele positive x_i vorhanden sind (vgl. auch S. 27). Zwei Punkte $x, y \in [0, n)$ gehören damit genau dann zum selben Abschnitt in der Partition von $[0, n)$, wenn ihre Labels übereinstimmen, also $c(x) = c(y)$ gilt.

d stellt ebenfalls eine Labelfunktion dar, die den diskreten „Atomen“ $1, \dots, n$ ganze

Zahlen als Label zuordnet. Eine (diskrete) Partition wird hieraus in ganz analoger Weise mittels einer Funktion

$$\pi_d: \Omega_{cde}^{(n)} \longrightarrow \mathcal{P}_n \quad (4.7)$$

gewonnen, die die Anzahl der zu jedem $k \in \mathbb{Z}$ gehörigen Punkte berechnet und diese Werte anschließend nach nichtaufsteigender Größe sortiert:

$$\pi_d(c, d, e) := \text{sort} \left(\left(\#d^{-1}(\{k\}) \right)_{k \in \mathbb{Z}} \right). \quad (4.8)$$

Der *Kopplungsindikator* e zeigt schließlich an, ob Kopplung zwischen kontinuierlicher und diskreter Partition besteht ($e = 0$) oder nicht ($e = 1$). Dabei bedeutet „Kopplung“ nicht, dass bspw. die l_1 -Abweichung zwischen $\pi_c(c, d, e)$ und $\frac{1}{n}\pi_d(c, d, e)$ beschränkt ist oder die Partitionen gar identisch sind. Vielmehr definieren wir die beiden Partitionen als gekoppelt, solange wir begründeterweise annehmen können, dass die Partitionen einander ausreichend ähneln. Was im Einzelnen darunter zu verstehen ist, werden wir im Abschnitt 4.3 darlegen, wenn wir die Markov-Kette auf $\Omega_{cde}^{(n)}$ einführen und der Indikator e seine eigentliche Bedeutung erlangt.

4.2 Approximation von kontinuierlichen Partitionen durch diskrete

Für die angestrebte Kopplung von CCF- und DCF⁽ⁿ⁾-Prozess ist es wichtig, die Partitionen so zu wählen, dass sie eine gewisse Ähnlichkeit zueinander aufweisen. Zu einer gegebenen kontinuierlichen Partitionen $p \in \Omega_1$ werden wir daher nun eine diskrete Partition konstruieren, die p approximiert. Dies geschieht mithilfe zweier Funktionen c und d , die so gewählt werden, dass $\pi_c(c, d, e) = p$ gilt und $\pi_d(c, d, e)$ eine diskrete Approximation der skalierten Partition $np \in n\Omega_1$ darstellt. Das Finden von c und d bewerkstelligt für $n \in \mathbb{N}$ eine Funktion

$$\Phi^{(n)} = (\Phi_1^{(n)}, \Phi_2^{(n)}, 0): \Omega_1 \longrightarrow \Omega_{cde}^{(n)}. \quad (4.9)$$

$\Phi_1^{(n)}(p)$ ordnet jedem Abschnitt np_j von $[0, n)$ das Label j zu:

$$\Phi_1^{(n)}(p)(x) := \sum_{j \geq 1} j \mathbf{1}_{[0, np_j) + n \sum_{i=1}^{j-1} p_i}(x), \quad x \in [0, n). \quad (4.10)$$

Anhand der Definition ist unmittelbar klar, dass $\pi_c(\Phi_1^{(n)}(p), \Phi_2^{(n)}(p), 0) = p$ gilt. Die Funktion $\Phi_2^{(n)}(p)$ approximiert, indem sie an der Stelle m das Label von $\Phi_1^{(n)}(p)$ an der Stelle $m - 1$ übernimmt:

$$\Phi_2^{(n)}(p)(m) := \Phi_1^{(n)}(p)(m - 1), \quad m \in \{1, \dots, n\}. \quad (4.11)$$

Die so gewonnene Funktion $d := \Phi_2^{(n)}(p)$ ist so zu verstehen, dass das Label für das Atom m auf das Intervall $(m - 1, m] \subset (0, n]$ ausgedehnt anzusehen ist. Wir erhalten dadurch eine diskrete Partition von $(0, n]$, die direkt mit der kontinuierlichen Partition $c := \Phi_1^{(n)}(p)$ verglichen werden kann¹. Dies ist in der folgenden Abbildung veranschaulicht:

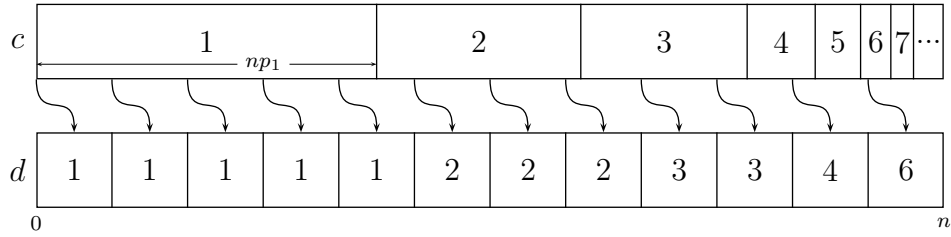


Abbildung 4.1: Approximation einer kontinuierlichen Partition $c = \Phi_1^{(n)}(p)$ durch eine diskrete Partition $d = \Phi_2^{(n)}(p)$ (für $n = 12$ und ein $p \in \Omega_1$). Die Pfeile deuten an, wie die Labels von c für d übernommen werden.

Die dritte Komponente von $\Phi^{(n)}$ ist der Wert des zu $\Phi_1^{(n)}(p)$ und $\Phi_2^{(n)}(p)$ gehörenden Kopplungsindikators. Dass dieser für alle $p \in \Omega_1$ gleich null ist, bedeutet, dass die zugehörigen Partitionen immer als gekoppelt definiert werden. Dies ist offenbar sinnvoll, da durch die Approximation in der Tat Ähnlichkeit zwischen den Partitionen besteht.

¹Es wird hier nicht immer genau unterschieden, welcher Typ Intervallgrenze, d. h. offen oder geschlossen, vorliegt. Dies ist aber für die im Folgenden eingeführten Wahrscheinlichkeiten ohne Belang.

Ist nun auf Ω_1 ein Wahrscheinlichkeitsmaß μ gegeben, so erhalten wir vermöge $\Phi^{(n)}$ daraus ein Bildmaß (vgl. [2], S. 22)

$$\mu^{\Phi^{(n)}} := \mu \circ \Phi^{(n)-1}, \quad (4.12)$$

das wir im Folgenden mit $Q_\mu^{(n)}$ bezeichnen. Mit μ ist $Q_\mu^{(n)}$ ebenfalls ein Wahrscheinlichkeitsmaß.

4.3 Markov-Kette auf $\Omega_{cde}^{(n)}$

Nun stehen alle Hilfsmittel bereit, um auf $\Omega_{cde}^{(n)}$ eine Markov-Kette $(c_k, d_k, e_k)_{k \geq 0}$ zu definieren. Diese Kette wird so gewählt, dass in ihr ein CCF- und ein DCF $^{(n)}$ -Prozess mit möglichst analoger Transitionen in dem Sinne „enthalten“ sind, dass die Projektionen π_c und π_d durch

$$p(k) := \pi_c(c_k, d_k, e_k) \quad \text{und} \quad \ell(k) := \pi_d(c_k, d_k, e_k) \quad (4.13)$$

eine CCF-Markov-Kette $(p(k))_{k \geq 0}$ auf Ω_1 sowie eine DCF $^{(n)}$ -Markov-Kette $(\ell(k))_{k \geq 0}$ auf \mathcal{P}_n liefern. Das Ziel hierbei ist, dass die $\ell(k)$ den $p(k)$ möglichst ähnlich sind. Wir sagen, dass *Kopplung* zwischen beiden Prozesse besteht, solange wir davon ausgehen können, dass $\ell(k)$ und $p(k)$ einander hinreichend ähnlich sind. Auskunft über den Kopplungszustand gibt das „Fähnchen“ e_k ($= 0$ für gekoppelt bzw. $= 1$ für Kopplung zusammengebrochen).

Die Definition von $(c_k, d_k, e_k)_{k \geq 0}$ ist nun wie folgt: Als Anfangsverteilung wählen wir das oben beschriebene $Q_\mu^{(n)}$. Damit liegt nach Wahl von $\Phi^{(n)}$ im Anfangszustand immer Kopplung vor: Startet unser Prozess im Zustand $(c_0, d_0, e_0) := \Phi^{(n)}(p)$ für ein $p \in \Omega_1$, so ist die Projektion $\ell(0) = \pi_d(\Phi^{(n)}(p))$ eine diskrete Approximation von p ; beide Partitionen sind sich also hinreichend ähnlich.

Den Übergangskern von $(c_k, d_k, e_k)_{k \geq 0}$ auf $\Omega_{c,d,e}^{(n)}$ wählen wir gerade so, dass die beiden Projektionen π_c und π_d einen CCF- bzw. DCF $^{(n)}$ -Prozess liefern. Hierzu gilt es, die in Kapitel 2 beschriebenen Split-Merge-Prozesse entsprechend auf $\Omega_{c,d,e}^{(n)}$ umzusetzen.

Ist (c, d, e) ein beliebiger gegebener Zustand, so sei $(\bar{c}, \bar{d}, \bar{e})$ der Zustand, zu dem

der Prozess als Nächstes springt. Wir beschreiben zunächst die Konstruktion für die kontinuierliche Partition² c . Seien dazu ξ_1 und ξ_2 zwei unabhängige, auf $[0, n)$ gleichverteilte Zufallsvariablen. Die neue kontinuierliche Partition \bar{c} wird dann wie folgt definiert:

Falls $c(\xi_1) \neq c(\xi_2)$, gelte

$$\bar{c}(x) := \begin{cases} c(\xi_1), & \text{falls } c(x) = c(\xi_2), \\ c(x), & \text{sonst,} \end{cases} \quad (4.14)$$

und falls $c(\xi_1) = c(\xi_2)$, gelte

$$\bar{c}(x) := \begin{cases} \text{neu}(c, d), & \text{falls } c(x) = c(\xi_1) \text{ und } x > \xi_1, \\ c(x), & \text{sonst.} \end{cases} \quad (4.15)$$

$\text{neu}(c, d)$ bezeichne hierbei ein beliebiges $k \in \mathbb{Z}$ mit $k \notin c([0, n)) \cup d(\{1, \dots, n\})$. Nach Wahl von c ist es immer möglich, ein solches k zu finden.

Erhalten also die Punkte ξ_i unter c verschiedene Labels, so werden die beiden zugehörigen Urbildmengen $c^{-1}(\{c(\xi_1)\})$ und $c^{-1}(\{c(\xi_2)\})$ zusammengefügt und erhalten beide das Label $c(\xi_1)$ (Merge).

Erhalten hingegen die Punkte ξ_i unter c dasselbe Label, so wird die zugehörige Urbildmenge $c^{-1}(\{c(\xi_1)\}) = c^{-1}(\{c(\xi_2)\})$ aufgeteilt (Split). Als Teilungspunkt wird ξ_1 wiederverwendet, der bedingt auf $c(\xi_1) = k$ für ein $k \in \mathbb{Z}$ gleichverteilt ist auf $c^{-1}(\{k\})$. Der links von ξ_1 liegende Teil behält sein Label, während der rechts von ξ_1 liegende Teil ein neues Label k erhält. Welchem Teil der Teilungspunkt selbst zugeschlagen wird, ist unerheblich, da er eine Lebesgue-Nullmenge bildet; wir vereinbaren, dass er dem linken Teil angefügt wird.

Es bleibt noch nachzuweisen, dass $(p(k))_{k \geq 0}$ ein CCF-Prozess ist: Die Wahrscheinlichkeit, dass durch ξ_1 der i -te Abschnitt in $\pi_c(c, d, e)$ ausgewählt wird, ist gerade

² Wir werden im Folgenden nicht immer exakt zwischen den Labelfunktionen c, d und den von ihnen erzeugten Partitionen π_c, π_d unterscheiden, und, wenn keine Verwechslungen zu befürchten sind, c und d selbst als „Partition“ bezeichnen.

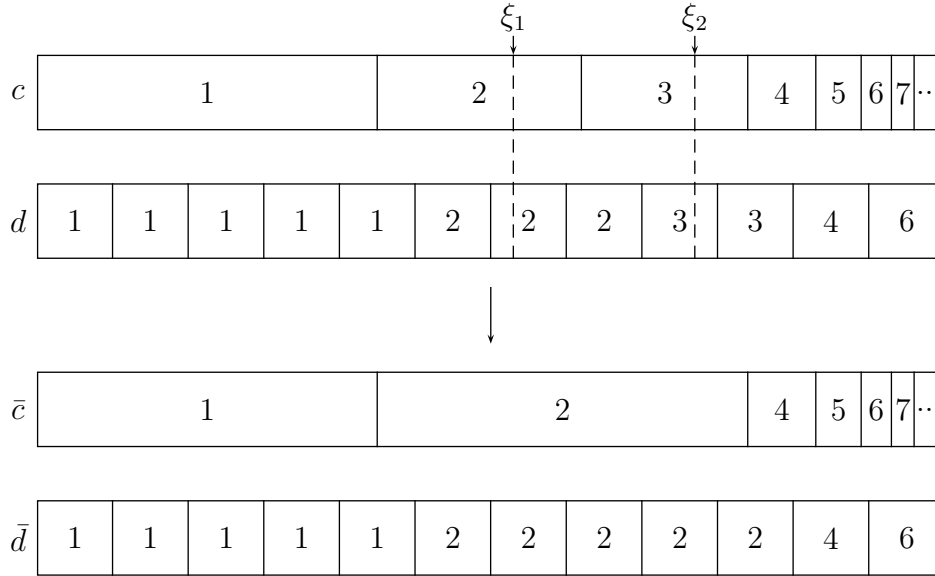


Abbildung 4.2: In den Partitionen c und d findet ein Merge-Übergang statt: Die Abschnitte mit den Labels 2 und 3 werden vereinigt.

$\mathbb{K}(c^{-1}(\{i\}))/n$, was mit $(\pi_c(c, d, e))_i$ übereinstimmt, da $\pi_c(c, d, e)$ ja bereits sortiert ist. Dieselbe Überlegung gilt für ξ_2 . Damit ist gezeigt, dass das Auswählen der Abschnitte längenverzerrt vonstatten geht.

Je nachdem, ob verschiedene Abschnitte oder zweimal derselbe ausgewählt wird, werden entweder zwei Abschnitte mit den Längen $\mathbb{K}(c^{-1}(\{c(\xi_1)\}))$ u. $\mathbb{K}(c^{-1}(\{c(\xi_2)\}))$ zu einem Abschnitt der Länge $\mathbb{K}(c^{-1}(\{c(\xi_1)\})) + \mathbb{K}(c^{-1}(\{c(\xi_2)\})) = \mathbb{K}(\bar{c}^{-1}(\{c(\xi_1)\}))$ zusammengefügt, oder es wird ein Abschnitt der Länge $\mathbb{K}(c^{-1}(\{c(\xi_1)\}))$ in zwei Abschnitte mit den Längen $\mathbb{K}(\bar{c}^{-1}(\{c(\xi_1)\}))$ und $\mathbb{K}(\bar{c}^{-1}(\{\text{neu}(c, d)\}))$ aufgeteilt.

Wir müssen noch zeigen, dass der Split-Übergang den betreffenden Abschnitt gleichmäßig aufteilt: Dies ist jedoch sofort klar, da die Gleichverteilung eingeschränkt auf $c^{-1}(\{c(\xi_1)\})$ wiederum eine Gleichverteilung auf dieser Menge liefert.

Damit ist gezeigt, dass $(p(k))_{k \geq 0}$ ein CCF-Prozess ist.

Als Nächstes ist nun die Konstruktion für die diskrete Partition d vorzunehmen. Die neue Partition \bar{d} wird folgendermaßen definiert; wie bereits erwähnt, wird das Label des Punktes m im Sinne der Approximation der kontinuierlichen Partition mittels der obere Gaußklammer $\lceil \cdot \rceil$ auf das Intervall $(m - 1, m]$ ausgedehnt:

Falls $d(\lceil \xi_1 \rceil) \neq d(\lceil \xi_2 \rceil)$, gelte

$$\bar{d}(m) := \begin{cases} d(\lceil \xi_1 \rceil), & \text{falls } d(m) = d(\lceil \xi_2 \rceil), \\ d(m), & \text{sonst,} \end{cases} \quad (4.16)$$

und falls $d(\lceil \xi_1 \rceil) = d(\lceil \xi_2 \rceil)$ und $\lceil \xi_1 \rceil \neq \lceil \xi_2 \rceil$, gelte

$$\bar{d}(m) := \begin{cases} \text{neu}(c, d), & \text{falls } d(m) = d(\lceil \xi_1 \rceil) \text{ und } m > \lceil \xi_1 \rceil \\ & \text{oder } m = \lceil \xi_1 \rceil \text{ und } \xi_1 < \lfloor \xi_1 \rfloor + \frac{\#d^{-1}(\{d(\lceil \xi_1 \rceil)\}) \cap [1, \xi_1]}{\#d^{-1}(\{d(\lceil \xi_1 \rceil)\}) - 1}, \\ d(m), & \text{sonst.} \end{cases} \quad (4.17)$$

Erhalten also die Punkte $\lceil \xi_1 \rceil$ und $\lceil \xi_2 \rceil$ unter d verschiedene Labels, so ist die Konstruktion analog zu der im kontinuierlichen Fall: Die beiden zugehörigen Urbildmengen $d^{-1}(\{d(\lceil \xi_1 \rceil)\})$ und $d^{-1}(\{d(\lceil \xi_2 \rceil)\})$ werden zusammengefügt und erhalten beide das Label $d(\lceil \xi_1 \rceil)$ (Merge).

Erhalten hingegen die Punkte $\lceil \xi_i \rceil$ unter d dasselbe Label, so ist das Vorgehen komplizierter als im kontinuierlichen Fall, was darin begründet zu sehen ist, dass zum einen das Teilungsatom keine (Zählmaß)-Nullmenge mehr darstellt. Zum anderen kann es passieren, dass nicht gesichert ist, dass ein aufzuspaltender Abschnitt aus mehr als einem Atom besteht:

Stimmen zwar die d -Labels der Punkte überein, nicht aber die gewählten Atome in $\{1, \dots, n\}$, so besteht die zugehörige Urbildmenge $d^{-1}(\{d(\lceil \xi_1 \rceil)\}) = d^{-1}(\{d(\lceil \xi_2 \rceil)\})$ aus mindestens zwei Atomen und kann somit aufgeteilt werden (Split). Als Teilungsatom wird wie im kontinuierlichen Fall $\lceil \xi_1 \rceil$ wiederverwendet, das bedingt auf $d(\lceil \xi_1 \rceil) = k$ für ein $k \in \mathbb{Z}$ gleichverteilt ist auf $d^{-1}(\{d(\lceil \xi_1 \rceil)\})$. Der links von $\lceil \xi_1 \rceil$ liegende Teil behält sein Label, während der rechts liegende Teil ein neues Label $\text{neu}(c, d)$ erhält. Welchem Teil das Teilungsatom selbst zugeschlagen wird, ist nun von entscheidender Bedeutung; die Konstruktion (4.17) ist so gewählt, dass es sowohl links als auch rechts zugeschlagen werden kann, und zwar genau so, dass die Aufteilung wie für den DCF⁽ⁿ⁾-Prozess gefordert erfolgt, was wir weiter unten noch näher ausführen werden.

Im Falle, dass sowohl derselbe Bereich als auch dasselbe Atom gewählt wird, d. h.

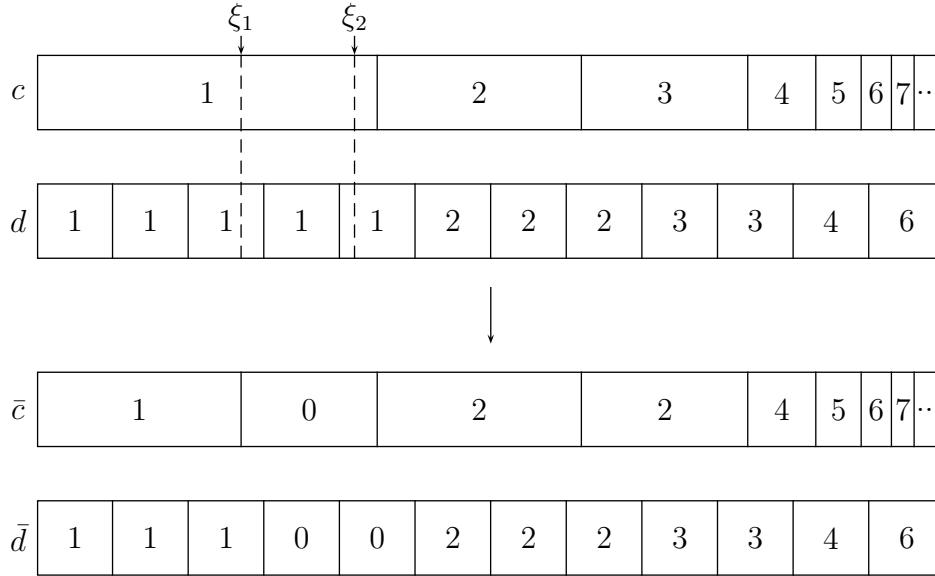


Abbildung 4.3: In den Partitionen c und d findet ein Split-Übergang statt: Der Abschnitt mit dem Label 1 wird in zwei Abschnitte mit den Labels 1 und $0 = \text{neu}(c, d)$ aufgespalten.

$d([\xi_1]) = d([\xi_2])$ und $[\xi_1] = [\xi_2]$, ist nicht mehr sichergestellt, dass der Bereich sinnvoll aufgeteilt werden kann, da er u. U. nur aus einem Atom besteht. In diesem Falle werden ξ_1 und ξ_2 verworfen, und es wird eine „Ersatzziehung“ durchgeführt. Dazu verwenden wir zwei Zufallsvariablen ζ_1 und ζ_2 , die unabhängig von ξ_1, ξ_2 und allen anderen Zufallsvariablen zu wählen sind. Weiter seien (ζ_1, ζ_2) gleichverteilt auf $[0, n)^2 \setminus \bigcup_{j=1}^n [j-1, j)^2$, es werden also zwei zufällige Atome *ohne* Zurücklegen aus $\{1, \dots, n\}$ gewählt. Die neue Partition \bar{d} wird dann wie oben definiert, wobei (ζ_1, ζ_2) für (ξ_1, ξ_2) einzusetzen ist. Nach Konstruktion ist nun der Split-Übergang in jeden Fall sinnvoll durchführbar.

Es bleibt noch nachzuweisen, dass $(\ell(k))_{k \geq 0}$ ein $\text{DCF}^{(n)}$ -Prozess ist: Die Wahrscheinlichkeit, dass durch $[\xi_1]$ bzw. $[\zeta_1]$ der i -te Abschnitt in $\pi_d(c, d, e)$ ausgewählt wird, ist gerade $\#(d^{-1}(\{i\}))/n$, was mit $(\pi_d(c, d, k))_i$ übereinstimmt, da $\pi_c(c, d, k)$ ja bereits sortiert ist. Dieselbe Überlegung greift auch für $[\xi_2]$ bzw. $[\zeta_2]$. Damit ist gezeigt, dass das Auswählen der Abschnitte mit den korrekten Wahrscheinlichkeiten geschieht. Durch die, wenn notwendig, zur Anwendung kommende „Ersatzziehung“ ist zudem sichergestellt, dass das Auswählen der beiden Abschnitte wie gefordert

ohne Zurücklegen stattfindet.

Je nachdem, ob verschiedene Abschnitte oder zweimal derselbe ausgewählt wird, werden entweder zwei Abschnitte mit den Längen $\#d^{-1}(d(\lceil \xi_1 \rceil))$ und $\#d^{-1}(d(\lceil \xi_2 \rceil))$ zu einem Abschnitt mit der Länge $\#d^{-1}(d(\lceil \xi_1 \rceil)) + \#d^{-1}(d(\lceil \xi_2 \rceil)) = \#\bar{d}^{-1}(d(\lceil \xi_1 \rceil))$ zusammengefügt oder es wird ein Abschnitt mit der Länge $\#d^{-1}(d(\lceil \xi_1 \rceil))$ in zwei Abschnitte mit den Längen $\#\bar{d}^{-1}(d(\lceil \xi_1 \rceil))$ und $\#\bar{d}^{-1}(d(\text{neu}(c, d)))$ aufgeteilt.

Wir müssen nun nur noch zeigen, dass der Split-Übergang dergestalt ist, wie für $\text{DCF}^{(n)}$ gefordert. Dies lässt sich einfach nachrechnen; wir führen diese Rechnung mit den ξ_i durch, im Falle der Ersatzziehung sind dafür die ζ_i einzusetzen, was aber an dem Ergebnis nichts ändert.

Wir nehmen zunächst an, dass die Atome der Menge $d^{-1}(\{d(\lceil \xi_1 \rceil)\})$ nicht über die ganze Menge $\{1, \dots, n\}$ verteilt sind, wie es üblicherweise der Fall sein wird, sondern am linken Ende konzentriert sind: $d^{-1}(\{d(\lceil \xi_1 \rceil)\}) = \{1, \dots, a\}$ mit $a := \#d^{-1}(\{d(\lceil \xi_1 \rceil)\})$. Gemäß Definition wird diese Menge genau dann in die Mengen $\{1, \dots, j\}$ und $\{j+1, \dots, n\}$ aufgeteilt, wenn entweder j Teilungsatom ist und der linken Menge zugeschlagen wird, d. h. falls gilt³

$$\lceil \xi_1 \rceil = j \quad \text{und} \quad \xi_1 \geq \lfloor \xi_1 \rfloor + \frac{\#(\{1, \dots, a\} \cap [1, \xi_1])}{\#\{1, \dots, a\} - 1} = j - 1 + \frac{j - 1}{a - 1},$$

oder wenn $j+1$ das Teilungsatom ist, aber der rechten Menge zugeschlagen wird, also falls gilt

$$\lceil \xi_1 \rceil = j + 1 \quad \text{und} \quad \xi_1 < \lfloor \xi_1 \rfloor + \frac{\#(\{1, \dots, a\} \cap [1, \xi_1])}{\#\{1, \dots, a\} - 1} = j + \frac{j}{a - 1}.$$

Dies lässt sich zusammenfassen als

$$\begin{aligned} j - 1 + \frac{j - 1}{a - 1} &\leq \xi_1 < j + \frac{j}{a - 1} \\ \iff j - 1 + \frac{j - 1}{a - 1} &\leq \xi_1 < j - 1 + \frac{j - 1}{a - 1} + \frac{a}{a - 1}. \end{aligned}$$

Die Wahrscheinlichkeit für dieses Ereignis, bedingt auf $\{\xi_1 \in [0, a)\}$, ist also gerade

³Der Fall $\lceil \xi_1 \rceil = \xi_1$ wird hier und im Folgenden nicht weiter betrachtet, da seine Wahrscheinlichkeit verschwindet.

$\frac{1}{a-1}$. Das bedeutet, dass die $a - 1$ Möglichkeiten, die Menge $\{1, \dots, a\}$ aufzuteilen, alle dieselbe Wahrscheinlichkeit, nämlich $\frac{1}{a-1}$, haben.

Der allgemeine Fall, wo $d^{-1}(\{d(\lceil \xi_1 \rceil)\})$ eine beliebige Menge ist, ist etwas aufwändiger zu handhaben, kann jedoch analog gezeigt werden, weshalb wir hier darauf verzichten.

Der Split-Übergang verläuft also wie gefordert, und $(\ell(k))_{k \geq 0}$ ist damit ein $\text{DCF}^{(n)}$ -Prozess.

In den Abbildungen 4.2 und 4.3 sind ein Merge- bzw. ein Split-Übergang in den Partitionen c und d dargestellt.

Schließlich bleibt noch der Kopplungs-Indikator e zu definieren, dessen nächster Zustand \bar{e} wie folgt festgesetzt wird:

$$\bar{e} = \begin{cases} 1, & \text{falls } c(\xi_1) \neq d(\xi_1) \text{ oder } c(\xi_2) \neq d(\xi_2) \text{ oder } \lceil \xi_1 \rceil = \lceil \xi_2 \rceil, \\ 0 & \text{sonst.} \end{cases} \quad (4.18)$$

Das bedeutet: Per Definition von $\Phi^{(n)}$ sind die Prozesse im Anfangszustand gekoppelt (vgl. Gleichung (4.10)). Ob diese Kopplung nach einem Übergang zum nächsten Zustand immer noch besteht, hängt davon ab, ob wir davon ausgehen können, dass der Übergang von kontinuierlicher und diskreter Partition ausreichend ähnlich ist: Dies ist der Fall, wenn in beiden Partitionen *korrespondierende* Abschnitte aufgeteilt bzw. vereinigt werden. In allen anderen Fällen, d. h. wenn nicht korrespondierende Abschnitte ausgewählt werden ($c(\xi_1) \neq d(\xi_1)$ bzw. $c(\xi_2) \neq d(\xi_2)$) oder wenn $\text{DCF}^{(n)}$ -Split nicht korrekt ausgeführt werden kann und auf die Ersatzziehung zurückgegriffen werden muss ($\lceil \xi_1 \rceil = \lceil \xi_2 \rceil$), ist nicht mehr gesichert, dass kontinuierliche und diskrete Partition noch ausreichend ähnlich sind; die Kopplung wird dann als zusammengebrochen definiert. In Abbildung 4.4 sind die verschiedenen Möglichkeiten der Entkopplung illustriert.

Wir setzen schließlich

$$\tau := \inf\{k \geq 1 : e_k = 1\} \quad (4.19)$$

für die *Entkopplungszeit* der beiden Prozesse $(p(k))_{k \geq 0}$ und $(\ell(k))_{k \geq 0}$. Es ist offensichtlich, dass τ nach Konstruktion eine Stoppzeit ist.

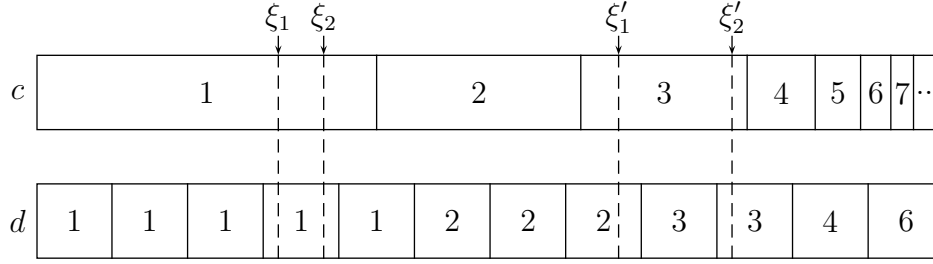


Abbildung 4.4: Zwei Situationen, in denen es zur Entkopplung kommt: In d wird zweimal dasselbe Atome gewählt (ξ_1, ξ_2) bzw. in c und d werden nicht korrespondierende Abschnitte ausgewählt (ξ'_1, ξ'_2) .

4.4 Beweis des Theorems

Beweis von (4.2). Um für $n \in \mathbb{N}$ und $p \in \Omega_1$ die Gesamtanzahl $N_{\ell(0)}$ der Abschnitte in $\ell(0) = \pi_d(\Phi^{(n)}(p))$ abzuschätzen, machen wir uns zunutze, dass alle Abschnitte in p mit $p_i < 1/n$ in der skalierten Partition np eine Gesamtlänge von $\sum_{i \geq 1} np_i \mathbf{1}_{\{np_i < 1\}}$ einnehmen und daher nach Konstruktion höchstens zu $\lfloor \sum_{i \geq 1} np_i \mathbf{1}_{\{np_i < 1\}} \rfloor$ Abschnitten (jeweils mit der Länge eins) in $\ell(0)$ führen können. Allgemein gilt zunächst

$$N_{\ell(0)} = \underbrace{\#\Phi_2^{(n)}(p)(\{1, \dots, n\})}_{\substack{\text{Anzahl der verschiedenen Bildpunkte} \\ \text{(Labels) von } \{1, \dots, n\} \text{ in } \mathbb{Z}}} . \quad (4.20)$$

Die p_i mit $p_i \geq 1/n$ haben unter der Skalierung mit n eine Länge ≥ 1 und führen daher alle jeweils zu einem Bildpunkt in \mathbb{Z} ; den Rest schätzen wir wie oben ab ($\alpha < 1/2$):

$$\begin{aligned} N_{\ell(0)} &\leq \sum_{i \geq 1} (\mathbf{1}_{\{np_i \geq 1\}} + np_i \mathbf{1}_{\{np_i < 1\}}) \\ &\leq \sum_{i \geq 1} ((np_i)^\alpha \mathbf{1}_{\{np_i \geq 1\}} + (np_i)^\alpha \mathbf{1}_{\{np_i < 1\}}) \\ &= n^\alpha \sum_{i \geq 1} p_i^\alpha. \end{aligned} \quad (4.21)$$

Unter Verwendung des Transformationssatzes (vgl. [2], S. 57) für $\Phi^{(n)}: \Omega_1 \longrightarrow \Omega_{cde}^{(n)}$

und mithilfe der Voraussetzung $\int \sum_{i \geq 1} p_i^\alpha d\mu < \infty$ erhalten wir

$$\begin{aligned} \frac{1}{n^\alpha} E_\mu^{(n)}[N_{\pi_d(\cdot)}] &= \frac{1}{n^\alpha} \int_{\Omega_{cde}^{(n)}} N_{\pi_d(c,d,e)} Q_\mu^{(n)}(d(c,d,e)) \\ &= \frac{1}{n^\alpha} \int_{\Omega_1} N_{\pi_d(\Phi^{(n)}(p))} \mu(dp) \\ &\leq \int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha d\mu < \infty, \end{aligned}$$

also

$$E_\mu^{(n)}[N_{\pi_d(\cdot)}] = O(n^\alpha). \quad (4.22)$$

Der Übersichtlichkeit halber schreiben wir $E_\mu^{(n)}[N_{\ell(0)}]$ für $E_\mu^{(n)}[N_{\pi_d(\cdot)}]$. Dank der Markov-Ungleichung ergibt sich

$$Q_\mu^{(n)}[N_{\ell(0)} > n^\beta] \leq n^{-\beta} E_\mu^{(n)}[N_{\ell(0)}] = O(n^{\alpha-\beta})$$

und damit für alle $\beta > \alpha$

$$Q_\mu^{(n)}[N_{\ell(0)} > n^\beta] \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (4.23)$$

□

Beweis von (4.3). Wir bezeichnen mit

$$\rho_k := \mathbb{X}(\{x \in [0, n) : c_k(x) \neq d_k(\lceil x \rceil)\}), \quad k \geq 0 \quad (4.24)$$

die *Abweichung von c_k und d_k* , also das Maß der Teilmenge von $[0, n)$, wo die Labels der kontinuierlichen und der diskreten Partition nicht übereinstimmen. In der diskreten Partition wird dabei – wie bereits erwähnt – das Label des Atoms k als Label für das Intervall $[k-1, k)$ interpretiert. Für $k = 0$ erhalten wir den Rundungsfehler, der durch die Approximation von c_0 durch d_0 entsteht; seine Größe ist gerade die Länge des grau eingefärbten Bereichs in Abbildung 4.5.

Nun ist klar, dass gemäß Konstruktion jeder Abschnitt in d_0 höchstens eine Abweichung der Länge eins gegenüber c_0 hervorrufen kann, wir erhalten also

$$\rho_0 \leq N_{\ell(0)}. \quad (4.25)$$

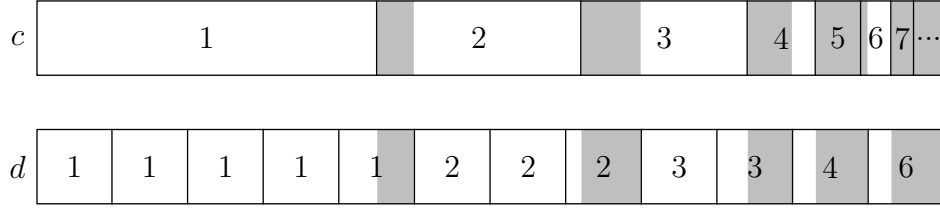


Abbildung 4.5: Eine kontinuierliche Partition c und eine diskrete Partition d . Die graue Unterlegung kennzeichnet den Bereich, in dem die Labels von c und d nicht übereinstimmen.

Solange c_k und d_k noch gekoppelt sind, d. h. solange $k < \tau$, kann ρ_k in jedem Schritt höchstens um 1 zunehmen: Bei einem Merge-Übergang nimmt ρ_k überhaupt nicht zu, sondern allenfalls ab. Bei einem Split-Übergang wird die kontinuierliche Partition bei ξ_1 unterteilt, die diskrete Partition hingegen bei $\lfloor \xi_1 \rfloor$ bzw. bei $\lceil \xi_1 \rceil$; der neu entstehende Fehler beträgt also maximal

$$\mathbb{X}((\lfloor \xi_1 \rfloor, \lceil \xi_1 \rceil)) = 1.$$

Es folgt $\rho_{k+1} \leq \rho_k + 1$ für $k < \tau$ und damit induktiv

$$\rho_k \leq \rho_0 + k \quad \text{auf dem Ereignis } \{k < \tau\}. \quad (4.26)$$

Wir wollen nun die mittlere l_1 -Abweichung von $p(k)$ und $\ell(k)/n$ zum Zeitpunkt $k := \lfloor n^\beta \rfloor$ abschätzen. Da für $q, \tilde{q} \in \Omega_1$ stets $|q - \tilde{q}|_1 \leq 2$ gilt, ergibt sich:

$$\begin{aligned} & E_\mu^{(n)} \left[\left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 \right] \\ &= \int_{\Omega_{cde}^{(n)}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} \\ &= \int_{\{\tau > \lfloor n^\beta \rfloor\}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} + \int_{\{\tau \leq \lfloor n^\beta \rfloor\}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} \\ &\leq \int_{\{\tau > \lfloor n^\beta \rfloor\}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} + 2 \int \mathbf{1}_{\{\tau \leq \lfloor n^\beta \rfloor\}} dQ_\mu^{(n)} \\ &= \int_{\{\tau > \lfloor n^\beta \rfloor\}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} + 2 Q_\mu^{(n)}[\tau \leq \lfloor n^\beta \rfloor]. \end{aligned} \quad (4.27)$$

Wir zeigen, dass für $n \rightarrow \infty$ beide Terme verschwinden. Dies ist letztlich darauf zurückzuführen, dass mit größer werdendem n die Approximationsgüte verbessert wird, wodurch sowohl die Abweichung zwischen beiden Prozessen als auch die Wahrscheinlichkeit einer Entkopplung sinken.

Wir schätzen zunächst den ersten Term ab. Dazu benötigen wir das folgende Lemma:

4.2 Lemma. *Für zwei Folgen nichtnegativer reeller Zahlen $p = (p_i)_{i \geq 1}$ und $q = (q_i)_{i \geq 1}$ mit endlicher l_1 -Norm gilt*

$$|p - q|_1 \geq |\text{sort}(p) - \text{sort}(q)|_1. \quad (4.28)$$

Beweis. Wir führen sort zunächst für p aus; q ordnen wir in analoger Weise um. Der Wert von $|p - q|_1 = \sum_{i \geq 1} |p_i - q_i|$ ändert sich dadurch nicht, da alle Summanden positiv sind und die Reihe somit absolut summierbar ist. Der Einfachheit halber bezeichnen wir die so umsortierten Folgen wiederum mit p und q .

Nun sortieren wir q ; dazu suchen wir das größte Element, etwa q_n , und vertauschen es mit dem ersten Element der Folge, q_1 . Wir zeigen nun

$$|p_1 - q_1| + |p_n - q_n| \leq |p_1 - q_n| + |p_n - q_1|. \quad (4.29)$$

Da die Gleichung symmetrisch in p und q ist, genügt es, die folgenden Fälle zu betrachten (andernfalls vertausche $p_1 \leftrightarrow q_n$ und $p_n \leftrightarrow q_1$):

1. $p_1 \geq p_n \geq q_n \geq q_1$: Es folgt $|p_1 - q_1| = |p_1 - q_n| + |q_n - q_1|$ und $|p_n - q_n| = |p_n - q_1| - |q_n - q_1|$, also gilt Gleichheit in (4.29).
2. $p_1 \geq q_n \geq p_n \geq q_1$: Es folgt $|p_1 - q_1| = |p_1 - q_n| + |q_n - q_1|$ und $|p_n - q_n| \geq |p_n - q_1| - |q_n - q_1|$, Gleichung (4.29) ist erfüllt.
3. $p_1 \geq q_n \geq q_1 \geq p_n$: Es folgt $|p_1 - q_1| \geq |p_1 - q_n|$ und $|p_n - q_n| \geq |p_n - q_1|$.

Induktiv lässt sich so $|p - q|_1 \geq |\text{sort}(p) - \text{sort}(q)|_1$ zeigen, allerdings nur für Folgen q , die in endlich vielen Schritten sortiert werden können. Um das Ergebnis auf allgemeine Folgen q ausdehnen zu können, definieren wir zunächst für $N \in \mathbb{N}$ den Operator sort_N , der genau N Sortierschritte ausführt, und wählen N so groß, dass

$\sum_{i=N+1}^{\infty} |p_i| < \varepsilon$ und $\sum_{i=N+1}^{\infty} |\text{sort}(q)_i| < \varepsilon$ gilt für ein $\varepsilon > 0$. Da nach Annahme p und q summierbar sind, ist dies immer möglich. Somit haben wir

$$\sum_{i=N+1}^{\infty} |p_i - \text{sort}(q)_i| \leq \sum_{i=N+1}^{\infty} (|p_i| + |\text{sort}(q)_i|) < 2\varepsilon,$$

und insgesamt ergibt sich

$$\begin{aligned} |p - \text{sort}(q)|_1 &= \sum_{i=1}^N |p_i - \text{sort}(q)_i| + \sum_{i=N+1}^{\infty} |p_i - \text{sort}(q)_i| \\ &= \sum_{i=1}^N |p_i - \text{sort}_N(q)_i| + \sum_{i=N+1}^{\infty} |p_i - \text{sort}(q)_i| \\ &\leq |p - \text{sort}_N(q)|_1 + 2\varepsilon \\ &\leq |p - q|_1 + 2\varepsilon. \end{aligned}$$

Mit $\varepsilon \searrow 0$ folgt sodann die Behauptung. \square

Auf dem Ereignis $\{\lfloor n^\beta \rfloor < \tau\}$, d. h. in gekoppeltem Zustand, können wir die Abweichung zwischen den Prozessen durch $N_{\ell(0)} + \lfloor n^\beta \rfloor$ abschätzen (vgl. Gleichungen (4.25) und (4.26)). Da diese Werte noch durch den Faktor $1/n$ rückskaliert werden müssen, erhalten wir das gewünschte Ergebnis. Es gilt

$$\begin{aligned} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 &= \frac{1}{n} \left| \text{sort} \left((\mathbb{X}(c_{\lfloor n^\beta \rfloor}^{-1}(\{i\})))_{i \geq 1} \right) - \text{sort} \left((\#d_{(\lfloor n^\beta \rfloor)}^{-1}(\{i\}))_{i \geq 1} \right) \right|_1 \\ &\leq \frac{1}{n} \left| (\mathbb{X}(c_{\lfloor n^\beta \rfloor}^{-1}(\{i\})))_{i \geq 1} - (\#d_{(\lfloor n^\beta \rfloor)}^{-1}(\{i\}))_{i \geq 1} \right|_1 \\ &= \frac{1}{n} \sum_{i \geq 1} \left| \mathbb{X}(c_{\lfloor n^\beta \rfloor}^{-1}(\{i\})) - \#d_{(\lfloor n^\beta \rfloor)}^{-1}(\{i\}) \right| \\ &= \frac{1}{n} \sum_{i \geq 1} \left| \mathbb{X}(\{x : c_{\lfloor n^\beta \rfloor}(x) = i\}) - \mathbb{X}(\{x : d_{(\lfloor n^\beta \rfloor)}(\lceil x \rceil) = i\}) \right|. \\ &\leq \frac{1}{2n} \sum_{i \geq 1} \left(\mathbb{X}(\{x : c_{\lfloor n^\beta \rfloor}(x) = i, c_{\lfloor n^\beta \rfloor}(x) \neq d_{\lfloor n^\beta \rfloor}(\lceil x \rceil)\}) \right. \\ &\quad \left. + \mathbb{X}(\{x : d_{\lfloor n^\beta \rfloor}(\lceil x \rceil) = i, c_{\lfloor n^\beta \rfloor}(x) \neq d_{\lfloor n^\beta \rfloor}(\lceil x \rceil)\}) \right), \end{aligned}$$

wobei für die letzte Zeile $|\mathbb{X}(A) - \mathbb{X}(B)| \leq \frac{1}{2}(\mathbb{X}(A \cap B^c) + \mathbb{X}(B \cap A^c))$ verwendet wurde. Dies lässt sich weiter fortführen zu

$$\begin{aligned} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 &\leq \frac{1}{n} \mathbb{X}(\{x \in [0, n) : c_{\lfloor n^\beta \rfloor}(x) \neq d_{\lfloor n^\beta \rfloor}(x)\}) \\ &= \frac{1}{n} \rho_{\lfloor n^\beta \rfloor} \leq \frac{2}{n} (\rho_0 + \lfloor n^\beta \rfloor) \\ &\leq \frac{1}{n} (N_{\ell(0)} + \lfloor n^\beta \rfloor), \end{aligned}$$

und es folgt mit Gleichung (4.22)

$$\begin{aligned} \int_{\{\tau > \lfloor n^\beta \rfloor\}} \left| p(\lfloor n^\beta \rfloor) - \frac{\ell(\lfloor n^\beta \rfloor)}{n} \right|_1 dQ_\mu^{(n)} &\leq \frac{1}{n} E_\mu^{(n)} (N_{\ell(0)} + \lfloor n^\beta \rfloor) \\ &= O(n^{\alpha-1} + n^{\beta-1}) \rightarrow 0 \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Zu zeigen ist nun noch, dass auch der Term $Q_\mu^{(n)}[\tau \leq \lfloor n^\beta \rfloor]$ für $n \rightarrow \infty$ verschwindet. Dieser gibt die Wahrscheinlichkeit an, dass $(p(k))_{k \geq 0}$ und $(\ell(k))_{k \geq 0}$ zum Zeitpunkt $\lfloor n^\beta \rfloor$ bereits entkoppelt sind.

Nehme dazu o. B. d. A. an, dass $\alpha < \beta < 1/2$ gilt. Andernfalls wähle $\bar{\beta}$ mit $1/2 > \bar{\beta} > \alpha$ und führe die nun folgende Rechnung mit $\bar{\beta}$ durch (siehe weiter unten).

Wir überlegen uns für beliebiges $k \in \mathbb{N}$ die Wahrscheinlichkeit, dass ein Prozess, der im k -ten Schritt noch gekoppelt ist, im $k+1$ -ten Schritt entkoppelt: Gegeben ρ_0, \dots, ρ_k , müssen dazu entweder ξ_1 oder ξ_2 zwei verschiedene Abschnitte in c_k und d_k auswählen, was jeweils mit Wahrscheinlichkeit ρ_k/n geschieht, oder in d_k wird dasselbe Atom zweimal ausgewählt, also $\lceil \xi_1 \rceil = \lceil \xi_2 \rceil$, was mit Wahrscheinlichkeit $1/n$ passiert:

$$\begin{aligned} &Q_\mu^{(n)}[\tau = k+1 | \tau > k] \\ &= Q_\mu^{(n)}(\{c(\xi_1) \neq d(\lceil \xi_1 \rceil)\} \cup \{c(\xi_2) \neq d(\lceil \xi_2 \rceil)\} \cup \{\lceil \xi_1 \rceil = \lceil \xi_2 \rceil\}) \\ &\leq \frac{\rho_k}{n} + \frac{\rho_k}{n} + \frac{1}{n} \\ &= \frac{2\rho_k + 1}{n}. \end{aligned} \tag{4.30}$$

Auf dem Ereignis $\{\rho_0 < n^\beta\}$ lässt sich dies weiter abschätzen

$$\begin{aligned} \frac{2\rho_k + 1}{n} &\leq \frac{2(\rho_0 + k) + 1}{n} \\ &\leq \frac{2(n^\beta + k) + 1}{n} \\ &= 2n^{\beta-1} + \frac{2k + 1}{n}, \end{aligned}$$

und für $k := \lfloor n^\beta \rfloor$ gilt überdies

$$\begin{aligned} &\leq 2n^{\beta-1} + 2n^{\beta-1} + \frac{1}{n} \\ &\leq 2n^{\beta-1} + 2n^{\beta-1} + n^{\beta-1} \\ &= 5n^{\beta-1}. \end{aligned} \tag{4.31}$$

Damit erhalten wir induktiv

$$\begin{aligned} &Q_\mu^{(n)}[\tau > k, \rho_0 < n^\beta] \\ &= Q_\mu^{(n)}(\{\tau \neq k\} \cap \{\tau > k-1, \rho_0 < n^\beta\}) \\ &= Q_\mu^{(n)}[\tau \neq k | \tau > k-1, \rho_0 < n^\beta] \cdot Q_\mu^{(n)}[\tau > k-1, \rho_0 < n^\beta] \\ &= (1 - Q_\mu^{(n)}[\tau = k | \tau > k-1, \rho_0 < n^\beta]) Q_\mu^{(n)}[\tau > k-1, \rho_0 < n^\beta] \\ &\geq (1 - 5n^{\beta-1}) Q_\mu^{(n)}[\tau > k-1, \rho_0 < n^\beta] \\ &\vdots \\ &\geq (1 - 5n^{\beta-1})^k Q_\mu^{(n)}[\tau > 0, \rho_0 < n^\beta] \\ &= (1 - 5n^{\beta-1})^k Q_\mu^{(n)}[\rho_0 < n^\beta] \end{aligned} \tag{4.32}$$

mit $k = \lfloor n^\beta \rfloor$ und somit insbesondere

$$\begin{aligned} Q_\mu^{(n)}[\tau > \lfloor n^\beta \rfloor] &\geq Q_\mu^{(n)}[\tau > \lfloor n^\beta \rfloor, \rho_0 < n^\beta] \\ &\geq (1 - 5n^{\beta-1})^{n^\beta} Q_\mu^{(n)}[\rho_0 < n^\beta] \\ &= ((1 - 5n^{\beta-1})^{n^{1-\beta}})^{n^{2\beta-1}} Q_\mu^{(n)}[\rho_0 < n^\beta]. \end{aligned} \tag{4.33}$$

Aus der Analysis ist bekannt $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x$. Da wegen $1 - \beta > 0$ die Folge

$(\lfloor n^{1-\beta} \rfloor)_{n \geq 0}$ eine Teilfolge von $(n)_{n \geq 0}$ ist, gilt somit

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{\lfloor n^{1-\beta} \rfloor} \right)^{\lfloor n^{1-\beta} \rfloor} = e^x$$

und aufgrund der Monotonie von $(1 + \frac{x}{n})^n$ auch

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n^{1-\beta}} \right)^{n^{1-\beta}} = e^x.$$

Sei $\delta > 0$. Wähle $N \in \mathbb{N}$ hinreichend groß, so dass

$$\left(1 - \frac{5}{n^{1-\beta}} \right)^{n^{1-\beta}} \geq \delta \quad \text{für alle } n \geq N$$

gilt. Wegen $2\beta - 1 < 0$ folgt $n^{2\beta-1} \searrow 0$ für $n \rightarrow \infty$ und wir erhalten

$$\left((1 - 5n^{\beta-1})^{n^{1-\beta}} \right)^{n^{2\beta-1}} \rightarrow 1 \quad \text{für } n \rightarrow \infty.$$

Der zweite Faktor in Gleichung (4.33) konvergiert für $\beta > \alpha$ wegen Gleichung (4.23) und (4.25) ebenfalls gegen 1. Es gilt also

$$Q_\mu^{(n)}[\tau > \lfloor n^\beta \rfloor] \rightarrow 1 \quad \text{für } n \rightarrow \infty.$$

Für $\beta < \alpha$ haben wir $\bar{\beta} > \beta$ mit $\bar{\beta} > \alpha$ und erhalten dann wegen $n^{\bar{\beta}} > n^\beta$

$$Q_\mu^{(n)}[\tau > \lfloor n^\beta \rfloor] \geq Q_\mu^{(n)}[\tau > \lfloor n^{\bar{\beta}} \rfloor] \rightarrow 1 \quad \text{für } n \rightarrow \infty.$$

Damit ist gezeigt, dass auch der zweite Term in Gleichung (4.27) verschwindet, und der Beweis ist vollständig. \square

5 Gruppendarstellungen und Charaktere

Da der Beweis der schnellen $\text{DCF}^{(n)}$ -Konvergenz Resultate aus der Darstellungstheorie der symmetrischen Gruppe verwendet, geben wir im Folgenden eine kurze Einführung in das Thema und stellen die in Kapitel 6 benötigten Ergebnisse bereit.

5.1 Darstellungen endlicher Gruppen

Die folgenden Ausführungen sind im Wesentlichen dem Buch von Diaconis [6], Kap. 2 entnommen.

Eine Gruppendarstellung ordnet Elementen einer Gruppe lineare Abbildungen zu. So ist es möglich, gruppentheoretische Probleme auf Probleme der linearen Algebra zu reduzieren.

5.1 Definition. Sei G eine endliche Gruppe. Eine *Darstellung* (engl. *representation*) (ρ, V) von G ist ein Homomorphismus ρ von G in die Gruppe der invertierbaren linearen Abbildungen eines endlichdimensionalen reellen oder komplexen Vektorraumes V

$$\rho: G \longrightarrow GL(V) \quad \text{mit} \quad \rho(st) = \rho(s)\rho(t) \quad \text{für alle } s, t \in G. \quad (5.1)$$

Der Vektorraum V wird als *Darstellungsraum* oder auch als *G -Modul* bezeichnet, und die Dimension von V wird auch *Dimension von ρ* genannt und mit d_ρ bezeichnet. Gelegentlich schreiben wir, falls keine Verwechslungen zu befürchten sind, für eine Darstellung (ρ, V) lediglich ρ .

Direkt aus der Definition folgt (*id* bezeichne das neutrale Element von G , I die Einheitsmatrix)

$$\rho(id) = I, \quad \rho(s^{-1}) = \rho(s)^{-1} \quad \text{für alle } s \in G. \quad (5.2)$$

Ein Untervektorraum $W \subset V$ heißt *stabil* unter G , wenn gilt

$$\rho(s)W \subset W \quad \text{für alle } s \in G. \quad (5.3)$$

Ist (ρ, V) eine Darstellung und W ein stabiler Unterraum von V , so heißt $(\rho|_W, W)$ *Subdarstellung* von (ρ, V) . Die Einschränkungen von ρ auf $W = \{0\}$ und auf $W = V$ sind die trivialen Subdarstellungen. Lässt eine Darstellung keine nichttrivialen Subdarstellungen zu, so heißt sie *irreduzibel*. Die auf einem beliebigen eindimensionalen Vektorraum V definierte *Einheitsdarstellung* $\rho(s) = I$ für alle $s \in G$ ist irreduzibel und wird mit 1 bezeichnet.

Um die Aktion von G auf V zu betrachten, genügt es, V in eine geeignete direkte Summe von stabilen Untervektorräumen zu zerlegen und die Aktion auf den einzelnen Summanden zu betrachten. Dabei heißt V die *direkte Summe* der Untervektorräume $(W_i)_{i=1, \dots, n}$, in Zeichen

$$V = \bigoplus_{i=1}^n W_i,$$

wenn gilt $V = \sum_i W_i$ und $W_j \cap \sum_{i \neq j} W_i = \{0\}$ für alle j . Im Spezialfall $V = W_1 \oplus W_2$ heißen W_1 und W_2 *komplementär*.

Wir konstruieren zunächst ein neues Skalarprodukt: Sei (ρ, V) eine beliebige Darstellung und sei $\langle \cdot, \cdot \rangle_1$ das Skalarprodukt von V . Wir setzen

$$\langle u, v \rangle := \sum_{s \in G} \langle \rho(s)u, \rho(s)v \rangle_1 \quad \text{für } u, v \in V \quad (5.4)$$

und erhalten so offenbar wiederum ein Skalarprodukt auf V . Dieses Skalarprodukt ist invariant unter $\rho(t)$ für alle $t \in G$, da gilt

$$\langle \rho(t)u, \rho(t)v \rangle = \sum_{s \in G} \langle \rho(st)u, \rho(st)v \rangle_1 = \langle u, v \rangle. \quad (5.5)$$

Nun lässt sich zeigen:

5.2 Satz. *Sei $\rho: G \longrightarrow GL(V)$ eine lineare Darstellung von G in V und sei W ein Untervektorraum von V , der stabil ist unter G . Dann gibt es ein Komplement W^0 von W , das ebenfalls stabil ist unter G .*

Beweis. Wir wählen für W^0 das orthogonale Komplement W^\perp von W bezüglich $\langle \cdot, \cdot \rangle$: $W^0 := \{v \in V \mid \langle v, w \rangle = 0 \quad \forall w \in W\}$. Dann gilt $V = W \oplus W^0$, und wir haben für beliebige $v \in W^0, s \in G$

$$\langle \rho(s)v, w \rangle = \langle v, \rho(s^{-1})w \rangle = 0 \quad \text{für alle } w \in W$$

wegen Gleichung (5.5), und somit $\rho(s)W^0 \subset W^0$. □

Wir sagen dann auch: Die Darstellung (ρ, V) ist die direkte Summe der Darstellungen $(\rho|_W, W)$ und $(\rho|_{W^0}, W^0)$.

Allgemein definieren wir für zwei Darstellungen (ρ, V) und (η, W) derselben Gruppe G mit $V \cap W = \{0\}$ die *direkte Summe*

$$(\rho, V) \oplus (\eta, W) := (\rho \oplus \eta, V \oplus W) \quad (5.6)$$

von (ρ, V) und (η, W) vermöge

$$\rho \oplus \eta(s)(v, w) := (\rho(s)v, \eta(s)w) \quad \text{für } s \in G, v \in V, w \in W. \quad (5.7)$$

Dies ist wiederum eine Darstellung von G . Die Dimension der Darstellung beträgt $d_\rho + d_\eta$, und (ρ, V) und (η, W) sind als Subdarstellungen enthalten.

In der Situation von Satz 5.2 können wir die Zerlegung verfeinern, indem wir W und W^0 weiter aufteilen. Dies lässt sich induktiv so lange fortsetzen, bis V in eine direkte Summe zerfällt, in der die Einschränkung von ρ auf jeden einzelnen Summanden eine irreduzible Darstellung liefert. Somit erhalten wir das oben angekündigte Resultat:

5.3 Satz. *Jede Darstellung ist eine direkte Summe irreduzibler Darstellungen.*

Die Invarianz des inneren Produkts $\langle \cdot, \cdot \rangle$ liefert noch eine weitere wichtige Eigenschaft: Für eine Orthonormalbasis $(e_i)_{i \in I}$ von V bezüglich $\langle \cdot, \cdot \rangle$ und beliebiges $s \in G$ gilt für die Matrix $\rho(s) = (r_{ij}(s))$

$$\begin{aligned} (\overline{\rho(s)}^\top \rho(s))_{ij} &= \sum_k \overline{r_{ik}(s)}^\top r_{kj}(s) \\ &= \sum_{k,l} \overline{r_{ki}(s)} r_{lj}(s) \langle e_k, e_l \rangle \end{aligned}$$

$$\begin{aligned}
&= \left\langle \sum_k r_{ki}(s) e_k, \sum_l r_{lj}(s) e_l \right\rangle \\
&= \langle \rho(s) e_i, \rho(s) e_j \rangle = \delta_{ij},
\end{aligned}$$

also $\rho(s)^{-1} = \bar{\rho}^\top$. Eine Matrix bzw. die zugehörige lineare Abbildung mit dieser Eigenschaft heißen bekanntlich *unitär*. Wir können somit ohne Einschränkung immer davon ausgehen, dass eine Darstellung unitär ist.

Betrachte nun zwei Darstellungen $\rho: G \longrightarrow GL(V)$ und $\tau: G \longrightarrow GL(W)$. ρ und τ heißen *äquivalent*, falls es eine bijektive lineare Abbildung $f: V \longrightarrow W$ gibt, so dass

$$\tau(s) = f \circ \rho(s) \circ f^{-1} \quad \text{für alle } s \in G \quad (5.8)$$

gilt. Eine wichtige Aussage macht das folgende Lemma:

5.4 Lemma (Schur). *Seien $\rho^1: G \longrightarrow GL(V_1)$ und $\rho^2: G \longrightarrow GL(V_2)$ zwei irreduzible Darstellungen von G und sei $f: V_1 \longrightarrow V_2$ eine lineare Abbildung dergestalt, dass gilt*

$$\rho^2(s) \circ f = f \circ \rho^1(s) \quad \text{für alle } s \in G. \quad (5.9)$$

Dann gilt:

- (1) *Sind ρ^1 und ρ^2 nicht äquivalent, so ist $f = 0$.*
- (2) *Ist $V_1 = V_2$ und $\rho^1 = \rho^2$, so ist $f = \lambda I$ für ein $\lambda \in \mathbb{C}$.*

Beweis. Wir schreiben der Übersichtlichkeit halber ρ_s für $\rho(s)$. Es zeigt sich, dass sowohl der Kern von f in V_1 als auch das Bild $f(V_1) \subset V_2$ unter ρ_s^1 bzw. ρ_s^2 invariante Unterräume sind (für alle $s \in G$): Aus $f(v) = 0$ folgt schon $f \circ \rho_s^1(v) = \rho_s^2 \circ f(v) = 0$, $\rho_s^1(v)$ liegt also ebenfalls im Kern. Für $w = f(v)$ folgt, dass $\rho_s^2(w) = f \circ \rho_s^1(v)$ ebenfalls schon im Bild liegt. Aus der jeweiligen Irreduzibilität lässt sich nun folgern, dass sowohl der Kern als auch das Bild entweder trivial oder schon der ganze Raum sind.

Um (1) zu zeigen, nehmen wir $f \neq 0$ an. Dann sind $\text{Ker}(f) = 0$ und $\text{Im}(f) = V_2$, f ist also ein Isomorphismus und ρ^1 und ρ^2 sind äquivalent. Um (2) zu zeigen, nehmen wir wiederum $f \neq 0$ an (für $f = 0$ ist die Aussage trivialerweise erfüllt). Dann hat f einen nichtverschwindenden Eigenwert λ . Wegen $\rho_s^2 \circ \lambda I = \lambda I \circ \rho_s^1$ gilt für die Abbildung $\tilde{f} := f - \lambda I$ ebenfalls $\rho_s^2 \circ \tilde{f} = \tilde{f} \circ \rho_s^1$ für alle $s \in G$. Da \tilde{f}

nach Konstruktion einen nichttrivialen Kern hat, muss schon $\tilde{f} \equiv 0$ gelten, und wir erhalten $f = \lambda I$. \square

Mithilfe des Lemmas von Schur lässt sich zeigen, dass die Matrixeinträge der irreduziblen Darstellungen einer Gruppe G eine orthogonale Basis für Funktionen $G \rightarrow \mathbb{C}$ bilden.

Wir zeigen zunächst, dass die Matrixeinträge ein Orthogonalsystem bilden. Dazu formulieren wir als Erstes das folgende Korollar:

5.5 Korollar. *Seien ρ^1 und ρ^2 wie im Lemma 5.4 und sei h eine lineare Abbildung von V_1 nach V_2 . Setze*

$$h^0 := \frac{1}{|G|} \sum_{t \in G} (\rho^2(t))^{-1} \circ h \circ \rho^1(t). \quad (5.10)$$

Dann gilt:

(1) Sind ρ^1 und ρ^2 nicht äquivalent, so gilt $h^0 = 0$.

(2) Ist $V_1 = V_2$ und $\rho^1 = \rho^2$, so ist $h^0 = \lambda I$ mit $\lambda = \frac{\text{Spur}(h)}{d_{\rho^1}} = \frac{\text{Spur}(h)}{d_{\rho^2}}$.

Beweis. Für jedes $s \in G$ gilt

$$\begin{aligned} \rho^2(s^{-1})h^0\rho^1(s) &= \frac{1}{|G|} \sum_{t \in G} \rho^2(s^{-1})\rho^2(t^{-1})h\rho^1(t)\rho^1(s) \\ &= \frac{1}{|G|} \sum_{t \in G} \rho^2(s^{-1}t^{-1})h\rho^1(ts) \\ &= \frac{1}{|G|} \sum_{t \in G} (\rho^2(ts))^{-1}h\rho^1(ts) = h^0, \end{aligned}$$

also

$$h^0 \circ \rho^1(s) = \rho^2(s) \circ h^0. \quad (5.11)$$

Sind ρ^1 und ρ^2 nicht äquivalent, so ist nach Schurs Lemma $h^0 = 0$, was (1) zeigt. Ist $V_1 = V_2$ und $\rho_1 = \rho_2 =: \rho$, so gilt nach dem zweiten Teil von Schurs Lemma $h^0 = \lambda I$. Bilden wir bei obiger Gleichung (5.11) auf beiden Seiten die Spur, so ergibt sich

$$\lambda d_\rho = \lambda \text{Spur}(I) = \text{Spur}(h^0) = \frac{1}{|G|} \sum_{t \in G} \text{Spur}(\rho(t)^{-1} \circ h \circ \rho(t)) = \text{Spur}(h). \quad \square$$

Seien nun ρ^1, ρ^2, h und h^0 in Matrixdarstellung gegeben (bezüglich geeigneter Basen von V_1 und V_2):

$$\begin{aligned}\rho^1(t) &= (r_{ij}^1(t)), & \rho^2(t) &= (r_{ij}^2(t)) \quad \text{für } t \in G, \\ h &= (x_{ij}), & h^0 &= (x_{ij}^0).\end{aligned}$$

Damit lässt sich Gleichung (5.10) schreiben als

$$\begin{aligned}x_{i_2 i_1}^0 &= \frac{1}{|G|} \sum_{t \in G} \sum_{j_1, j_2} r_{i_2 j_2}^2(t^{-1}) x_{j_2 j_1} r_{j_1 i_1}^1(t) \\ &= \sum_{j_1, j_2} \left(\frac{1}{|G|} \sum_{t \in G} r_{i_2 j_2}^2(t^{-1}) r_{j_1 i_1}^1(t) \right) x_{j_2 j_1}.\end{aligned}\tag{5.12}$$

Im Falle (1) gilt $h^0 \equiv 0$ für jedes h . Dies kann aber nur der Fall sein, wenn die Koeffizienten von $x_{j_2 j_1}$ stets verschwinden. Wir erhalten:

5.6 Korollar. Seien $\rho^1 = (r_{ij}^1)$ und $\rho^2 = (r_{ij}^2)$ wie in Lemma 5.4.

(1) Sind ρ^1 und ρ^2 nicht äquivalent, so gilt

$$\frac{1}{|G|} \sum_{t \in G} r_{i_2 j_2}^2(t^{-1}) r_{j_1 i_1}^1(t) = 0 \quad \text{für alle } i_1, i_2, j_1, j_2.\tag{5.13}$$

(2) Ist $V_1 = V_2$ und $\rho^1 = \rho^2 = (r_{ij})$, so gilt

$$\frac{1}{|G|} \sum_{t \in G} r_{i_2 j_2}(t^{-1}) r_{j_1 i_1}(t) = \begin{cases} \frac{1}{d_\rho} & \text{für } i_1 = i_2 \text{ und } j_1 = j_2, \\ 0 & \text{sonst.} \end{cases}\tag{5.14}$$

Beweis. Für den Fall (1) ist bereits alles gezeigt. Im Fall (2) gilt für jede Wahl von h nach Korollar 5.5 $h^0 = \lambda I$, also $x_{i_2 i_1}^0 = \lambda \delta_{i_2 i_1}$, mit $\lambda = \frac{1}{d_\rho} \text{Spur}(h) = \frac{1}{d_\rho} \sum_{j_1, j_2} \delta_{j_2 j_1} x_{j_2 j_1}$. Setzen wir für $x_{i_2 i_1}^0$ den Ausdruck (5.12) ein, so erhalten wir

$$\sum_{j_1, j_2} \left(\frac{1}{|G|} \sum_{t \in G} r_{i_2 j_2}^2(t^{-1}) r_{j_1 i_1}^1(t) \right) x_{j_2 j_1} = \sum_{j_1, j_2} \frac{1}{d_\rho} \delta_{i_2 i_1} \delta_{j_2 j_1} x_{j_2 j_1}.$$

Da h beliebig war, müssen die Koeffizienten von $x_{j_2 j_1}$ auf beiden Seiten übereinstim-

men, und wir erhalten das angegebene Ergebnis. \square

Die Gleichungen (5.13) und (5.14) nehmen eine besonders übersichtliche Form an, wenn die beteiligten Darstellungen unitär sind, was wir ohne Einschränkung annehmen können (vgl. S. 62). Dann gilt $r_{ij}(s^{-1}) = r_{ij}(s)^{-1} = \overline{r_{ji}(s)}$ für alle $s \in G$. Da es sich bei den Matrixeinträgen um Funktionen von G nach \mathbb{C} handelt, führen wir nun das übliche innere Produkt auf dem Vektorraum derselben ein. Für $\phi, \psi: G \rightarrow \mathbb{C}$ setzen wir

$$(\phi|\psi) := \frac{1}{|G|} \sum_{t \in G} \phi(t) \overline{\psi(t)}. \quad (5.15)$$

Damit lässt sich das vorangehende Korollar wie folgt neu formulieren:

5.7 Korollar. *Seien $\rho^1 = (r_{ij}^1)$ und $\rho^2 = (r_{ij}^2)$ wie in Lemma 5.4, wobei zusätzlich angenommen sei, dass die Darstellungen unitär sind.*

(1) *Sind ρ^1 und ρ^2 nicht äquivalent, so gilt*

$$(r_{i_1 j_1}^1 | r_{i_2 j_2}^2) = 0 \quad \text{für alle } i_1, i_2, j_1, j_2. \quad (5.16)$$

(2) *Ist $V_1 = V_2$ und $\rho^1 = \rho^2 = (r_{ij})$, so gilt*

$$(r_{i_1 j_1} | r_{i_2 j_2}) = \frac{1}{d_\rho} \delta_{i_1 i_2} \delta_{j_1 j_2}. \quad (5.17)$$

Damit ist insbesondere gezeigt, dass die Matrixeinträge der unitären irreduziblen Darstellungen einer Gruppe G ein Orthogonalsystem bilden. Am Ende des nächsten Abschnitts werden wir zeigen, dass sie sogar eine Orthogonalbasis für Funktionen von G nach \mathbb{C} bilden.

5.2 Charakter einer Darstellung

Als Nächstes wird der Begriff des Charakters eingeführt, der in unserem Beweis der DCF⁽ⁿ⁾-Konvergenz entscheidende Verwendung findet (vgl. [6], Kap. 2).

5.8 Definition. Sei (ρ, V) die Darstellung einer endlichen Gruppe G . Dann heißt

die Abbildung

$$\chi_\rho: G \longrightarrow \mathbb{C}, \quad \chi_\rho(s) := \text{Spur } \rho(s) = \sum_i r_{ii}(s), \quad (5.18)$$

die jedem $s \in G$ die Spur der Matrix $\rho(s) = (r_{ij}(s))$ zuordnet, *Charakter* χ_ρ der *Darstellung* (ρ, V) .

Ein Charakter ist unabhängig von der gewählten Basis. Dies folgt aus der Basisinvarianz der Spur, denn für zwei Matrizen $A = (a_{ij}), B = (b_{kl})$ gilt (vgl. [9], S. 221)

$$\text{Spur}(A \cdot B) = \sum_{i,j} a_{ij} b_{ji} = \sum_{l,k} b_{lk} a_{kl} = \text{Spur}(B \cdot A).$$

und für ein invertierbares B (die Matrix einer Basistransformation) folgt somit

$$\text{Spur}(B^{-1} \cdot A \cdot B) = \text{Spur}(A).$$

Dies bedeutet weiter, dass äquivalente Darstellungen denselben Charakter besitzen: Sind ρ und τ zwei äquivalente Darstellungen, so gibt es eine bijektive lineare Abbildung f mit $\rho = f \circ \tau \circ f^{-1}$, und wir erhalten für die zugehörigen Charaktere

$$\chi_\rho = \text{Spur } \rho = \text{Spur}(f \circ \tau \circ f^{-1}) = \text{Spur } \tau = \chi_\tau. \quad (5.19)$$

Der nächste Satz beschreibt grundlegende Eigenschaften der Charaktere:

5.9 Satz. *Sei χ der Charakter einer Darstellung ρ mit Dimension d_ρ . Dann gilt für $s, t \in G$:*

$$(1) \quad \chi(id) = d_\rho,$$

$$(2) \quad \chi(s^{-1}) = \overline{\chi(s)},$$

$$(3) \quad \chi(tst^{-1}) = \chi(s).$$

Beweis. (1) $\rho(id) = I$ und $\text{Spur}(I) = d_\rho$. (2) Wähle ein $n \in \mathbb{N}$ mit $s^n = id$. Somit gilt $\rho(s^n) = I$, und es folgt, dass die Eigenwerte λ_i von $\rho(s)$ allesamt Einheitswurzeln

sind. Somit gilt $1 = |\lambda_i|^2 = \lambda_i \overline{\lambda_i}$ für alle i , und wir erhalten

$$\overline{\chi(s)} = \text{Spur } \overline{\rho(s)} = \sum_i \overline{\lambda_i} = \sum_i \frac{1}{\lambda_i} = \text{Spur } \rho(s)^{-1} = \chi(s^{-1}).$$

$$(3) \chi(tst^{-1}) = \text{Spur } \rho(tst^{-1}) = \text{Spur } (\rho(t)\rho(s)\rho(t)^{-1}) = \text{Spur } \rho(s) = \chi(s). \quad \square$$

Weiter ist festzustellen, dass sich aus Schurs Lemma (in der Form von Korollar 5.7) eine Orthonormalitätsrelation für die Charaktere gewinnen lässt:

5.10 Satz. *Die Charaktere der irreduziblen Darstellungen einer Gruppe G (modulo Äquivalenz) sind orthonormal bezüglich (\cdot, \cdot) als Funktionen von G nach \mathbb{C} .*

Beweis. Seien ρ, ρ' unitäre irreduzible Darstellungen mit Charakteren χ, χ' . Sind die Darstellungen in Matrixform $\rho(t) = (r_{ij}(t))$ bzw. $\rho'(t) = (r'_{ij}(t))$ gegeben, so haben wir

$$(\chi|\chi') = \frac{1}{|G|} \sum_{t \in G} \sum_{i,j} r_{ii}(t) \overline{r'_{jj}(t)} = \sum_{i,j} \frac{1}{|G|} \sum_{t \in G} r_{ii}(t) \overline{r'_{jj}(t)} = \sum_{i,j} (r_{ii}|r'_{jj}).$$

Ist $\rho' = \rho$ (und damit auch $\chi' = \chi$), so gilt $r'_{ij} = r_{ij}$, und es folgt mit Korollar 5.7

$$(\chi|\chi) = \sum_{i,j} (r_{ii}|r_{jj}) = \sum_{i,j} \frac{1}{d_\rho} \delta_{ij} = 1.$$

Für ρ, ρ' äquivalent erhalten wir wegen $\chi \equiv \chi'$ dasselbe Ergebnis (vgl. Gleichung (5.19)). Sind hingegen ρ und ρ' nicht äquivalent, so gilt nach Korollar 5.7 $(r_{ii}|r'_{jj}) = 0$ für alle i, j und somit insbesondere $(\chi|\chi') = 0$. \square

Aus der Orthogonalitätsrelation erhalten wir:

5.11 Satz. *Sei (ρ, V) eine Darstellung von G , die sich nach Satz 5.3 in eine direkte Summe irreduzibler Subdarstellungen (ρ_i, W_i) zerlegen lässt, d. h.*

$$(\rho, V) = \bigoplus_{i=1}^k (\rho_i, W_i). \quad (5.20)$$

Ist dann (τ, W) eine weitere irreduzible Darstellung von G , so gilt für die Charaktere ϕ und χ von (ρ, V) bzw. (τ, W) : Die Anzahl der Subdarstellungen (ρ_i, W_i) , die zu (τ, W) äquivalent sind, ist gerade $(\phi|\chi)$.

Beweis. Für zwei Darstellungen ρ_1 und ρ_2 gilt: Die Matrix von $\rho_1 \oplus \rho_2$ ist bei geeigneter Wahl der Basen gerade $\begin{pmatrix} \rho_1 & 0 \\ 0 & \rho_2 \end{pmatrix}$, und wir erhalten für den zugehörigen Charakter

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2}.$$

Sei nun χ_i der Charakter von ρ_i für $i = 1, \dots, k$. Dann gilt $\phi = \chi_1 + \dots + \chi_k$, und wir erhalten

$$(\phi|\chi) = (\chi_1|\chi) + \dots + (\chi_k|\chi). \quad (5.21)$$

Der Wert eines einzelnen Summanden $(\chi_i|\chi)$ ist entweder 1 oder 0, je nachdem, ob (ρ_i, W_i) und (τ, W) äquivalente Darstellungen sind oder nicht. \square

Insbesondere heißt dies, dass die Anzahl der W_i isomorph zu W nicht von der Zerlegung (d. h. von der gewählten Basis) abhängt, da χ und ϕ und somit auch $(\phi|\chi)$ nicht von den gewählten Basen abhängen.

Weiter erhalten wir folgendes wichtige Resultat hinsichtlich der Äquivalenz von Darstellungen:

5.12 Korollar. *Zwei Darstellungen ρ_1, ρ_2 sind genau dann äquivalent, wenn sie denselben Charakter ϕ besitzen.*

Beweis. „ \Leftarrow “ Für jede irreduzible Darstellung (τ, W) mit Charakter χ gilt: Die Anzahl der zu (τ, W) äquivalenten irreduziblen Subdarstellungen von ρ_1 bzw. von ρ_2 ist gleich $(\phi|\chi)$. Somit sind bei einer Zerlegung von ρ_1 bzw. ρ_2 in eine direkte Summe irreduzibler Subdarstellungen alle Summanden äquivalent, und damit sind ρ_1 und ρ_2 selbst schon äquivalent.

„ \Rightarrow “ vgl. Gleichung (5.19). \square

Gemäß des obigen Satzes können wir die Zerlegung einer Darstellung (ρ, V) in irreduzible Subdarstellungen nach Satz 5.3 modulo Äquivalenz immer schreiben als

$$(\rho, V) = \bigoplus_{i=1}^n m_i(\rho_i, W_i), \quad (5.22)$$

falls (ρ, V) dieselbe (äquivalente) irreduzible Subdarstellung (ρ_i, W_i) jeweils m_i -mal enthält. Ist ϕ der Charakter von ρ und χ_i derjenige von ρ_i , so gilt wegen $\phi = \sum_i m_i \chi_i$

$$(\phi|\phi) = \sum_{i,j=1}^n m_i m_j (\chi_i|\chi_j) = \sum_{i=1}^n m_i^2. \quad (5.23)$$

Ist $(\phi|\phi) = 1$, so ist ρ selbst schon irreduzibel:

5.13 Satz. *Ist ϕ der Charakter einer Darstellung, so ist die Darstellung genau dann irreduzibel, wenn $(\phi|\phi) = 1$ gilt.*

Mithilfe der Charaktere können wir nun – wie bereits angekündigt – zeigen, dass die Matrixeinträge der unitären irreduziblen Darstellungen einer Gruppe G eine Orthonormalbasis für Funktionen $G \rightarrow \mathbb{C}$ bilden. Dazu müssen wir noch nachweisen, dass es gerade $|G|$ solcher Matrixeinträge gibt (vgl. S. 65).

Wir definieren zunächst die *reguläre Darstellung* (ρ_G, V_G) einer Gruppe G . Als Darstellungsraum wählen wir den von einer Orthonormalbasis $(e_s)_{s \in G}$ aufgespannten Vektorraum V_G und setzen

$$\rho_G(s)e_t := e_{sot} \quad \text{für } s, t \in G. \quad (5.24)$$

Beachte, dass sich V_G mit der Menge aller Funktionen $G \rightarrow \mathbb{C}$ identifizieren lässt.

Nach Konstruktion ist die reguläre Darstellung „maximal“ in dem Sinne, dass keine Informationen über G verloren gehen. Somit sind alle irreduziblen Darstellungen bereits in der regulären Darstellung enthalten (modulo Äquivalenz). Dies ergibt sich aus dem folgenden Satz:

5.14 Satz. *Der Charakter χ_G der regulären Darstellung einer Gruppe G ist gegeben durch*

$$\chi_G(s) = \begin{cases} |G| & \text{für } s = id, \\ 0 & \text{für alle } s \neq id. \end{cases} \quad (5.25)$$

Beweis. Es gilt $\rho_G(id) = I$, also $\text{Spur}(\rho_G(id)) = |G|$. Für alle $s \neq id$ haben wir $\rho(s)e_t = e_{sot} \neq e_t$, somit verschwinden alle Diagonaleinträge von $\rho(s)$ und es folgt $\text{Spur}(\rho_G(s)) = 0$. \square

5.15 Korollar. *Jede irreduzible Darstellung (ρ, W) einer Gruppe G ist äquivalent zu mindestens einer Subdarstellung der regulären Darstellung. Die Anzahl der zu (ρ, W) äquivalenten Subdarstellungen von (ρ_G, V_G) beträgt gerade d_ρ .*

Beweis. Sei χ der Charakter von ρ . Dann gilt (vgl. Satz 5.9)

$$(\chi_G | \chi) = \frac{1}{|G|} \sum_{s \in G} \chi_G(s) \overline{\chi(s)} = \overline{\chi(id)} = d_\rho.$$

Satz 5.11 liefert sodann die Behauptung. \square

Daraus folgt insbesondere, dass es nur endlich viele irreduzible Darstellungen einer Gruppe gibt. Das folgende Korollar macht nun die entscheidende Aussage über deren Grade:

5.16 Korollar. *Seien (ρ_i, W_i) , $i = 1, \dots, k$, die irreduziblen Darstellungen einer Gruppe G (modulo Äquivalenz). Dann gilt:*

- (i) *Die Grade d_{ρ_i} erfüllen $\sum_{i=1}^k d_{\rho_i}^2 = |G|$.*
- (ii) *Für $s \in G$ mit $s \neq id$ gilt $\sum_{i=1}^k d_{\rho_i} \chi_i(s) = 0$.*

Beweis. Gemäß Korollar 5.15 gilt $\chi_G(s) = \sum_i d_{\rho_i} \chi_i(s)$; beachte dazu auch die Gleichungen (5.22) und (5.21). Um (i) zu zeigen, wähle $s = 1$. Für (ii) wähle ein beliebiges $s \neq id$. \square

Da eine Darstellung immer so gewählt werden kann, dass sie unitär ist (vgl. S. 62), erhalten wir zusammen mit der entsprechenden Fassung des Lemmas von Schur schließlich das angekündigte Ergebnis:

5.17 Korollar. *Die Matrixeinträge der unitären irreduziblen Darstellungen einer Gruppe G bilden eine Orthogonalbasis für den Vektorraum der Funktionen $G \rightarrow \mathbb{C}$.*

Beweis. Aus Korollar 5.7 folgt, dass die Matrixeinträge der unitären irreduziblen Darstellungen ein Orthogonalsystem bilden. Dieses System ist hinreichend groß, um den betreffenden Vektorraum auch aufzuspinnen, denn nach Korollar 5.16 gibt es $\sum_i d_{\rho_i}^2 = |G|$ solcher Matrixeinträge. Dies stimmt gerade mit der Dimension des Vektorraumes der Funktionen $G \rightarrow \mathbb{C}$ überein. \square

5.3 Anzahl der irreduziblen Darstellungen einer Gruppe

Nach dieser kurzen Einführung in die Grundzüge der Darstellungstheorie endlicher Gruppen wollen wir nun die Anzahl der irreduziblen Darstellungen einer Gruppe genauer bestimmen. Wir werden zeigen, dass diese im Falle der symmetrischen Gruppe S_n gleich der Anzahl der Partitionen von $\{1, \dots, n\}$ ist.

Zunächst führen wir eine Äquivalenzrelation auf der Gruppe G ein: Zwei Elemente $s, t \in G$ heißen *konjugiert*, wenn es ein $u \in G$ gibt, so dass gilt

$$u \circ s \circ u^{-1} = t. \quad (5.26)$$

Hierdurch erhalten wir eine Zerlegung von G in (disjunkte) Äquivalenzklassen. Ist G abelsch, so umfassen alle Äquivalenzklassen bloß ein einziges Element. Im allgemeinen Fall können die Äquivalenzklassen durchaus größer sein, so zum Beispiel bei der S_n . Bevor wir uns der Frage nach der Anzahl der irreduziblen Darstellungen der S_n widmen, geben wir zunächst einige allgemeine Eigenschaften dieser Gruppe an (vgl. [5], S. 251 ff.).

Ein Element $\pi \in S_n$ wird häufig in der Form

$$\begin{pmatrix} 1 & \dots & n \\ \pi(1) & \dots & \pi(n) \end{pmatrix}$$

geschrieben. Eine Permutation $\pi \in S_n$ heißt ein *Zyklus*, wenn es paarweise verschiedene Zahlen $x_1, \dots, x_r \in \{1, \dots, n\}$, $r \geq 2$, gibt mit

$$\begin{aligned} \pi(x_i) &= x_{i+1} \quad \text{für } 1 \leq i < r, \\ \pi(x_r) &= x_1, \\ \pi(x) &= x \quad \text{für } x \in \{1, \dots, n\} \setminus \{x_1, \dots, x_r\}. \end{aligned}$$

In diesem Fall nennt man π genauer einen *r-Zyklus* und verwendet für π die Schreib-

weise (x_1, \dots, x_r) . Zwei Zyklen (x_1, \dots, x_r) und (y_1, \dots, y_s) heißen *fremd*, wenn

$$\{x_1, \dots, x_r\} \cap \{y_1, \dots, y_s\} = \emptyset$$

gilt. Ein 2-Zyklus heißt auch *Transposition*. Wir notieren ohne Beweis:

5.18 Satz. *Sei $n \geq 2$.*

- (i) *Sind $\pi_1, \pi_2 \in S_n$ fremde Zyklen, so gilt $\pi_1 \circ \pi_2 = \pi_2 \circ \pi_1$.*
- (ii) *Jedes $\pi \in S_n$ ist Produkt paarweise fremder Zyklen. Diese sind eindeutig durch π bestimmt (bis auf Reihenfolge).*
- (iii) *Jedes $\pi \in S_n$ ist Produkt von Transpositionen.*

Aussage (ii) gestattet es uns, eine Permutation $\pi \in S_n$ in *Zyklenschreibweise* anzugeben, d. h. als Produkt paarweise fremder Zyklen. In dieser Schreibweise lässt sich die Konjugationsoperation auf S_n besonders übersichtlich darstellen (vgl. [6], S. 14): Schreiben wir für ein $\pi \in S_n$ mit $a, b, c, d, e, f \in \{1, \dots, n\}$

$$\pi = (a \dots b)(c \dots d) \cdots (e \dots f),$$

so ergibt sich unter Konjugation mit einer beliebigen Permutation $\sigma \in S_n$

$$\sigma \circ \pi \circ \sigma^{-1} = (\sigma(a) \dots \sigma(b)) (\sigma(c) \dots \sigma(d)) \cdots (\sigma(e) \dots \sigma(f)).$$

Man erkennt, dass π und $\sigma\pi\sigma^{-1}$ für jedes $r \in \{1, \dots, n\}$ dieselbe Anzahl an r -Zyklen enthalten; wir sagen auch: π und $\sigma\pi\sigma^{-1}$ sind von demselben *Zyklentyp*. Umgekehrt lässt sich zu zwei Permutationen $\pi_1, \pi_2 \in S_n$ desselben Zyklentyps offensichtlich immer ein $\sigma \in S_n$ angeben mit $\sigma \circ \pi_1 \circ \sigma^{-1} = \pi_2$. Damit ist gezeigt:

5.19 Satz. *Zwei Permutationen aus S_n sind genau dann konjugiert zueinander, wenn sie vom selben Zyklentyp sind.*

Beachte, dass für die Definition des Zyklentyps die Reihenfolge der Zyklen ohne Bedeutung ist. Wir können daher o. B. d. A. annehmen, dass diese nach nichtaufsteigender Länge angeordnet sind. Dann ist offensichtlich, dass die (geordneten)

Zyklentypen in bijektiver Weise den (geordneten) Partitionen von $\{1, \dots, n\}$ entsprechen, und wir haben:

5.20 Korollar. *Die Anzahl der Partitionen von $\{1, \dots, n\}$ und die der Konjugationsklassen von S_n stimmen überein.*

Nun wenden wir uns wieder einer allgemeinen endlichen Gruppe zu. Es gilt, eine Verbindung von der Anzahl der Konjugationsklassen zu der der irreduziblen Darstellungen zu schaffen. Wir führen hierzu (für eine Gruppe G) den Begriff der Klassenfunktion ein: Eine Funktion $f: G \rightarrow \mathbb{C}$ heißt *Klassenfunktion*, wenn sie auf den Konjugationsklassen von G konstant ist. Nach Satz 5.9 sind die Charaktere Klassenfunktionen.

Für eine Klassenfunktion f führen wir weiter die *Fouriertransformierte* \hat{f} ein, die eine Darstellung (ρ, V) von G auf eine lineare Abbildung in $GL(V)$ abbildet vermöge

$$\hat{f}(\rho) := \sum_{s \in G} f(s) \rho(s). \quad (5.27)$$

Damit können wir nun den folgenden Satz beweisen, der beim Beweis der $\text{DCF}^{(n)}$ -Konvergenz in Kapitel 6 Verwendung findet (vgl. Gleichung (6.14)):

5.21 Satz. *Sei f eine Klassenfunktion auf einer Gruppe G und sei $\rho: G \rightarrow GL(V)$ eine irreduzible Darstellung von G mit Charakter χ_ρ . Dann gilt*

$$\hat{f}(\rho) = \lambda I \quad \text{mit} \quad \lambda = \frac{1}{d_\rho} \sum_{s \in G} f(s) \chi_\rho(s) = \frac{|G|}{d_\rho} (f | \overline{\chi_\rho}). \quad (5.28)$$

Beweis. Es gilt

$$\begin{aligned} \rho(t) \hat{f}(\rho) \rho(t)^{-1} &= \sum_{s \in G} \rho(t) f(s) \rho(s) \rho(t^{-1}) = \sum_{s \in G} f(s) \rho(tst^{-1}) \\ &= \sum_{s \in G} f(t^{-1}st) \rho(s) = \sum_{s \in G} f(s) \rho(s) = \hat{f}(\rho) \end{aligned}$$

für alle $t \in G$, und wir können somit den zweiten Teil von Schurs Lemma 5.4 zur Anwendung bringen und erhalten $\hat{f}(\rho) = \lambda I$. Bilden wir hiervon auf beiden Seiten

die Spur, so erhalten wir

$$\lambda d_\rho = \text{Spur}(\lambda I) = \text{Spur}(\hat{f}(\rho)) = \text{Spur}\left(\sum_{s \in G} f(s)\rho(s)\right) = \sum_{s \in G} f(s)\chi_\rho(s).$$

Wie wir bereits wissen, sind die Charaktere orthonormale Klassenfunktionen. Wir zeigen, dass sie überdies den Raum aller Klassenfunktionen aufspannen:

5.22 Satz. *Die Charaktere der irreduziblen Darstellungen bilden eine orthonormale Basis für die Klassenfunktionen.*

Beweis. Wir zeigen indirekt, dass sich jede Klassenfunktion als Linearkombination von derartigen Charakteren schreiben lässt. Seien für ein $k \in \mathbb{N}$ ρ_1, \dots, ρ_k die irreduziblen Darstellungen (bzw. genauer: ein Vetretersystem der Äquivalenzklassen irreduzibler Darstellungen) einer Gruppe und seien χ_1, \dots, χ_k die zugehörigen Charaktere. Weiter sei f eine Klassenfunktion mit $(f|\chi_i) = 0$ für $i = 1, \dots, k$. Dann erhalten wir mit Satz 5.21 $\hat{f}(\rho) = 0$ für jede irreduzible Darstellung ρ und die inverse Fouriertransformation liefert $f = 0$. \square

Damit sind wir fast am Ziel: Die Anzahl der irreduziblen Darstellungen ist also gleich der Dimension des Vektorraumes der Klassenfunktionen. Diese ist aber offenbar gleich der Anzahl der Konjugationsklassen, da eine Klassenfunktion auf den einzelnen Konjugationsklassen jeden beliebigen Wert annehmen kann. Es gilt somit:

5.23 Satz. *Die Anzahl der irreduziblen Darstellungen (modulo Äquivalenz) einer Gruppe G ist gleich der Anzahl der Konjugationsklassen von G .*

Für den Spezialfall $G = S_n$ erhalten wir daraus in Kombination mit Korollar 5.20:

5.24 Korollar. *Die Anzahl der irreduziblen Darstellungen der S_n ist gleich der Anzahl der Partitionen von $\{1, \dots, n\}$.*

5.4 Darstellungen der S_n

Wir wenden uns nun den irreduziblen Darstellungen der Permutationsgruppe S_n zu. Nachdem wir jetzt um ihre Anzahl wissen, werden wir versuchen, sie – soweit wie

möglich – auch anzugeben. Wir überlegen uns zunächst, welche Darstellungen wir sofort angeben können (vgl. [6], Kap. 2.):

- Die *Einheitsdarstellung* ist eindimensional (vgl. S. 60). Wir erhalten sie, indem wir jeder Permutation die identische Abbildung zuordnen:

$$\rho(\pi) = I \quad \text{für alle } \pi \in S_n.$$

- Die *alternierende Darstellung* ist ebenfalls eindimensional. Jeder Permutation $\pi \in S_n$ wird die Multiplikation mit ihrem Signum $\text{sgn}(\pi)$ zugeordnet:

$$\rho(\pi) = \text{sgn}(\pi) \cdot I \quad \text{für alle } \pi \in S_n.$$

Nach Eigenschaften des Signums (vgl. dazu auch [5], S. 252 f.) ist diese Abbildung ein Homomorphismus:

$$\rho(\pi_1 \pi_2) = \text{sgn}(\pi_1 \pi_2) I = \text{sgn}(\pi_1) I \cdot \text{sgn}(\pi_2) I = \rho(\pi_1) \rho(\pi_2).$$

- Die *Permutationsdarstellung* ist eine n -dimensionale Darstellung. Um sie zu definieren, sei e_1, \dots, e_n die Standardbasis von \mathbb{R}^n . Wir setzen für $\pi \in S_n$

$$\rho(\pi)e_i = e_{\pi(i)}.$$

Für die Matrix der linearen Abbildung $\rho(\pi)$ erhalten wir $(\rho(\pi)_{ij}) = (\delta_{i\pi(j)})$.

Mit der Einheitsdarstellung und der alternierenden Darstellung sind damit sogar bereits zwei irreduzible Darstellungen gefunden. Die Permutationsdarstellung hingegen ist nicht irreduzibel – der eindimensionale Unterraum

$$W := \text{span}(e_1 + \dots + e_n)$$

ist invariant unter $\rho(\pi)$ für alle $\pi \in S_n$ (dies führt wieder auf die Einheitsdarstellung). Jedoch enthält sie eine weitere irreduzible Subdarstellung, die sich relativ leicht

angeben lässt. Nach Satz 5.2 ist das orthogonale Komplement von W

$$W^\perp = \left\{ x \in \mathbb{R}^n : \sum_i x_i = 0 \right\}$$

(beachte hierzu $\langle \sum_i x_i e_i, \sum_j e_j \rangle = \sum_{i,j} x_i \delta_{ij} = \sum_i x_i$) ebenfalls invariant. Schränken wir nun die Permutationsdarstellung auf W^\perp ein, so erhalten wir eine irreduzible Darstellung, wie sich zeigen lässt¹. Für die S_3 sind so alle (drei möglichen) irreduziblen Darstellungen gefunden.

Im allgemeinen Fall gibt es jedoch weit mehr irreduzible Darstellungen. Insgesamt existiert nach Korollar 5.24 zu jeder Partition von $\{1, \dots, n\}$ eine irreduzible Darstellung von S_n . Dies werden wir im Folgenden näher ausführen. Dazu ist einige theoretische Vorarbeit notwendig, bei der wir uns wiederum auf [6], Kap. 7 wie auch auf [10] stützen. Zusätzlich verweisen wir noch auf das Buch von James und Kerber [14], das dieses Thema sehr ausführlich behandelt.

Zunächst führen wir eine geeignete Diagramm-Schreibweise für eine Partition $\lambda = (\lambda_1, \dots, \lambda_r) \in \mathcal{P}_n$ ein: Das *Young-Diagramm* von λ ist eine geordnete, linksbündige Menge von Kästchen, bestehend aus N_λ Reihen derart, dass die i -te Reihe λ_i Kästchen umfasst (vgl. Abbildung 5.1). Wir bezeichnen das Diagramm von λ mit $[\lambda]$. Die Kästchen erhalten die Koordinaten (i, j) , wobei i die Zeile von oben gezählt anzeigt und j die Spalte, gezählt von links.

Werden nun in das Young-Diagramm $[\lambda]$ die Zahlen $1, \dots, n \in \mathbb{N}$ in beliebiger Weise in die einzelnen Kästchen eingetragen, so sprechen wir von einem λ -*Tableau*. Offenbar gibt es genau $n!$ λ -Tableaus. Lassen wir nun eine Permutation $\pi \in S_n$ auf ein λ -Tableau wirken, indem wir den Eintrag i durch $\pi(i)$ ersetzen, so definiert dies eine *Aktion* oder *Operation* der symmetrischen Gruppen S_n auf der Menge der λ -Tabloide (vgl. [5], S. 239).

¹ Sei $w_1 := e_1 - e_2$ und $w_2 := e_2 - e_3$. Dann gilt $W^\perp = \text{span}\{w_1, w_2\}$: Für $v = xe_1 + ye_2 + ze_3 \in W^\perp$ lässt sich schreiben $v = xe_1 + ye_2 + (-x - y)e_3 = x(e_1 - e_2) + (x + y)(e_2 - e_3)$. Sei nun $0 \neq (x, y, z) \in W^\perp$ (wir nehmen o. B. d. A. $x \neq 0$ an) und sei $W_1 := \text{span}\{(x, y, z)\}$ invariant unter S_n . Dann liegen $(1, y', z')$ und damit $(y', 1, z')$ und $(1 - y', y' - 1, 0)$ ebenfalls in W_1 . Ist $y' \neq 1$, so liegen $(1, -1, 0) = e_1 - e_2$ und damit auch $(0, 1, -1) = e_2 - e_3$ in W_1 , und wir haben $W_1 = W^\perp$. Ist $y' = 1$, so gilt $(1, 1, z') \in W_1 \rightsquigarrow (1 - z', z' - 1, 0) \in W_1 \rightsquigarrow (1, -1, 0) \in W_1 \rightsquigarrow (1, 1, -2) \in W_1 \rightsquigarrow (0, -1, 1) = -(e_2 - e_3) \in W_1 \rightsquigarrow (1, -1, 0) = e_1 - e_2 \in W_1$, also $W_1 = W^\perp$.

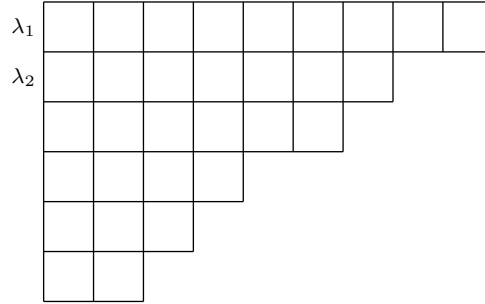


Abbildung 5.1: Das Young-Diagramm der Partition $\lambda = (9, 7, 6, 4, 3, 2, 0, \dots) \in \mathcal{P}_{31}$.

Es liegt nun nahe, auf der Menge der Tableaus eine Äquivalenzrelation einzuführen: Zwei λ -Tableaus t_1 und t_2 mit $\lambda \in \mathcal{P}_n$ heißen *äquivalent*, wenn alle Zeilen von t_1 und t_2 jeweils dieselben Einträge enthalten. Die Äquivalenzklasse eines gegebenen λ -Tableaus t nennen wir ein λ -*Tabloid* und bezeichnen es mit $\{t\}$. Man kann sich ein Tabloid also vorstellen als ein Tableau mit ungeordneten Zeileneinträgen. Die oben eingeführte Aktion von S_n auf der Menge der Tableaus setzt sich in offensichtlicher Weise zu einer Aktion auf der Menge der Tabloide fort. Diese Abbildung ist unabhängig vom gewählten Repräsentanten.

Wir fassen nun die Menge der λ -Tabloide $\{t\}$ zu gegebenem $\lambda \in \mathcal{P}_n$ als Basisvektoren $e_{\{t\}}$ eines Vektorraumes auf. Den so aufgespannten Raum bezeichnen wir mit M^λ , d. h.

$$M^\lambda := \text{span} \{e_{\{t\}} : \{t\} \text{ ist ein } \lambda\text{-Tabloid}\}. \quad (5.29)$$

Für die Aktion von S_n auf M^λ setzen wir

$$\pi(e_{\{t\}}) := e_{\pi(\{t\})}, \quad (5.30)$$

wobei $\pi(\{t\})$ die Aktion von S_n auf der Menge der λ -Tabloide bezeichnet.

Damit haben wir das nötige Handwerkszeug beisammen, um zu einem gegebenen $\lambda \in \mathcal{P}_n$ die *Permutationsdarstellung* $(\rho_\lambda, M^\lambda)$ von S_n in M^λ zu definieren:

$$\rho_\lambda: S_n \longrightarrow GL(M^\lambda) \quad \text{vermöge} \quad \rho_\lambda(\pi)e_{\{t\}} := \pi(e_{\{t\}}) = e_{\pi(\{t\})}. \quad (5.31)$$

Die Darstellung entspricht der Aktion der S_n auf der Menge der λ -Tabloide. Aus diesem Grunde werden wir den Begriff „Darstellung“ nicht immer explizit erwähnen,

sondern lediglich von der Aktion von S_n auf einer Menge sprechen.

Ein Großteil ist jetzt geschafft: Die so definierte Darstellung $(\rho_\lambda, M^\lambda)$ ist zwar reduzibel, enthält aber die irreduzible Darstellung, die wir suchen. Um den entsprechenden Untervektorraum angeben zu können, sind weitere Vorbereitungen vonnöten:

Wir definieren für $\lambda \in \mathcal{P}_n$ zu einem gegebenen λ -Tableau t den *Spaltenstabilisator* $C_t \subset S_n$, der alle Permutationen umfasst, die die Einträge von t nur innerhalb der Spalten umordnen.

Damit können wir nun für jedes λ -Tableau t das *Polytabloid* $e_t \in M^\lambda$ definieren:

$$e_t := \sum_{\pi \in C_t} \text{sgn}(\pi) e_{\pi(\{t\})}. \quad (5.32)$$

Nach Wahl der Tabloide und des Spaltenstabilisators gilt

$$\sigma(e_t) = \sum_{\pi \in C_t} \text{sgn}(\pi) e_{\sigma \circ \pi(\{t\})} = \sum_{\pi \in C_t} \text{sgn}(\pi) e_{\pi \circ \sigma(\{t\})} = e_{\sigma(t)}$$

für alle $\sigma \in S_n$, und es folgt, dass der Raum

$$S^\lambda := \text{span} \{e_t : t \text{ ist ein } \lambda\text{-Tableau}\} \subset M^\lambda \quad (5.33)$$

invariant ist unter der Aktion von S_n ; S^λ heißt auch *Specht-Modul*.

Mit $(\rho_\lambda|_{S^\lambda}, S^\lambda)$ ist nunmehr die zu $\lambda \in \mathcal{P}_n$ gehörige irreduzible Darstellung der S_n gefunden. Um die Irreduzibilität zu zeigen, benötigen wir eine (partielle) Ordnungsrelation auf \mathcal{P}_n : Sind λ und μ zwei Partitionen in \mathcal{P}_n , so sagen wir, λ *dominiert* μ (in Zeichen $\lambda \supseteq \mu$), wenn gilt

$$\begin{aligned} \lambda_1 &\geq \mu_1 \\ \lambda_1 + \lambda_2 &\geq \mu_1 + \mu_2 \\ &\vdots \end{aligned}$$

Es lässt sich zeigen, dass $\lambda \supseteq \mu$ genau dann gilt, wenn wir $[\lambda]$ aus $[\mu]$ erhalten können, indem wir nacheinander Kästchen am rechten Rand abwärts bewegen, so dass jeder Zwischenzustand wieder das Young-Diagramm einer Partition darstellt. Es ist leicht zu sehen, dass aus $\lambda \supseteq \mu$ und $\lambda \subseteq \mu$ schon $\lambda = \mu$ folgt.

Das folgende Lemma gibt eine Charakterisierung der Ordnungsrelation:

5.25 Lemma. *Seien $\lambda, \mu \in \mathcal{P}_n$. Weiter sei t_1 ein λ -Tableau und t_2 ein μ -Tableau derart, dass für jedes k alle Einträge aus der k -ten Zeile von t_2 in verschiedenen Spalten von t_1 enthalten sind. Dann gilt $\lambda \supseteq \mu$.*

Beweis. Da sich die Einträge der ersten Zeile von t_2 in verschiedenen Spalten von t_1 befinden, folgt $\lambda_1 \geq \mu_1$. Die Einträge in der zweiten Zeile von t_2 sind ebenfalls in verschiedenen Spalten von t_1 enthalten, womit folgt, dass jede Spalte von t_1 maximal zwei Einträge aus den ersten beiden Zeilen von t_2 enthält. Verschieben wir nun gedanklich die Kästchen mit diesen Einträgen innerhalb der Spalten von t_1 nach oben, so finden sie in den ersten beiden Zeilen Platz, es gilt also $\lambda_1 + \lambda_2 \geq \mu_1 + \mu_2$. Allgemein lässt sich so für jedes k zeigen, dass jede Spalte von t_1 höchstens k Einträge aus der k -ten Zeile von t_2 enthält, womit $\sum_{i=1}^k \lambda_i \geq \sum_{i=1}^k \mu_i$ gezeigt ist. Insgesamt folgt $\lambda \supseteq \mu$. \square

5.26 Lemma. *Seien $\lambda, \mu \in \mathcal{P}_n$. Weiter seien t ein λ -Tableau und s ein μ -Tableau derart, dass gilt*

$$\sum_{\pi \in C_t} \text{sgn}(\pi) e_{\pi(\{s\})} \neq 0. \quad (5.34)$$

Dann folgt $\lambda \supseteq \mu$. Falls schon $\lambda = \mu$, so gilt

$$\sum_{\pi \in C_t} \text{sgn}(\pi) e_{\pi(\{s\})} = \pm e_t. \quad (5.35)$$

Beweis. Wir nehmen an, es gebe $a, b \in \{1, \dots, n\}$, so dass sich a und b in derselben Zeile von s und in derselben Spalte von t befinden. Dann folgt

$$(id - (ab))e_{\{s\}} = e_{\{s\}} - e_{\{s\}} = 0.$$

Überdies ist die von id und (ab) erzeugte Gruppe $H := \langle id, (ab) \rangle$ eine Untergruppe von C_t . Da (ab) von der Ordnung 2 ist, umfasst H nur diese beiden Elemente. Bezeichnen wir die Repräsentanten der Linksnebenklassen von H in C_t mit $\sigma_1, \dots, \sigma_k$, so können

wir C_t schreiben als (vgl. [5], S. 15 f.)

$$C_t = \bigcup_{i=1}^k \sigma_i H.$$

Setzen wir dies in die obige Summe ein, so erhalten wir

$$\begin{aligned} \sum_{\pi \in C_t} \operatorname{sgn}(\pi) e_{\pi(\{s\})} &= \sum_{i=1}^k \operatorname{sgn}(\sigma_i \circ \operatorname{id}) \sigma_i \circ \operatorname{id}(e_{\{s\}}) + \sum_{i=1}^k \operatorname{sgn}(\sigma_i \circ (ab)) \sigma_i \circ (ab)(e_{\{s\}}) \\ &= \sum_{i=1}^k \operatorname{sgn}(\sigma_i) \sigma_i \circ (\operatorname{id} - (ab))(e_{\{s\}}) = 0. \end{aligned}$$

Dies ist jedoch nach Annahme ausgeschlossen, und es folgt somit, dass die Einträge einer jeden Zeile von s sich in verschiedenen Spalten von t befinden. Mit dem vorangehenden Lemma erhalten wir $\lambda \supseteq \mu$.

Nehme nun an, dass $\lambda = \mu$ gilt und dass die Summe nicht verschwindet. Mit derselben Argumentation wie oben erhalten wir wiederum, dass sich die Einträge einer jeden Zeile von s in verschiedenen Spalten von t befinden. Es gibt daher eine eindeutig bestimmte Permutation $\pi^* \in C_t$ mit $\pi^*(\{t\}) = \{s\}$. Wir haben somit

$$\sum_{\pi \in C_t} \operatorname{sgn}(\pi) e_{\pi(\{s\})} = \sum_{\pi \in C_t} \operatorname{sgn}(\pi) e_{\pi \circ \pi^*(\{t\})} = \operatorname{sgn}(\pi^*)^{-1} e_t = \pm e_t.$$

5.27 Lemma. *Sei $v \in M^\mu$ für ein $\mu \in \mathcal{P}_n$ und sei t ein μ -Tableau. Dann gibt es ein $c \in \mathbb{R}$ mit*

$$\sum_{\pi \in C_t} \operatorname{sgn}(\pi) \pi(v) = c \cdot e_t. \quad (5.36)$$

Beweis. v lässt sich als endliche Linearkombination von μ -Basistabloten $e_{\{s\}}$ schreiben: $v = \sum_s c_s e_{\{s\}}$ mit $c_s \in \mathbb{R}$. Auf jeden einzelnen dieser Basisvektoren wenden wir das vorhergehende Lemma an und erhalten

$$\sum_{\pi \in C_t} \operatorname{sgn}(\pi) \pi(v) = \sum_s c_s \sum_{\pi \in C_t} \operatorname{sgn}(\pi) e_{\pi(\{s\})} = \sum_s c_s d_s e_t$$

mit $d_s \in \{+1, -1, 0\}$. □

Als Nächstes führen wir nun ein inneres Produkt $\langle \cdot, \cdot \rangle$ auf M^μ ein, unter dem die $e_{\{s\}}$ orthonormal sind:

$$\langle e_{\{s\}}, e_{\{t\}} \rangle := \begin{cases} 1 & \text{falls } \{s\} = \{t\}, \\ 0 & \text{sonst.} \end{cases}$$

Dieses Produkt ist invariant unter der Aktion von S_n , da $\pi(\{s\}) = \pi(\{t\})$ genau dann gilt, wenn $\{s\} = \{t\}$ ist. Damit haben wir für beliebige $u, v \in M^\mu$ mit der abkürzenden Schreibweise $A_t(\cdot) := \sum_{\pi \in C_t} \text{sgn}(\pi) \pi(\cdot)$ und wegen $\text{sgn}(\pi^{-1}) = \text{sgn}(\pi)$

$$\langle A_t u, v \rangle = \sum_{\pi \in C_t} \text{sgn}(\pi) \langle \pi(u), v \rangle = \sum_{\pi \in C_t} \text{sgn}(\pi) \langle u, \pi^{-1}(v) \rangle = \langle u, A_t v \rangle. \quad (5.37)$$

Mithilfe dieses Produktes lässt sich der folgende Satz beweisen:

5.28 Satz (Submodul-Satz). *Sei U ein unter der Aktion von S_n invarianter Untervektorraum von M^λ für ein $\lambda \in \mathcal{P}_n$. Dann gilt entweder $S^\lambda \subset U$ oder $U \subset S^{\lambda^\perp}$. Insbesondere ist $(\rho_\lambda|_{S^\lambda}, S^\lambda)$ eine irreduzible Subdarstellung von $(\rho_\lambda, M^\lambda)$.*

Beweis. Sei $u \in U$ und t ein λ -Tableau. Nach Lemma 5.27 gilt $A_t u = c \cdot e_t$ für ein $c \in \mathbb{R}$. Ist es möglich, u so zu wählen, dass c nicht verschwindet, so folgt wegen der Invarianz von U

$$e_t = \frac{1}{c} \sum_{\pi \in C_t} \text{sgn}(\pi) \pi(u) \in U.$$

Damit gilt weiter $e_{\pi(t)} = \pi(e_t) \in U$ für alle $\pi \in S_n$, und somit liegt e_s für jedes λ -Tableau s in U . Dies beweist $S^\lambda \subset U$.

Gilt hingegen $A_t u = 0$ für alle Wahlen von t und u , so haben wir

$$0 = \langle A_t u, e_{\{t\}} \rangle = \langle u, A_t e_{\{t\}} \rangle = \langle u, e_t \rangle$$

für alle $u \in U$ und alle λ -Tableaus t . Damit ist $U \subset S^{\lambda^\perp}$ gezeigt.

Die Irreduzibilität von $(\rho_\lambda|_{S^\lambda}, S^\lambda)$ lässt sich nun wie folgt zeigen: Ist $W \neq \{0\}$ ein unter ρ_λ invarianter Untervektorraum von S^λ , so gilt schon $W \supset S^\lambda$, also $W \equiv S^\lambda$. $(\rho_\lambda|_{S^\lambda}, S^\lambda)$ enthält also nur die trivialen Subdarstellungen. \square

Damit können wir nun zu jedem $\lambda \in \mathcal{P}_n$ mittels S^λ eine irreduzible Darstellung angeben. Es bleibt jedoch noch nachzuweisen, dass damit in der Tat *alle* irreduziblen Darstellungen gefunden sind. Wie bereits angegeben, ist die Anzahl dieser gleich der Anzahl der Partitionen von $\{1, \dots, n\}$. Wir müssen also nur noch zeigen, dass die $(\rho_\lambda, S^\lambda)$ für $\lambda \in \mathcal{P}_n$ (paarweise) nicht äquivalent sind. Hierfür benötigen wir zwei weitere Lemmata:

5.29 Lemma. *Seien $\lambda, \mu \in \mathcal{P}_n$ und sei $T: M^\lambda \longrightarrow M^\mu$ eine lineare Abbildung, die mit der Aktion von S_n auf M^λ bzw. M^μ vertauscht. Ist $S^\lambda \not\subseteq \text{Ker } T$, so gilt $\lambda \supseteq \mu$. Ist schon $\lambda = \mu$, so gilt: Die Einschränkung von T auf S^λ ist ein skalares Vielfaches der Identität.*

Beweis. Sei $u \in M^\lambda$ mit $T(u) = 0$. Dann haben wir $T(\pi(u)) = \pi(T(u)) = 0$; $\text{Ker } T$ ist also invariant unter der Aktion von S_n . Satz 5.28 liefert sodann $\text{Ker } T \subset S^{\lambda^\perp}$, da $S^\lambda \subseteq \text{Ker } T$ nach Annahme ausgeschlossen ist. Damit gilt für jedes λ -Tableau t

$$0 \neq Te_t = TA_te_{\{t\}} = A_tTe_{\{t\}}.$$

Nun ist aber $Te_{\{t\}}$ eine Linearkombination von μ -Tabloiden $e_{\{s\}}$, und für mindestens ein solches muss somit gelten

$$A_te_{\{s\}} \neq 0.$$

Mit Lemma 5.26 folgt daher $\lambda \supseteq \mu$ bzw. $Te_t = ce_t$ im Falle von $\lambda = \mu$. □

5.30 Lemma. *Seien $\lambda, \mu \in \mathcal{P}_n$ und sei $T: S^\lambda \longrightarrow S^\mu$ eine lineare Abbildung, die mit der Aktion von S_n auf S^λ bzw. S^μ vertauscht. Ist $T \neq 0$, so gilt $\lambda \supseteq \mu$.*

Beweis. Jedes solches T kann zu einer Abbildung $\bar{T}: M^\lambda \longrightarrow M^\mu$ fortgesetzt werden, indem wir

$$\bar{T} := \begin{cases} T & \text{auf } S^\lambda, \\ 0 & \text{auf } S^{\lambda^\perp} \end{cases}$$

definieren. Da S^λ und S^{λ^\perp} invariant sind unter S_n , vertauscht \bar{T} ebenfalls mit der Aktion von S_n . Ist nun $T \neq 0$ und damit auch $\bar{T} \neq 0$, so folgt $S^\lambda \not\subseteq \text{Ker } \bar{T}$, und Lemma 5.29 liefert $\lambda \supseteq \mu$. □

Damit können wir schließlich zeigen:

5.31 Satz. Die Darstellungen $(\rho_\lambda|_{S^\lambda}, S^\lambda)$, $\lambda \in \mathcal{P}_n$ liefern alle irreduziblen Darstellungen der S_n .

Beweis. Seien $\lambda, \mu \in \mathcal{P}_n$. Wir zeigen, dass aus der Äquivalenz von $(\rho_\lambda|_{S^\lambda}, S^\lambda)$ und $(\rho_\mu|_{S^\mu}, S^\mu)$ bereits $\lambda = \mu$ folgt. Angenommen, ρ_λ und ρ_μ sind äquivalente Darstellungen, so gibt es eine bijektive lineare Abbildung

$$T: S^\lambda \longrightarrow S^\mu \quad \text{mit} \quad T \circ \rho_\lambda(\pi) = \rho_\mu(\pi) \circ T \quad \text{für alle } \pi \in S_n.$$

Aufgrund der Bijektivität von T können wir das vorhergehende Lemma in beiden Richtungen anwenden und erhalten $\lambda \supseteq \mu$ sowie $\lambda \subseteq \mu$, was zusammen $\lambda = \mu$ liefert. \square

Wir haben damit unsere gewünschte Beziehung zwischen Partitionen und irreduziblen Darstellungen gefunden. Leider sind die betreffenden Specht-Moduln recht unanschaulich. Wir hoffen, dem etwas abhelfen zu können, indem wir noch kurz eine Basis angeben:

Wir definieren t als ein *Standardtableau*, wenn die Einträge im zugehörigen Young-Diagramm sowohl entlang der Zeilen nach rechts als auch entlang der Spalten nach unten abnehmen. Es lässt sich zeigen (vgl. [14], S. 303 f.), dass zu gegebenem $\lambda \in \mathcal{P}_n$ die Menge $\{e_t | t \text{ ist ein Standardtableau der Form } \lambda\}$ eine Basis von S^λ bildet.

Im Anhang A.2 findet sich schließlich eine Tabelle, die die Dimension aller irreduziblen Darstellungen der S_{10} sowie den Wert des zugehörigen Charakters auf der Konjugationsklasse der Transpositionen angibt.

5.5 $DCF^{(n)}$ als Markov-Kette auf S_n

Wir wollen nun noch einmal genauer auf die Beziehung zwischen (diskreten) Partitionen aus \mathcal{P}_n und Konjugationsklassen von S_n eingehen und darstellen, inwiefern der $DCF^{(n)}$ -Prozess als Prozess auf den Konjugationsklassen von S_n verstanden werden kann. Mit diesem Wissen bestimmen wir schließlich die invariante Verteilung von $DCF^{(n)}$.

Wie bereits gezeigt, entspricht jeder Partition von $\{1, \dots, n\}$ genau eine Konjugationsklasse von S_n , welche wiederum genau einen Zyklentyp repräsentiert. Wir

können daher die Konjugationsklassen bzw. die Zyklentypen mit den Partitionen identifizieren. Insbesondere stimmen die Anzahl der Partitionen von $\{1, \dots, n\}$ und die der Konjugationsklassen von S_n überein. Für ein gegebenes $\pi \in S_n$ existiert somit (genau) eine Partition $\gamma \in \mathcal{P}_n$, die die Konjugationsklasse von $\pi \in S_n$ repräsentiert. Betrachten wir nun die Aktion von S_n auf sich selbst

$$S_n \times S_n \longrightarrow S_n, \quad (\sigma, \pi) \longmapsto \sigma \circ \pi, \quad (5.38)$$

so ist offensichtlich, dass diese den Zyklentyp einer Permutation und damit die zugehörige Konjugationsklasse ändern kann. Dies gilt auch, wenn wir statt allgemeiner Permutationen σ nur Transpositionen τ zulassen, die überdies zufällig sind, d. h. unter einem Maß $\kappa^{(n)} \in \mathcal{M}_1(S_n)$ definiert durch

$$\kappa^{(n)}(\sigma) := \begin{cases} \frac{2}{n(n-1)} & \text{für } \sigma = \tau \text{ Transposition} \\ 0 & \text{sonst,} \end{cases} \quad (5.39)$$

gewählt werden. Durch die wiederholte (Links-)Aktion zufälliger Transpositionen auf S_n erhalten wir sodann eine Markov-Kette auf S_n . Für den zugehörigen Übergangskern

$$M^{(n)}: S_n \times \mathcal{B}_{S_n} \longrightarrow [0, 1]. \quad (5.40)$$

gilt: Zu einem Übergang $\sigma \rightarrow \sigma'$ kommt es genau dann, wenn die Permutation $\sigma' \circ \sigma^{-1}$ (von links) auf σ wirkt, somit ist die Wahrscheinlichkeit für diesen Übergang gerade die Wahrscheinlichkeit von $\sigma' \circ \sigma^{-1}$ unter $\kappa^{(n)}$. Wir erhalten also

$$M^{(n)}(\sigma, \sigma') = \kappa^{(n)}(\sigma' \circ \sigma^{-1}). \quad (5.41)$$

Den so gewonnenen Prozess können wir nun nach \mathcal{P}_n abbilden, indem wir jedem $\sigma \in S_n$ die dem Zyklentyp von σ entsprechende Partition $\gamma \in \mathcal{P}_n$ zuordnen. Es gilt:

5.32 Lemma. *Die sukzessive Aktion zufälliger Transpositionen auf S_n , abgebildet nach \mathcal{P}_n , liefert eine Markov-Kette auf \mathcal{P}_n . Für den zugehörigen Übergangskern*

$\mathbf{M}^{(n)}: \mathcal{P}_n \times \mathcal{B}_{\mathcal{P}_n} \longrightarrow [0, 1]$ gilt

$$\mathbf{M}^{(n)}(\gamma, \gamma') = \frac{1}{|K_\gamma|} \sum_{\sigma \in K_\gamma} \sum_{\sigma' \in K_{\gamma'}} \mathbf{M}^{(n)}(\sigma, \sigma'), \quad (5.42)$$

wobei K_γ die zu $\gamma \in \mathcal{P}_n$ gehörige Konjugationsklasse von S_n bezeichnet.

Beweis. Wir zeigen, dass die Übergänge in \mathcal{P}_n nicht von der betrachteten Permutation in S_n , d. h. dem gewählten Repräsentanten einer Konjugationsklasse, abhängen.

Seien also $\gamma, \gamma' \in \mathcal{P}_n$, und sei σ eine Permutation aus K_γ . Weiter sei τ_1 eine Transposition, die σ auf eine Permutation $\sigma' \in K_{\gamma'}$ abbildet. Ist nun ρ eine weitere Permutation aus K_γ , so gibt es ein $\pi \in S_n$ mit $\rho = \pi \circ \sigma \circ \pi^{-1}$. Setzen wir daher $\tilde{\tau}_1 := \pi \circ \tau_1 \circ \pi^{-1}$, folgt

$$\rho' := \tilde{\tau}_1 \circ \rho = \pi \circ \tau_1 \circ \sigma \circ \pi^{-1} = \pi \circ \sigma' \circ \pi^{-1} \in K_{\gamma'}.$$

ρ wird also durch die Transposition $\tilde{\tau}_1$ ebenfalls auf eine Permutation in $K_{\gamma'}$ abgebildet. Existiert noch ein weiterer Übergang $\sigma \rightarrow \sigma'' \in K_{\gamma'}$ mittels einer Transposition $\tau_2 \neq \tau_1$, so erhalten wir wiederum vermöge $\tilde{\tau}_2 := \pi \circ \tau_2 \circ \pi^{-1} \neq \tilde{\tau}_1$ einen Übergang von ρ nach $\rho'' := \pi \circ \sigma'' \circ \pi^{-1} \in K_{\gamma'}$. Somit gibt es zu jedem Übergang von σ nach $K_{\gamma'}$ immer auch einen Übergang von ρ nach $K_{\gamma'}$. Da alle (möglichen) Übergänge dieselbe Wahrscheinlichkeit haben, erhalten wir

$$\mathbf{M}^{(n)}(\rho, K_{\gamma'}) \geq \mathbf{M}^{(n)}(\sigma, K_{\gamma'}),$$

woraus insgesamt

$$\mathbf{M}^{(n)}(\rho, K_{\gamma'}) = \mathbf{M}^{(n)}(\sigma, K_{\gamma'}) \quad \text{für alle } \rho, \sigma \in K_\gamma$$

folgt.

Für den Übergangskern $\mathbf{M}^{(n)}$ des Prozesses auf \mathcal{P}_n gilt daher

$$\mathbf{M}^{(n)}(\gamma, \gamma') = \mathbf{M}^{(n)}(\sigma, K_{\gamma'}),$$

wobei σ eine beliebige Permutation aus K_γ ist. Dies können wir auch schreiben als

$$\begin{aligned}\mathbf{M}^{(n)}(\gamma, \gamma') &= \frac{1}{|K_\gamma|} \sum_{\sigma \in K_\gamma} \mathbf{M}^{(n)}(\sigma, K_{\gamma'}) \\ &= \frac{1}{|K_\gamma|} \sum_{\sigma \in K_\gamma} \sum_{\sigma' \in K_{\gamma'}} \mathbf{M}^{(n)}(\sigma, \sigma').\end{aligned}\quad \square$$

Obiges Ergebnis lässt sich weiter präzisieren:

5.33 Satz. *Der $DCF^{(n)}$ -Prozess auf \mathcal{P}_n entspricht der sukzessiven Aktion zufälliger Transpositionen auf S_n , abgebildet nach \mathcal{P}_n .*

Beweis. Wir zeigen, dass die Wahrscheinlichkeiten für einen Übergang $\gamma \rightarrow \gamma'$ in \mathcal{P}_n übereinstimmen, d. h. $\mathbf{M}^{(n)}(\gamma, \gamma') = \mathbf{M}^{(n)}(\gamma, \gamma')$. Dazu genügt es gemäß obigen Ausführungen, ein beliebiges $\sigma \in K_\gamma$ zu betrachten.

Wir überlegen uns zunächst, wie σ durch eine Transposition $\tau = (ij)$ abgebildet wird. Es gilt: Liegen i und j in verschiedenen Zyklen von σ , d. h., enthält σ Zyklen

$$(a \dots b i \dots c) \quad \text{und} \quad (d \dots e j \dots f),$$

so werden diese zu einem Zyklus der Form

$$(a \dots b j \dots f d \dots e i \dots c)$$

verbunden. Liegen i und j hingegen in demselben Zyklus, d. h., σ enthält einen Zyklus

$$(a \dots b i \dots c j \dots d),$$

so wird dieser in zwei Zyklen

$$(a \dots b j \dots d) \quad \text{und} \quad (i \dots c)$$

aufgespaltet.

Bilden wir diese Übergänge nach \mathcal{P}_n ab, so gilt: Im ersten Fall werden zwei Abschnitte von γ verschmolzen; im zweiten Fall wird ein Abschnitt γ in zwei neue Abschnitte aufgeteilt. Nach Wahl von $\kappa^{(n)}$ sind die Wahrscheinlichkeiten hierfür

offensichtlich gleich den entsprechenden Wahrscheinlichkeiten des DCF⁽ⁿ⁾-Übergangs (vgl. Kap. 2.4).

Es bleibt noch zu zeigen, dass bei der Aufspaltung eines Zyklus die Längen der entstehenden Teilzyklen genauso verteilt sind, wie es dem DCF⁽ⁿ⁾-Übergang entspricht.

Wir nehmen dazu o. B. d. A. an, der Zyklus $(1, \dots, k)$ mit $k \leq n$ werde aufgespalten. Die Transposition (ij) stellen wir mittels zweier Zufallsvariablen ξ_1, ξ_2 dar, die wir unabhängig und ohne Zurücklegen gleichmäßig aus $\{1, \dots, k\}$ wählen, und setzen $(ij) := (\xi_1, \xi_2)$. Offenbar ist dann die Wahrscheinlichkeit für jede Transposition aus S_k dieselbe wie unter $\kappa^{(k)}$ (bzw. wie unter $\kappa^{(n)}$ bedingt auf S_k).

Wir betrachten nun die Länge des Bruchzyklus, der ξ_1 enthält. Der entstehende Zyklus (ξ_1, \dots) hat genau dann die Länge m , wenn sich für ein $l \in \{1, \dots, k\}$ der Eintrag ξ_1 an l -ter Stelle von $(1, \dots, k)$ befindet und der Eintrag ξ_2 sich an $l + m$ -ter Stelle (bzw. an $l + m - k$ -ter Stelle, falls $l + m > k$). Für dieses Ereignis gibt es offenbar gerade k Realisierungen von ξ_1, ξ_2 , die jeweils die Wahrscheinlichkeit $\frac{1}{k(k-1)}$ haben. Folglich beträgt die Wahrscheinlichkeit, dass der Bruchzyklus (ξ_1, \dots) die Länge m hat, gerade $\frac{1}{k-1}$. Somit entspricht der Split-Übergang dem des DCF⁽ⁿ⁾-Prozesses. \square

Die Aussage des Satzes lässt sich auch so interpretieren: Der DCF⁽ⁿ⁾-Prozess entspricht der Evolution des Zyklentyps einer Permutation aus S_n unter der Aktion von zufälligen Transpositionen.

Hiermit lässt sich nun leicht die stationäre Verteilung des DCF⁽ⁿ⁾-Prozesses ermitteln. Wir bestimmen die unter zufälligen Transpositionen invariante Verteilung auf S_n und übertragen dieses Ergebnis anschließend auf \mathcal{P}_n . Es gilt:

5.34 Satz. *Die Gleichverteilung auf S_n ist die (eindeutig bestimmte) stationäre Verteilung unter der Aktion zufälliger Transpositionen.*

Beweis. Wir fassen die sukzessive Aktion zufälliger Transpositionen auf S_n wiederum als Markov-Kette mit Zustandsraum S_n und Übergangsmatrix $M^{(n)}$ auf. Da jede Permutation aus S_n endliches Produkt von Transpositionen ist (vgl. Satz 5.18), handelt es sich um eine irreduzible, aufgrund der Endlichkeit von S_n positiv rekurrente diskrete Markovkette. Dies impliziert direkt die Existenz einer eindeutig bestimmten stationären Verteilung (vgl. [1], S. 65). Es bleibt daher nur noch zu zeigen, dass

die Gleichverteilung (im Folgenden mit L_{S_n} bezeichnet) stationär ist. Für beliebiges $\pi \in S_n$ gilt

$$\begin{aligned} L_{S_n} M^{(n)}(\pi) &= \sum_{\sigma \in S_n} L_{S_n}(\sigma) \kappa^{(n)}(\pi \circ \sigma^{-1}) \\ &= \sum_{\tau \text{ Transp.}} L_{S_n}(\tau^{-1} \circ \pi) \kappa^{(n)}(\tau) \\ &= \sum_{\tau \text{ Transp.}} \frac{1}{n!} \cdot \frac{2}{n(n-1)} \\ &= \frac{1}{n!} \equiv L_{S_n}(\pi). \end{aligned}$$

Die Gleichverteilung L_{S_n} ist also stationär. □

Wir können daher folgern:

5.35 Korollar. *Das von der Gleichverteilung auf S_n auf den Konjugationsklassen von S_n induzierte Maß ist die (eindeutig bestimmte) stationäre Verteilung des $DCF^{(n)}$ -Prozesses.*

Beweis. Die Stationarität folgt direkt aus der Aussage des Satzes. Die Eindeutigkeit ergibt sich wie oben aus der Irreduzibilität und der Endlichkeit des Zustandsraumes (vgl. auch S. 13). □

Um das von der Gleichverteilung auf S_n induzierte Maß auf \mathcal{P}_n zu beschreiben, müssen wir die Mächtigkeit der einzelnen Konjugationsklassen bestimmen. Es gilt:

5.36 Lemma. *Sei $\gamma \in \mathcal{P}_n$. Die zu γ gehörende Konjugationsklasse von S_n umfasst gerade*

$$\frac{n!}{\prod_{k=1}^n k^{N_\gamma(k)} N_\gamma(k)!} \quad (5.43)$$

Elemente.

Beweis. Wir überlegen uns die Zahl der Permutationen aus S_n , die über den Zyklentyp $\gamma = (\gamma_i)_i$ verfügen, d. h. der Permutationen, die $N_\gamma(k)$ Zyklen der Länge k für $k = 1, \dots, n$ haben.

Zunächst können wir aus jeder Permutation $\pi \in S_n$ eine Permutation $\sigma \in S_n$ mit Zyklentyp γ konstruieren. Wir setzen einfach

$$\sigma := (\pi(1) \dots \pi(\gamma_1)) (\pi(\gamma_1 + 1) \dots \pi(\gamma_1 + \gamma_2)) \cdots (\pi(n - \gamma_j + 1) \dots \pi(n)),$$

wobei j der größte Index mit $\gamma_j \neq 0$ ist. Dies führt offenbar auf $n!$ Permutationen mit Zyklentyp γ , die jedoch nicht alle verschieden sind. Vielfachheiten resultieren aus folgenden beiden Umständen:

- (1) Identische Zyklen werden mehrfach gezählt: Ein Zyklus der Länge k kann auf k verschiedene Arten aufgeschrieben werden (je nachdem, wo der „Anfang“ gewählt wird). Will man diese Vielfachheit beheben, so führt dies zu einem Korrekturfaktor $(\prod_{k=1}^n k^{N_\gamma(k)})^{-1}$.
- (2) Die Reihenfolge von Zyklen gleicher Länge wird fälschlicherweise mit berücksichtigt; dies wird durch den Faktor $(\prod_{k=1}^n N_\gamma(k)!)^{-1}$ kompensiert.

Damit ergibt sich insgesamt das beschriebene Ergebnis. □

Unter der Gleichverteilung auf S_n trägt jedes Element einer Konjugationsklasse die Masse $\frac{1}{n!}$. Damit folgt schließlich:

5.37 Satz. *Die eindeutig bestimmte stationäre Verteilung des $DCF^{(n)}$ -Prozesses ist das durch die Ewens'sche Stichproben-Formel*

$$\pi_S^{(n)}(\gamma) = \frac{1}{\prod_{k=1}^n k^{N_\gamma(k)} N_\gamma(k)!}, \quad \gamma \in \mathcal{P}_n \tag{5.44}$$

gegebene Maß.

Für weitere Ausführungen zur invarianten Verteilung von $DCF^{(n)}$ verweisen wir auf Kapitel 2.6.

6 DCF⁽ⁿ⁾-Konvergenz

Wie bereits erwähnt, konvergiert die Verteilung eines DCF⁽ⁿ⁾-Prozesses $(\ell(k))_{k \geq 0}$ für $k \rightarrow \infty$ gegen $\pi_S^{(n)}$. Allerdings verschlechtert sich diese Konvergenzrate mit größer werdendem n , so dass es nicht direkt möglich ist, einen DCF⁽ⁿ⁾-Prozess in geeigneter Weise mit einem CCF-Prozess zu koppeln, um von der Grenzverteilung von DCF⁽ⁿ⁾ auf diejenige von CCF zu schließen. Um diese Schwierigkeit zu überwinden, benötigen wir eine *schnellere* Konvergenz. Diese erhält man, wenn man als Startpartitionen des DCF⁽ⁿ⁾-Prozesses nur solche mit insgesamt relativ wenigen Abschnitten verwendet. In diesem Falle ist die Wahrscheinlichkeit, dass sich der Prozess lange mit dem Vereinigen vieler kleiner Abschnitte aufhält, reduziert und seine Verteilung nähert sich schnell der Gleichgewichtsverteilung an. Eine solche Konvergenz werden wir allerdings nur auf bestimmten Mengen zeigen können, nämlich einer Familie \mathcal{C} von Ω_1 -Umgebungen. Später, d. h. in Kapitel 7, bleibt nur noch zu zeigen, dass \mathcal{C} umfassend genug ist, um ein Maß durch seine Werte darauf eindeutig festzulegen.

6.1 Vorbereitungen

Wir definieren für $n \in \mathbb{N}$ und $\beta \in (0, 1]$ den Raum $\mathcal{P}_{n,\beta} \subset \mathcal{P}_n$, der alle Partitionen enthält, die über weniger als n^β Abschnitte verfügen:

$$\mathcal{P}_{n,\beta} := \{\ell \in \mathcal{P}_n : N_\ell < n^\beta\} = \{\ell \in \mathcal{P}_n : \ell_{\lceil n^\beta \rceil} = 0\}. \quad (6.1)$$

Als Nächstes werden wir die Familie \mathcal{C} definieren. Sei dazu für $k \in \mathbb{N}$ eine Familie von k -dimensionalen Intervallen

$$I_k := \left\{ (\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1}^k : 0 < a_i < b_i < 1, \sum_{i=1}^k b_i < 1, a_k > 1 - \sum_{i=1}^k a_i \right\} \quad (6.2)$$

gegeben. Die Elemente aus I_k werden dazu dienen, den k -Anfang, d. h. die ersten k Abschnitte, einer Partition $p \in \Omega_1$ zu kontrollieren, indem wir

$$a_i < p_i < b_i \quad \text{für } i = 1, \dots, k$$

verlangen. Der k -Rest ist dann gleichfalls schon abschätzbar, denn die letzte Bedingung in (6.2) sorgt dafür, dass stets

$$\sum_{i>k} p_i = 1 - \sum_{i=1}^k p_i < 1 - \sum_{i=1}^k a_i < a_k$$

gilt.

Wir definieren daher für jedes $(\mathbf{a}, \mathbf{b}) \in I_k$

$$C_{\mathbf{a}, \mathbf{b}} := \{x \in \Omega_1 : x_i \in (a_i, b_i) \text{ für } i = 1, \dots, k\}, \quad (6.3)$$

also die Menge von Partitionen aus Ω_1 , deren k -Anfang in (\mathbf{a}, \mathbf{b}) liegt. $C_{\mathbf{a}, \mathbf{b}}$ ist genau dann nichtleer, wenn das k -dimensionale Intervall (\mathbf{a}, \mathbf{b}) mit der Sortierung verträglich ist, d. h., falls $a_i < \min_{1 \leq j \leq i} b_j$ gilt für $i = 1, \dots, k$. Ein $x \in \Omega_1$ liegt nämlich genau dann in I_k , wenn $a_i < x_i < b_i$ für $i = 1, \dots, k$ gilt, was äquivalent ist zu $a_i < x_i = \min_{1 \leq j \leq i} x_j < \min_{1 \leq j \leq i} b_j$. Ein nichtleeres $C_{\mathbf{a}, \mathbf{b}}$ mit $(\mathbf{a}, \mathbf{b}) \in I_k$ lässt sich damit auch schreiben als

$$C_{\mathbf{a}, \mathbf{b}} = \{x = (x', x'') : x' \in G_{\mathbf{a}, \mathbf{b}}, x'' \in (1 - |x'|_1)\Omega_1\}.$$

Hierbei bezeichnet (\cdot, \cdot) die Verkettung (Hintereinanderfügung), und die Menge $G_{\mathbf{a}, \mathbf{b}}$ enthält alle k -Anfänge von Partitionen, die mit (\mathbf{a}, \mathbf{b}) verträglich sind, also

$$G_{\mathbf{a}, \mathbf{b}} := \left\{x' = (x_i)_{i=1, \dots, k} \in \prod_{i=1}^k (a_i, b_i) : x_1 \geq x_2 \geq \dots \geq x_k\right\}. \quad (6.4)$$

Weiter definieren wir noch

$$\delta_{\mathbf{a}, \mathbf{b}} := \min \left\{1 - \sum_{i=1}^k b_i, a_k - \left(1 - \sum_{i=1}^k a_i\right)\right\}. \quad (6.5)$$

Dieser Wert gibt an, wie stark die Einschränkung ist, die dem k -Rest einer Partition in $C_{\mathbf{a},\mathbf{b}}$ auferlegt ist. Wir erhalten für alle $(x', x'') \in C_{\mathbf{a},\mathbf{b}}$ wegen $|x''|_1 = 1 - |x'|_1$

$$\begin{aligned} 1 - \sum_{i=1}^k b_i &< |x''|_1 < 1 - \sum_{i=1}^k a_i < x_k - a_k + 1 - \sum_{i=1}^k a_i \\ \Rightarrow \quad \delta_{\mathbf{a},\mathbf{b}} &< |x''|_1 < x_k - \delta_{\mathbf{a},\mathbf{b}}. \end{aligned} \quad (6.6)$$

Schließlich setzen wir

$$\mathcal{C} := \{C_{\mathbf{a},\mathbf{b}} : (\mathbf{a}, \mathbf{b}) \in I_k \text{ für ein } k \geq 1\}. \quad (6.7)$$

6.2 Das Konvergenz-Theorem und Bezüge zur Darstellungstheorie

Damit können wir nun das Ergebnis formulieren (vgl. [7], S. 926):

6.1 Theorem. *Sei $\beta \in (0, \frac{1}{2})$. Für jedes $n \in \mathbb{N}$ sei $(X^{(n)}(k))_{k \geq 0}$ eine $DCF^{(n)}$ -Markov-Kette mit Wahrscheinlichkeitsmaß $P^{(n)}$ und Anfangsverteilung $\mu_0^{(n)} \in \mathcal{M}_1(\mathcal{P}_{n,\beta})$. Dann gilt für $C \in \mathcal{C}$, $\beta' > \beta$ und eine Folge natürlicher Zahlen $(k_n)_{n \geq 1}$ mit $k_n \geq n^{\beta'}$*

$$\Delta_C^{(n)}(k_n) := P^{(n)}(X^{(n)}(k_n) \in nC) - \pi_S^{(n)}(nC) \rightarrow 0 \quad \text{für } n \rightarrow \infty. \quad (6.8)$$

Um den Beweis zu führen, verwenden wir Ergebnisse aus der Darstellungstheorie der S_n (vgl. Kap. 5). Beachte, dass alle hier vorkommenden Vektorräume, Funktionswerte, etc. reell sind.

Zunächst sei daran erinnert, dass die Partitionen $\lambda \in \mathcal{P}_n$ gerade den Äquivalenzklassen in S_n unter Konjugation entsprechen. Wir können daher die Konjugationsklassen durch die Partitionen indizieren und schreiben $(K_\lambda)_{\lambda \in \mathcal{P}_n}$. Den Vektorraum der Funktionen, die auf den Konjugationsklassen konstant sind, bezeichnen wir mit

$$F_n := \{f: S_n \longrightarrow \mathbb{R} : f(\sigma\pi\sigma^{-1}) = f(\pi) \quad \text{für alle } \pi, \sigma \in S_n\}. \quad (6.9)$$

Eine solche Klassenfunktion $f \in F_n$ können wir auch als Funktion auf \mathcal{P}_n auffassen,

indem wir für alle $\lambda \in \mathcal{P}_n$ ein beliebiges $\pi \in K_\lambda$ wählen und

$$f(\lambda) := f(\pi)$$

setzen. Den Raum der Funktionen auf \mathcal{P}_n bezeichnen wir mit

$$\mathcal{F}_n := \{f: \mathcal{P}_n \longrightarrow \mathbb{R}\}. \quad (6.10)$$

Die Charaktere $\chi: S_n \longrightarrow \mathbb{R}$ sind ebenfalls Klassenfunktionen und können somit auch als Elemente aus \mathcal{F}_n aufgefasst werden. Da zu jeder Partition $\lambda \in \mathcal{P}_n$ eine irreduzible Darstellung ρ_λ und ein entsprechender Charakter χ_λ gehören, können wir die Charaktere ebenfalls durch die Partitionen indizieren: $(\chi_\lambda)_{\lambda \in \mathcal{P}_n}$.

In Gleichung (5.15) haben wir auf F_n das Produkt

$$(f|g) = \frac{1}{|S_n|} \sum_{\pi \in S_n} f(\pi)g(\pi), \quad f, g \in F_n$$

eingeführt. Fassen wir nun f, g als Elemente aus \mathcal{F}_n auf und setzen

$$\langle f, g \rangle = \langle f, g \rangle_n := \sum_{\gamma \in \mathcal{P}_n} f(\gamma)g(\gamma)\pi_S^{(n)}(\gamma), \quad (6.11)$$

so stimmt der Wert der beiden Produkte überein – denn zerlegen wir die Summe über alle Permutationen aus S_n in eine Summation über alle Konjugationsklassen K_γ in S_n und über die jeweils in K_γ enthaltenen Permutationen, so gilt (vgl. Lemma 5.36)

$$\begin{aligned} (f|g) &= \frac{1}{|S_n|} \sum_{\gamma \in \mathcal{P}_n} \sum_{\pi \in K_\gamma} f(\pi)g(\pi) \\ &= \sum_{\gamma \in \mathcal{P}_n} f(\gamma)g(\gamma) \cdot \frac{|K_\gamma|}{n!} \\ &= \sum_{\gamma \in \mathcal{P}_n} f(\gamma)g(\gamma)\pi_S^{(n)}(\gamma) \equiv \langle f, g \rangle. \end{aligned}$$

Wie bereits gezeigt (vgl. Satz 5.22), bilden die Charaktere der irreduziblen Darstellungen eine Orthonormalbasis von F_n (bezüglich (\cdot, \cdot)). Offenbar bilden die Charaktere, aufgefasst als Elemente aus \mathcal{F}_n , dann ebenfalls eine Orthonormalbasis

von \mathcal{F}_n (bezüglich $\langle \cdot, \cdot \rangle$).

Dies gestattet es uns, die Zustände des $\text{DCF}^{(n)}$ -Prozesses, die wir als Elemente von \mathcal{F}_n auffassen können, in ihre Anteile auf den Basis-Charakteren zu zerlegen. Wir werden zeigen, dass die Charaktere, jeweils multipliziert mit $\pi_S^{(n)}$, Eigenvektoren der $\text{DCF}^{(n)}$ -Übergangsmatrix sind. Damit lässt sich die Evolution des Prozesses mithilfe der Eigenwerte von $M^{(n)}$ darzustellen.

Um die Eigenvektoren und Eigenwerte von $M^{(n)}$ zu ermitteln, greifen wir auf die in Kapitel 5.5 beschriebene Aktion zufälliger Transpositionen auf S_n zurück und betrachten die zugehörige Übergangsmatrix $M^{(n)}$, deren Eigenwerte und -funktionen wir als Erstes bestimmen.

$M^{(n)}$ können wir zunächst als Matrix einer linearen Abbildung

$$M^{(n)}: \mathcal{M}_1(S_n) \longrightarrow \mathcal{M}_1(S_n) \quad (6.12)$$

auffassen (vgl. Gleichung (5.40)). Wir schreiben dazu die Elemente aus $\mathcal{M}_1(S_n)$ als Vektoren mit über S_n indizierten Komponenten, was im Falle eines diskreten Maßes ohne Weiteres möglich ist, und bilden ab, indem wir von rechts mit der Matrix $M^{(n)}$ multiplizieren. Dies behalten wir auch im Folgenden bei. Allgemeiner können wir auch beliebige Funktionen auf S_n zulassen und erhalten durch $M^{(n)}$ eine Abbildung auf dem Vektorraum derselben.

In Kapitel 5.5 haben wir gezeigt, dass die Aktion zufälliger Transpositionen auf S_n eine stationäre Verteilung besitzt, was bedeutet, dass $M^{(n)}$ über mindestens einen nichtnegativen linken Eigenvektor verfügt. Dieser ist offenbar ein Eigenvektor zum Eigenwert 1, denn da $M^{(n)}$ eine stochastische Matrix ist, gehören alle nichtnegativen Eigenvektoren notwendigerweise zum Eigenwert 1. Weil die stationäre Verteilung eindeutig bestimmt ist, gibt es jedoch nur genau einen solchen Vektor, da jede andere nichtnegative normierte Eigenfunktion auf S_n ebenfalls ein invariantes Maß auf S_n definieren würde. Dennoch kann $M^{(n)}$ weitere Eigenvektoren haben, wenn es sich dabei um allgemeine Vektoren handelt. Wir werden nun zeigen, dass alle Charaktere χ_λ Eigenvektoren von $M^{(n)}$ sind:

6.2 Satz. *Sei $M^{(n)}$ die darstellende Matrix einer zufälligen Transposition auf S_n . Dann gilt:*

(i) Die Charaktere $(\chi_\lambda)_{\lambda \in \mathcal{P}_n}$ der irreduziblen Darstellungen von S_n sind Eigenfunktionen von $M^{(n)}$.

(ii) Für den zu χ_λ gehörigen Eigenwert $\theta_\lambda^{(n)}$ gilt $\theta_\lambda^{(n)} = \frac{\chi_\lambda(\tau)}{\chi_\lambda(id)}$, wobei τ eine beliebige Transposition aus S_n ist.

Beweis. (i) Für $\pi \in S_n$ und $\lambda \in \mathcal{P}_n$ gilt (vgl. Gleichung (5.41))

$$\begin{aligned}
 \chi_\lambda M^{(n)}(\pi) &= \sum_{\sigma \in S_n} \kappa^{(n)}(\pi \circ \sigma^{-1}) \chi_\lambda(\sigma) \\
 &= \sum_{\sigma \in S_n} \kappa^{(n)}(\sigma) \chi_\lambda(\sigma^{-1} \circ \pi) \\
 &= \sum_{\tau \in S_n} \kappa^{(n)}(\tau) \chi_\lambda(\tau^{-1} \circ \pi) \\
 &= \sum_{\tau \in S_n} \kappa^{(n)}(\tau) \cdot \text{Spur}(\rho_\lambda(\tau^{-1} \circ \pi)) \\
 &= \text{Spur} \left(\sum_{\tau \in S_n} \kappa^{(n)}(\tau) \rho_\lambda(\tau^{-1}) \rho_\lambda(\pi) \right), \tag{6.13}
 \end{aligned}$$

wobei im letzten Schritt ausgenutzt wurde, dass die Spur eine lineare Abbildung ist. Da für Transpositionen $\tau^{-1} = \tau$ gilt, haben wir

$$\sum_{\tau \in S_n} \kappa^{(n)}(\tau) \rho_\lambda(\tau^{-1}) = \sum_{\tau \in S_n} \kappa^{(n)}(\tau) \rho_\lambda(\tau) = \sum_{\sigma \in S_n} \kappa^{(n)}(\sigma) \rho_\lambda(\sigma),$$

was gerade der Fouriertransformierten $\widehat{\kappa}^{(n)}$ von $\kappa^{(n)}$ an der Stelle ρ_λ entspricht (vgl. Gleichung (5.27)). Wir können nun Satz 5.21 anwenden und erhalten

$$\widehat{\kappa}^{(n)}(\rho_\lambda) = \theta_\lambda^{(n)} \cdot I \quad \text{mit} \quad \theta_\lambda^{(n)} = \frac{1}{d_{\rho_\lambda}} \sum_{\pi \in S_n} \kappa^{(n)}(\pi) \chi_\lambda(\pi). \tag{6.14}$$

Mit diesen Ergebnissen lässt sich die obige Rechnung wie folgt fortführen:

$$\begin{aligned}
 \chi_\lambda M^{(n)}(\pi) &= \text{Spur}(\widehat{\kappa}^{(n)}(\rho_\lambda) \rho_\lambda(\pi)) \\
 &= \text{Spur}(\theta_\lambda^{(n)} \rho_\lambda(\pi)) \\
 &= \theta_\lambda^{(n)} \chi_\lambda(\pi), \tag{6.15}
 \end{aligned}$$

χ_λ ist also Eigenvektor von $M^{(n)}$ zum Eigenwert $\theta_\lambda^{(n)}$.

(ii) Für den Eigenwert $\theta_\lambda^{(n)}$ gilt mit $\chi_\lambda(id) = d_{\rho_\lambda}$ (vgl. Satz 5.9)

$$\theta_\lambda^{(n)} = \frac{1}{d_{\rho_\lambda}} \sum_{\pi \in S_n} \kappa^{(n)}(\pi) \chi_\lambda(\pi) = \frac{1}{d_{\rho_\lambda}} \sum_{\tau' \text{ Transpos.}} \kappa^{(n)}(\tau') \chi_\lambda(\tau') = \frac{\chi_\lambda(\tau)}{\chi_\lambda(id)}, \quad (6.16)$$

wobei τ eine beliebige Transposition in S_n ist; da die Charaktere Klassenfunktionen sind, gilt $\chi_\lambda(\tau') = \chi_\lambda(\tau)$ für alle Transpositionen $\tau, \tau' \in S_n$. \square

Dieses Ergebnis lässt sich direkt auf den $DCF^{(n)}$ -Prozess übertragen:

6.3 Korollar. *Die Eigenfunktionen der Übergangsmatrix des $DCF^{(n)}$ -Prozesses $\mathbb{M}^{(n)}$ sind die Charaktere $(\chi_\lambda)_{\lambda \in \mathcal{P}_n}$ der irreduziblen Darstellungen der S_n , jeweils multipliziert mit $\pi_S^{(n)}$. Die zugehörigen Eigenwerte stimmen überein mit denen von $M^{(n)}$.*

Beweis. Sei $\gamma \in \mathcal{P}_n$. Unter Verwendung von Gleichung (5.42) mit $\mathbf{M}^{(n)} \equiv \mathbb{M}^{(n)}$ ergibt sich

$$\begin{aligned} \pi_S^{(n)} \chi_\lambda \mathbb{M}^{(n)}(\gamma) &= \sum_{\gamma' \in \mathcal{P}_n} \pi_S^{(n)}(\gamma') \chi_\lambda(\gamma') \mathbb{M}^{(n)}(\gamma', \gamma) \\ &= \sum_{\gamma' \in \mathcal{P}_n} \frac{|K_{\gamma'}|}{n!} \chi_\lambda(\gamma') \frac{1}{|K_{\gamma'}|} \sum_{\sigma' \in K_{\gamma'}} \sum_{\sigma \in K_\gamma} M^{(n)}(\sigma', \sigma) \\ &= \sum_{\gamma' \in \mathcal{P}_n} \frac{1}{n!} \sum_{\sigma' \in K_{\gamma'}} \sum_{\sigma \in K_\gamma} \chi_\lambda(\sigma') M^{(n)}(\sigma', \sigma) \\ &= \frac{1}{n!} \sum_{\sigma \in K_\gamma} \sum_{\sigma' \in S_n} \chi_\lambda(\sigma') M^{(n)}(\sigma', \sigma) \\ &= \frac{1}{n!} \sum_{\sigma \in K_\gamma} \theta_\lambda^{(n)} \chi_\lambda(\sigma) \\ &= \frac{|K_\gamma|}{n!} \theta_\lambda^{(n)} \chi_\lambda(\gamma) \\ &= \theta_\lambda^{(n)} \pi_S^{(n)}(\gamma) \chi_\lambda(\gamma). \end{aligned} \quad \square$$

Da nun die Eigenvektoren und Eigenwerte von $\mathbb{M}^{(n)}$ gefunden sind, überlegen wir uns, inwieweit wir diese bzw. die Charaktere berechnen oder geeignet abschätzen können. Die Eigenwerte $\theta_\lambda^{(n)}$ lassen sich prinzipiell sogar auch bestimmen, ohne dass

man den Charakter χ_λ und die Dimension der Darstellung kennt; dies ist jedoch nicht immer möglich. Die entsprechende Formel lautet (vgl. [6], D-2, S. 40):

$$\theta_\lambda^{(n)} = \frac{1}{n(n-1)} \sum_{j=1}^n \lambda_j(\lambda_j - 2j + 1) = \frac{1}{n(n-1)} \left(\sum_{i=1}^n \lambda_i^2 - \sum_{j=1}^{\lambda_1} \lambda_j'^2 \right). \quad (6.17)$$

Dabei bezeichnet λ' die adjungierte Partition, die weiter unten definiert wird.

Für unsere Zwecke ist vor allem der Charakter der trivialen Partition $\lambda = (n, 0, \dots)$ wichtig. Diesen können wir anhand unserer Überlegungen in Kapitel 5.4 direkt bestimmen: Wir nutzen aus, dass es für diese Partition nur *ein* Standardtableau t gibt (vgl. S. 83); die zugehörige irreduzible Darstellung $(\rho_{(n,0,\dots)}, S^{(n,0,\dots)})$ mit $S^{(n,0,\dots)} = \text{span}\{e_t\}$ ist also eindimensional. Da das $(n, 0, \dots)$ -Tabloid $\{t\}$ über ungeordnete Zeileneinträge verfügt, gilt $\rho_{(n,0,\dots)}(\pi)(\{t\}) = \{t\}$ für alle $\pi \in S_n$. Somit wird das Polytabloid $e_t \equiv e_{\{t\}}$ (vgl. Gleichung (5.32)) unter allen $\rho_{(n,0,\dots)}(\pi)$ auf sich selbst abgebildet, wir haben also

$$\rho_{(n,0,\dots)}(\pi) = I \quad \text{für alle } \pi \in S_n. \quad (6.18)$$

Es folgt $\chi_{(n,0,\dots)} \equiv d_{\rho_{(n,0,\dots)}} = 1$ und damit auch $\theta_{(n,0,\dots)}^{(n)} = 1$.

Um auch im allgemeinen Fall die Charaktere bestimmen zu können, greifen wir noch einmal auf das Young-Diagramm $[\lambda]$ einer Partition $\lambda \in \mathcal{P}_n$ zurück (vgl. S. 76) und definieren einige neue Größen:

1. $\lambda' \in \mathcal{P}_n$ ist die Partition, die man erhält, wenn man das Young-Diagramm von λ transponiert: $[\lambda'] = [\lambda]^\top$.
2. $B_\lambda := \max \{i : (i, i) \in [\lambda]\} = \max \{i : \lambda_i \geq i\}$ ist die Diagonalenlänge von $[\lambda]$.
3. $R_\lambda(i, j) := \{(u, v) : i \leq u \leq \lambda'_j, j \vee \lambda_{u+1} \leq v \leq \lambda_u\}$ ist das Randsegment von $[\lambda]$, das durch (i, j) eingegrenzt wird.
4. $[\lambda_*^{(i,j)}] := [\lambda] \setminus R_\lambda(i, j)$ definiert eine Partition $\lambda_*^{(i,j)} \in \mathcal{P}_{n - \#R_\lambda(i,j)}$. Beachte dazu, dass das Entfernen eines Randsegments von $[\lambda]$ wieder das Young-Diagramm einer Partition liefert.

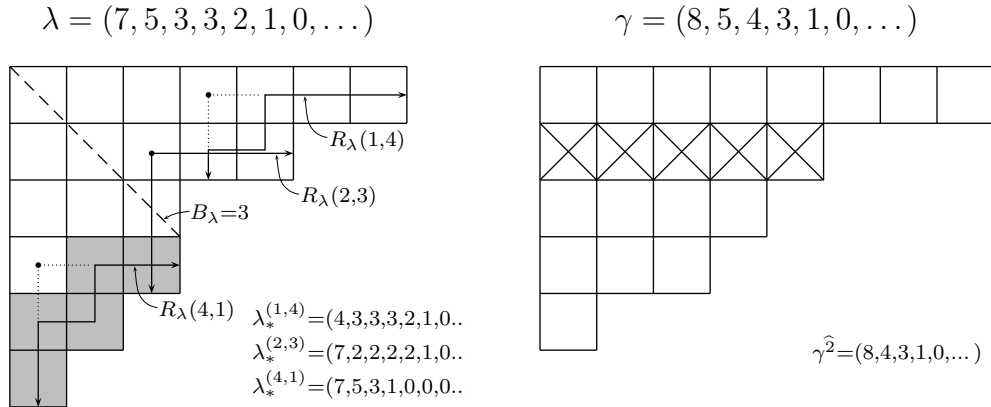


Abbildung 6.1: Die Young-Diagramme zweier Partitionen $\lambda, \gamma \in \mathcal{P}_{21}$. Die λ -Zellen $(1,4)$, $(2,3)$ und $(4,1)$ erzeugen jeweils ein Randsegment der Länge 5 ($R_\lambda(4,1)$ ist grau hinterlegt), die bei Anwendung der Murnaghan-Nakayama-Formel zusammen mit γ_2 gelöscht werden. Es gilt somit $\chi_\lambda = -\chi_{\lambda_*^{(1,4)}}(\gamma^2) + \chi_{\lambda_*^{(2,3)}}(\gamma^2) + \chi_{\lambda_*^{(4,1)}}(\gamma^2)$.

Schließlich definieren wir noch für $\gamma \in \mathcal{P}_n$ die Partition, die man aus γ erhält, wenn man den r -ten Abschnitt entfernt:

$$\gamma^{\hat{r}} := (\gamma_1, \dots, \gamma_{r-1}, \gamma_{r+1}, \dots). \quad (6.19)$$

In Abbildung 6.1 findet sich eine Veranschaulichung der neu eingeführten Größen.

Damit steht nun das Begriffsinventar für die folgende Formel bereit, die es uns erlaubt, Charaktere rekursiv zu berechnen (vgl. [10], Theorem 3.4):

6.4 Murnaghan-Nakayama-Formel. Für Partitionen $\lambda, \gamma \in \mathcal{P}_n$ und $1 \leq r \leq N_\gamma$ gilt

$$\chi_\lambda(\gamma) = \sum_{(i,j) : |R_\lambda(i,j)| = \gamma_r} (-1)^{\lambda_j' - i} \chi_{\lambda_*^{(i,j)}}(\gamma^{\hat{r}}), \quad (6.19)$$

wobei die Summe verschwindet, wenn über eine leere Indexmenge summiert wird, sowie $\chi_\emptyset(\emptyset) = 1$.

Dies bedeutet: Für gegebene $\lambda, \gamma \in \mathcal{P}_n$ lässt sich der Charakter $\chi_\lambda(\gamma)$ berechnen, indem wir ihn auf eine Summe von Charakteren reduzierter Partitionen zurückführen. Diese Reduktion geschieht, indem von $[\lambda]$ alle Randsegmente der Länge γ_{i_1} entfernt

werden, während gleichzeitig der Abschnitt γ_{i_1} von γ gelöscht wird. Dieses Vorgehen wird nun mit $\gamma_{i_2}, \gamma_{i_3} \dots$ wiederholt, bis letztlich nur noch Charaktere der leeren Partition χ_\emptyset , ausgewertet an \emptyset , übrigbleiben. Ist in einem bestimmten Reduktionsschritt eine solche Ausschöpfung von $[\lambda]$ durch γ nicht möglich, weil kein Randsegment der geforderten Länge existiert, verschwindet der betreffende Summand. Insgesamt ist es somit möglich, den Charakter $\chi_\lambda(\gamma)$ als Summe $\sum \pm \chi_\emptyset(\emptyset) = \sum \pm 1$ zu schreiben.

Insbesondere gilt $\chi_\lambda(\gamma) = 0$, falls es keine Möglichkeit gibt, $[\lambda]$ durch γ vollständig auszuschöpfen. Da jedes Randsegment R_λ maximal ein Kästchen auf der Diagonalen beinhaltet, muss γ mindestens über so viele Abschnitte verfügen, wie $[\lambda]$ Kästchen auf der Diagonalen besitzt, sonst kann $[\lambda]$ nicht vollständig ausgeschöpft werden und der Charakter verschwindet an der entsprechenden Stelle. Es gilt also für $\lambda, \gamma \in \mathcal{P}_n$

$$B_\lambda > N_\gamma \quad \Rightarrow \quad \chi_\lambda(\gamma) = 0. \quad (6.20)$$

Ein Beweis für die Murnaghan-Nakayama-Formel findet sich in [14], Kap. 2.4, S. 58 ff.

6.3 Beweis des Theorems

Bevor wir mit dem eigentlichen Beweis beginnen, ist es sinnvoll, eine Charakterisierung der Partitionen $\gamma \in \mathcal{P}_n$ vorzunehmen, die in $nC = nC_{\mathbf{a}, \mathbf{b}}$ für ein gegebenes $k \in \mathbb{N}$ und $(\mathbf{a}, \mathbf{b}) \in I_k$ liegen (wir nehmen an, $C \neq \emptyset$). Gemäß den Eigenschaften von $C_{\mathbf{a}, \mathbf{b}}$ kann jedes solches γ ausgedrückt werden als die Verkettung (γ', γ'') von einem $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$ und einem $\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}$. Hierbei bezeichnet $G_{\mathbf{a}, \mathbf{b}}^{(n)}$ die diskrete Entsprechung von $G_{\mathbf{a}, \mathbf{b}}$, d. h.

$$G_{\mathbf{a}, \mathbf{b}}^{(n)} := \{(\ell_i)_{i=1, \dots, k} \in \{0, \dots, n\}^n : \ell_1 \geq \dots \geq \ell_k, \ell_i \in (na_i, nb_i) \text{ für } i = 1, \dots, k\}. \quad (6.21)$$

Aus Gleichung (6.6) erhalten wir sodann:

- (i) $|\gamma'|_1 < n$,
- (ii) $\exists \delta = \delta(C) > 0$, so dass $\gamma'_k > (n - |\gamma'|_1) + \delta n$ und $|\gamma''|_1 = n - |\gamma'|_1 > \delta n$ gilt (wähle bspw. $\delta := \delta_{\mathbf{a}, \mathbf{b}}$).

Dieser Sachverhalt ist in der folgenden Abbildung illustriert:

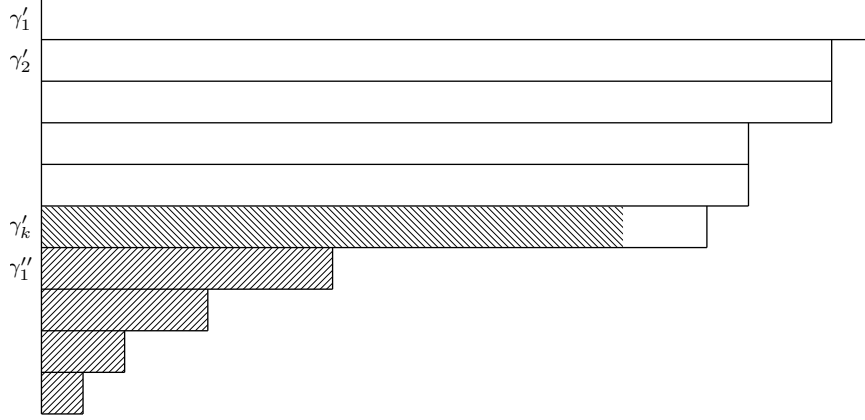


Abbildung 6.2: Eine Partition $\gamma = (\gamma', \gamma'')$ aus $nC_{\mathbf{a}, \mathbf{b}}$ (hier $n = 122$) lässt sich in einen Anteil $\gamma' \in nG_{\mathbf{a}, \mathbf{b}}^{(n)}$ und einen Anteil $\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}$ zerlegen. Nach Wahl von (\mathbf{a}, \mathbf{b}) ist $\gamma'' \neq \emptyset$ und es gilt $\gamma'_k > |\gamma''|_1$. Dies ist auch in der Abbildung wiedergegeben: Die beiden schraffierten Bereiche entsprechen jeweils $|\gamma''|_1$.

Beweis von Theorem 6.1. Um $\Delta_C^{(n)}(k_n)$ abzuschätzen, zerlegen wir die Zustände des $\text{DCF}^{(n)}$ -Prozesses mittels harmonischer Analyse in ihre Anteile bezüglich der Charaktere-Basis und berechnen die Evolution dieser Anteile über die Eigenwerte von $\mathbb{M}^{(n)}$.

Zunächst formen wir $P^{(n)}(X^{(n)}(k_n) \in nC)$ um. Sei also $C \in \mathcal{C}$ fest; wir setzen $f_n := \mathbf{1}_{nC}$. Da $\mu_0^{(n)}$ $\pi_S^{(n)}$ -stetig ist, sei weiter $g_0^{(n)}(\gamma) = \mu_0^{(n)}(\gamma)/\pi_S^{(n)}(\gamma)$ die Dichte von $\mu_0^{(n)}$ bezüglich $\pi_S^{(n)}$ (vgl. [2], S. 55). Dann gilt mit $P^{(n)X^{(n)}(k_n)} = \mu_0^{(n)}\mathbb{M}^{(n)k_n}$

$$\begin{aligned}
 P^{(n)}(X^{(n)}(k_n) \in nC) &= \sum_{\gamma \in \mathcal{P}_n} P^{(n)}(X^{(n)}(k_n) = \gamma) f_n(\gamma) \\
 &= \sum_{\gamma \in \mathcal{P}_n} \mu_0^{(n)} \mathbb{M}^{(n)k_n}(\gamma) f_n(\gamma) \\
 &= \sum_{\gamma, \gamma' \in \mathcal{P}_n} g_0^{(n)}(\gamma') \pi_S^{(n)}(\gamma') \mathbb{M}^{(n)k_n}(\gamma', \gamma) f_n(\gamma) \\
 &= \sum_{\gamma' \in \mathcal{P}_n} g_0^{(n)}(\gamma') \pi_S^{(n)}(\gamma') \mathbb{M}^{(n)k_n} f_n(\gamma') \\
 &= \langle g_0^{(n)}, \mathbb{M}^{(n)k_n} f_n \rangle.
 \end{aligned} \tag{6.22}$$

$g_0^{(n)}$ lässt sich in der Basis $(\chi_\lambda)_{\lambda \in \mathcal{P}_n}$ entwickeln zu $g_0^{(n)} = \sum_{\lambda \in \mathcal{P}_n} \langle g_0^{(n)}, \chi_\lambda \rangle \chi_\lambda$. Dies setzen wir in die obige Formel ein und erhalten

$$\begin{aligned} P^{(n)}(X^{(n)}(k_n) \in nC) &= \sum_{\lambda \in \mathcal{P}_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle \chi_\lambda, \mathbb{M}^{(n)k_n} f_n \rangle \\ &= \sum_{\lambda \in \mathcal{P}_n} \theta_\lambda^{(n)k_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle f_n, \chi_\lambda \rangle, \end{aligned} \quad (6.23)$$

denn es gilt mit Korollar 6.3

$$\begin{aligned} \langle \chi_\lambda, \mathbb{M}^{(n)k_n} f_n \rangle &= \sum_{\gamma \in \mathcal{P}_n} \pi_S^{(n)}(\gamma) \chi_\lambda(\gamma) \mathbb{M}^{(n)k_n} f_n(\gamma) \\ &= \sum_{\gamma, \gamma' \in \mathcal{P}_n} \pi_S^{(n)}(\gamma) \chi_\lambda(\gamma) \mathbb{M}^{(n)k_n}(\gamma, \gamma') f_n(\gamma') \\ &= \sum_{\gamma' \in \mathcal{P}_n} \theta_\lambda^{(n)k_n} \pi_S^{(n)}(\gamma') \chi_\lambda(\gamma') f_n(\gamma') \\ &= \theta_\lambda^{(n)k_n} \langle f_n, \chi_\lambda \rangle. \end{aligned} \quad (6.24)$$

Als Nächstes gilt es, den Term der Gleichgewichtsverteilung entsprechend umzuformen. $\pi_S^{(n)}$ entspricht gerade der Eigenfunktion $\pi_S^{(n)} \chi$ von $\mathbb{M}^{(n)}$ mit $\chi \equiv 1$. Dies ist der Charakter $\chi_{(n,0,\dots)}$ der zur trivialen Partition $(n, 0, \dots)$ gehörenden irreduziblen Darstellung der S_n ; der Eigenwert $\theta_{(n,0,\dots)}^{(n)}$ ist ebenfalls gleich 1 (vgl. auch Gleichung (6.18)). Wir können also schreiben

$$\begin{aligned} \pi_S^{(n)}(nC) &= \sum_{\gamma \in \mathcal{P}_n} f_n(\gamma) \pi_S^{(n)}(\gamma) \chi_{(n,0,\dots)}(\gamma) \\ &= \langle f_n, \chi_{(n,0,\dots)} \rangle \\ &= \theta_{(n,0,\dots)}^{(n)k_n} \langle f_n, \chi_{(n,0,\dots)} \rangle \\ &= \theta_{(n,0,\dots)}^{(n)k_n} \langle g_0^{(n)}, \chi_{(n,0,\dots)} \rangle \langle f_n, \chi_{(n,0,\dots)} \rangle, \end{aligned} \quad (6.25)$$

wobei für die letzte Zeile $1 = \sum_{\lambda \in \mathcal{P}_n} \mu_0^{(n)}(\lambda) = \sum_{\lambda \in \mathcal{P}_n} g_0^{(n)}(\lambda) \pi_S^{(n)}(\lambda) = \langle g_0^{(n)}, 1 \rangle$ verwendet wurde.

Damit entspricht die Abweichung $\Delta_C^{(n)}(k_n)$ gerade der Summe der Anteile des

Prozesses auf allen anderen Basis-Charakteren, d. h.

$$\Delta_C^{(n)}(k_n) = \sum_{\substack{\lambda \in \mathcal{P}_n \\ \lambda \neq (n, 0, \dots)}} \theta_\lambda^{(n)k_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle f_n, \chi_\lambda \rangle. \quad (6.26)$$

Diese Summe können wir zunächst vereinfachen: Da wir uns auf Startpartitionen γ beschränkt haben, die über höchstens n^β Abschnitte verfügen, gilt $\mu_0^{(n)}(\gamma) = 0$ und damit $g_0^{(n)}(\gamma) = 0$ für alle $\gamma \in \mathcal{P}_n$ mit $N_\gamma > n^\beta$. Für $N_\gamma \leq n^\beta$ liefert die Folgerung (6.20) zur Murnaghan-Nakayama-Formel $\chi_\lambda(\gamma) = 0$ für alle $\lambda \in \mathcal{P}_n$ mit $B_\lambda > n^\beta$. Damit reduziert sich die Summation auf

$$\Delta_C^{(n)}(k_n) = \sum_{\substack{\lambda \in \mathcal{P}_n \\ \lambda \neq (n, 0, \dots), B_\lambda \leq n^\beta}} \theta_\lambda^{(n)k_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle f_n, \chi_\lambda \rangle.$$

Der Beweis gelingt nun, indem wir die verbleibende Summationsmenge so aufteilen, dass die entsprechenden Basisvektoren entweder orthogonal zur Funktion f_n sind (vgl. Lemma 6.6) oder aber der Eigenwert für großes n hinreichend klein wird (Lemma 6.5). Damit lässt sich zeigen, dass die gesamte Summe für $n \rightarrow \infty$ verschwindet.

Wir wählen $\eta \in (0, 1)$ mit $1 - (\beta' - \beta) < \eta < 1$ und setzen $n_0 := 5^{1/(1-\eta)}$. Für $n \geq n_0$ zerlegen wir die Summationsmenge wie folgt:

$$\Delta_C^{(n)}(k_n) = \left(\sum_{\lambda \in \mathcal{P}'_n} + \sum_{\lambda \in \mathcal{P}''_n} + \sum_{\lambda \in \mathcal{P}'''_n} \right) \theta_\lambda^{(n)k_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle f_n, \chi_\lambda \rangle \quad (6.27)$$

mit

$$\begin{aligned} \mathcal{P}'_n &= \mathcal{P}'_n(\eta, \beta) := \{ \lambda \in \mathcal{P}_n : B_\lambda \leq n^\beta, \lambda_1 \leq n - 2n^\eta, N_\lambda \leq n - 2n^\eta \}, \\ \mathcal{P}''_n &= \mathcal{P}''_n(\eta, \beta) := \{ \lambda \in \mathcal{P}_n : B_\lambda \leq n^\beta, n - 2n^\eta < \lambda_1 < n \}, \\ \mathcal{P}'''_n &= \mathcal{P}'''_n(\eta, \beta) := \{ \lambda \in \mathcal{P}_n : B_\lambda \leq n^\beta, n - 2n^\eta < N_\lambda \}. \end{aligned}$$

Nach Wahl von n_0 sind diese drei Mengen paarweise disjunkt, und ihre Vereinigung ergibt gerade die gesamte Summationsmenge:

$$\mathcal{P}'_n \cup \mathcal{P}''_n \cup \mathcal{P}'''_n = \{ \lambda \in \mathcal{P}_n : \lambda \neq (n, 0, \dots), B_\lambda \leq n^\beta \}.$$

Beachte hierzu, dass für $\lambda \in \mathcal{P}_n'''$

$$\lambda_1 \leq n - N_\lambda + 1 < n - n + 2n^\eta + 1 = 2n^\eta + 1$$

und für $\lambda \in \mathcal{P}_n''$

$$\lambda_1 > n - 2n^\eta$$

gilt. Nach Wahl von n_0 gilt aber für $n \geq n_0$

$$\begin{aligned} 1 &= 4\left(5^{\frac{1}{1-\eta}}\right)^{\eta-1} + \frac{1}{5} \geq 4n^{\eta-1} + \frac{1}{n} \\ \Rightarrow \quad n - 2n^\eta &\geq 2n^\eta + 1, \end{aligned}$$

woraus insbesondere folgt, dass \mathcal{P}_n'' und \mathcal{P}_n''' disjunkt sind. Weiter folgt für $\lambda \in \mathcal{P}_n'''$ auch

$$\lambda_1 < 2n^\eta + 1 \leq n - 2n^\eta < n$$

und damit $(n, 0, \dots) \notin \mathcal{P}_n'''$.

Die Menge \mathcal{P}_n' ist so gewählt, dass λ_1 und N_λ gleichmäßig beschränkt sind. In diesem Fall können wir den zugehörigen Eigenwert $\theta_\lambda^{(n)}$ geeignet abschätzen:

6.5 Lemma. *Für alle $\lambda \in \mathcal{P}_n$ gilt $|\theta_\lambda^{(n)}| \leq \frac{\lambda_1 \vee N_\lambda}{n}$. Insbesondere gibt es ein $n_1 = n_1(\eta) \in \mathbb{N}$, so dass für $\lambda \in \mathcal{P}_n'$ gilt*

$$|\theta_\lambda^{(n)}| \leq e^{-n^{\eta-1}} \quad \forall n \geq n_1. \quad (6.28)$$

Die Mengen \mathcal{P}_n'' und \mathcal{P}_n''' enthalten die restlichen Partitionen: Dies sind solche, wo die Masse im zugehörigen Young-Diagramm entweder auf die erste Zeile konzentriert ist (\mathcal{P}_n'') oder auf die erste Spalte (\mathcal{P}_n'''). In beiden Fällen lässt sich zeigen, dass der zugehörige Charakter für große n orthogonal zu f_n ist:

6.6 Lemma. *Es gibt ein $n_2 = n_2(\beta, C) \in \mathbb{N}$, so dass für $\lambda \in \mathcal{P}_n'' \cup \mathcal{P}_n'''$ gilt*

$$\langle f_n, \chi_\lambda \rangle = 0 \quad \forall n \geq n_2. \quad (6.29)$$

Beweis von Lemma 6.5. Wir wollen den Eigenwert $\theta_\lambda^{(n)}$ des Charakters χ_λ abschätzen.

Aus der Formel (6.17) für $\theta_\lambda^{(n)}$ erhalten wir für beliebiges $\lambda \in \mathcal{P}_n$

$$\begin{aligned}\theta_\lambda^{(n)} &= \frac{1}{n(n-1)} \left(\sum_{i=1}^n \lambda_i^2 - \sum_{j=1}^{\lambda_1} \lambda_j'^2 \right) \\ &\leq \frac{1}{n(n-1)} \left(\lambda_1 \sum_{i=1}^n \lambda_i - \sum_{j=1}^{\lambda_1} 1 \right) \\ &\leq \frac{1}{n(n-1)} (\lambda_1 n - \lambda_1) = \frac{\lambda_1}{n}.\end{aligned}$$

Ersetzen wir in der ersten Zeile λ durch die transponierte Partition λ' , so ändert sich nur das Vorzeichen auf der linken Seite. Wir erhalten also $-\theta_\lambda^{(n)} = \theta_{\lambda'}^{(n)} \leq \lambda'_1/n = N_\lambda/n$, und damit $|\theta_\lambda^{(n)}| \leq \frac{\lambda_1 \vee N_\lambda}{n}$.

Für $\lambda \in \mathcal{P}'_n$ können wir dieses Ergebnis weiter verschärfen: Wähle $n_1 \in \mathbb{N}$ derart, dass $(1 - \frac{2}{n^{1-\eta}})^{n^{1-\eta}} \leq e^{-1}$ gilt für alle $n \geq n_1$. Wegen $\lambda_1 \vee N_\lambda \leq n - 2n^\eta$ folgt sodann

$$|\theta_\lambda^{(n)}| \leq 1 - \frac{2}{n^{1-\eta}} = \left(1 - \frac{2}{n^{1-\eta}}\right)^{n^{1-\eta} \cdot n^{\eta-1}} \leq e^{-n^{\eta-1}}$$

für alle $n \geq n_1$. □

Beweis von Lemma 6.6. Um die Aussage zu zeigen, gehen wir sinngemäß wie folgt vor: Wir nutzen aus, dass der Wert der Indikatorfunktion $f_n = \mathbf{1}_{nC}$ an der Stelle $\lambda \in \mathcal{P}_n$, also $f_n(\lambda)$, nach Eigenschaften von $C \in \mathcal{C}$ nur vom k -Anfang einer Partition λ abhängt (für ein bestimmtes $k \in \mathbb{N}$). Modifizieren wir also f_n so, dass der k -Anfang ignoriert wird und bezeichnen diese Funktion mit f_n^{k*} , so folgt $f_n^{k*} \equiv 1$. Dies entspricht jedoch gerade dem Charakter der trivialen Partition, und da die Charaktere orthonormal sind, folgt schon $\langle f_n^{k*}, \chi_{\lambda^*} \rangle = 0$ für alle anderen Charaktere.

Da f_n verändert wurde, muss der Charakter χ_λ ebenfalls passend modifiziert werden, was durch das χ_{λ^*} angedeutet ist. Hierzu werden wir die Murnaghan-Nakayama-Formel 6.4 verwenden, deren entsprechende Anwendbarkeit durch die spezielle Form von $\lambda \in \mathcal{P}''_n$ bzw. $\lambda \in \mathcal{P}'''_n$ sowie die von $C \in \mathcal{C}$ sichergestellt ist.

Für die eigentliche Beweisführung beginnen wir mit $\lambda \in \mathcal{P}''_n$. Sei also $C = C_{\mathbf{a}, \mathbf{b}}$ mit $(\mathbf{a}, \mathbf{b}) \in I_k$ für ein $k \in \mathbb{N}$ und sei $\delta := \delta_{\mathbf{a}, \mathbf{b}}$ (vgl. Gleichung (6.5)). Wie bereits geschildert, lässt sich nun $\gamma \in nC_{\mathbf{a}, \mathbf{b}}$ darstellen als $\gamma = (\gamma', \gamma'')$ mit $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$ und

$\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}$. Wir erhalten also

$$\begin{aligned} \langle f_n, \chi_\lambda \rangle &= \sum_{\gamma \in \mathcal{P}_n} \mathbf{1}_{nC}(\gamma) \chi_\lambda(\gamma) \pi_S^{(n)}(\gamma) \\ &= \sum_{\gamma \in nC_{\mathbf{a}, \mathbf{b}}} \chi_\lambda(\gamma) \pi_S^{(n)}(\gamma) \\ &= \sum_{\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}} \sum_{\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}} \chi_\lambda(\gamma', \gamma'') \pi_S^{(n)}(\gamma', \gamma''). \end{aligned} \quad (6.30)$$

Beachte, dass wegen $n - |\gamma'|_1 > \delta n$ (vgl. S. 100) die innere Summe stets nicht leer ist.

Wir wollen nun zeigen, dass für jedes $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$ die innere Summe für große n verschwindet. Wir nutzen dabei aus, dass gemäß Definition von \mathcal{P}_n'' die Masse von λ im Wesentlichen auf λ_1 konzentriert ist: Mit hinreichend großem $n_2'' = n_2''(\beta, \eta, C)$ gilt $\lambda_1 > n - 2n'' > n - \delta n$ für alle $\lambda \in \mathcal{P}_n''$ mit $n \geq n_2''$, woraus gleichzeitig folgt, dass die Gesamtmasse aller anderen Abschnitte höchstens $n - \lambda_1 < \delta n$ beträgt.

Andererseits gilt für $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$ nach Konstruktion $\gamma'_i \geq \gamma'_k > \delta n$ für $i = 1, \dots, k$ (vgl. S. 100). Wenn wir nun die Murnaghan-Nakayama-Formel k -mal auf $\chi_\lambda(\gamma', \gamma'')$ anwenden, indem wir sukzessive Randsegmente der Längen $\gamma'_1, \dots, \gamma'_k$ von λ abtrennen, ist damit sichergestellt, dass wir in jedem Reduktionsschritt nur genau *ein* Randsegment mit der vorgeschriebenen Länge vorfinden, nämlich die letzten γ'_i Kästchen von λ_1 . Insgesamt werden so $|\gamma'|_1$ Kästchen abgezogen, was wegen $|\gamma'|_1 < n - \delta n < \lambda_1$ immer möglich ist. Wir erhalten somit

$$\chi_\lambda(\gamma', \gamma'') = \chi_{\lambda^*}(\gamma'')$$

mit $\lambda^* \in \mathcal{P}_{n-|\gamma'|_1}$ definiert durch $\lambda_1^* := \lambda_1 - |\gamma'|_1$ und $\lambda_j^* := \lambda_j$ für $j \geq 2$.

Nun modifizieren wir noch den Faktor $\pi_S^{(n)}(\gamma', \gamma'')$: Gemäß Definition von $\pi_S^{(n)}$ und wegen $\gamma'_k > \gamma''_1$ können wir

$$\frac{\pi_S^{(n)}(\gamma', \gamma'')}{\pi_S^{(n-|\gamma'|_1)}(\gamma'')} = \frac{1}{\prod_{i=1}^k \gamma'_i \prod_{j=1}^k N_{\gamma'}(j)!} =: R(\gamma') < \infty.$$

setzen; damit lässt sich die innere Summe aus (6.30) wie folgt schreiben:

$$\begin{aligned} \sum_{\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}} \pi_S^{(n)}(\gamma', \gamma'') \chi_\lambda(\gamma', \gamma'') &= R(\gamma') \sum_{\gamma'' \in \mathcal{P}_{n-|\gamma'|_1}} \pi_S^{(n-|\gamma'|_1)}(\gamma'') \chi_{\lambda^*}(\gamma'') \\ &= R(\gamma') \langle \chi_{\lambda^*}, 1 \rangle_{n-|\gamma'|_1}. \end{aligned} \quad (6.31)$$

Wegen $\lambda \neq (n, 0, \dots)$ ist $\lambda^* \neq (n - |\gamma'|_1, 0, \dots)$, und wir erhalten somit aus der Orthonormalität der Charaktere $\langle \chi_{\lambda^*}, 1 \rangle_{n-|\gamma'|_1} = 0$. Da $\gamma' \in G_{\mathbf{a}, \mathbf{b}}$ beliebig war, ist damit gezeigt, dass die innere Summe verschwindet.

Ähnliches Vorgehen liefert dasselbe Ergebnis für $\lambda \in \mathcal{P}_n'''$. Wir nutzen aus, dass jetzt die Masse von λ überwiegend in der ersten Spalte von $[\lambda]$ konzentriert ist: Mit hinreichend großem $n_2''' = n_2'''(\beta, \eta, C)$ gilt $N_\lambda > n - 2n^\eta > n - \delta n + 1$ für alle $\lambda \in \mathcal{P}_n'''$ mit $n \geq n_2'''$. Die Gesamtmasse außerhalb der ersten Spalte beträgt damit höchstens $n - N_\lambda < \delta n - 1$.

Wenden wir nun für ein $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$ die Murnaghan-Nakayama-Formel wie oben an, so gibt es zu jedem γ'_i wiederum nur genau ein Randsegment der passenden Länge in λ , und zwar sind dies jetzt die untersten γ'_i Kästchen in der ersten Spalte. Auch hier ist wegen $|\gamma'|_1 = n - |\gamma''|_1 < n - \delta n < N_\lambda$ sichergestellt, dass es insgesamt ausreichend Abschnitte mit der Länge 1 in λ gibt. Wir erhalten also entsprechend

$$\chi_\lambda(\gamma', \gamma'') = \chi_{\lambda^{**}}(\gamma'')$$

mit $\lambda^{**} \in \mathcal{P}_{n-|\gamma'|_1}$ definiert durch $\lambda_j^{**} := \lambda_j$ für $1 \leq j \leq N_\lambda - |\gamma'|_1$, $\lambda_j^{**} := 0$ sonst.

Wir müssen nun nur noch nachweisen, dass $\langle \chi_{\lambda^{**}}, 1 \rangle = 0$ gilt, λ^{**} also nicht die triviale Partition $(n - |\gamma'|_1, 0, \dots)$ ist. Wäre dies der Fall, so müsste $N_\lambda = |\gamma'|_1 + 1$ gelten, was wegen

$$n - \delta n + 1 < N_\lambda = |\gamma'|_1 + 1 < n - \delta n + 1$$

auf einen Widerspruch führt. λ^{**} ist also nicht die triviale Partition, und die innere Summe verschwindet für alle $\gamma' \in G_{\mathbf{a}, \mathbf{b}}^{(n)}$.

Setzen wir schließlich $n_2 := n_2'' \vee n_2'''$, so folgt das gewünschte Ergebnis. \square

Nun können wir Gleichung (6.27) neu aufschreiben: Wegen Lemma 6.6 fallen für

großes n die Summationen über \mathcal{P}_n'' und \mathcal{P}_n''' weg, und wir erhalten

$$\Delta_C^{(n)}(k_n) = \sum_{\lambda \in \mathcal{P}_n'} \theta_\lambda^{(n)k_n} \langle g_0^{(n)}, \chi_\lambda \rangle \langle f_n, \chi_\lambda \rangle. \quad (6.32)$$

Der jetzt folgenden Anwendung des ersten Lemmas kommt eine Schlüsselstellung zu: Für den mit k_n potenzierten Eigenwert $\theta_\lambda^{(n)k_n}$ erhalten wir aufgrund der Voraussetzungen an $(k_n)_{n \geq 1}$ eine Abschätzung, die stark genug ist, um zu zeigen, dass trotz der wachsenden Summandenzahl $|\mathcal{P}_n'|$ und des Terms $\langle g_0^{(n)}, \chi_\lambda \rangle$ die verbleibende Summe für $n \rightarrow \infty$ verschwindet. Damit ist dann die gewünschte Konvergenz gezeigt.

Unter Verwendung von Lemma 6.5 erhalten wir zunächst

$$\begin{aligned} |\Delta_C^{(n)}(k_n)| &\leq \sum_{\lambda \in \mathcal{P}_n'} |\theta_\lambda^{(n)k_n}| |\langle g_0^{(n)}, \chi_\lambda \rangle| |\langle f_n, \chi_\lambda \rangle| \\ &\leq \sum_{\lambda \in \mathcal{P}_n'} (e^{-n^{\eta-1}})^{k_n} |\langle g_0^{(n)}, \chi_\lambda \rangle| |\langle f_n, \chi_\lambda \rangle|. \end{aligned}$$

Da nach Voraussetzung $k_n \geq n^{\beta'}$ gilt, folgt weiter

$$\begin{aligned} |\Delta_C^{(n)}(k_n)| &\leq \sum_{\lambda \in \mathcal{P}_n'} (e^{-n^{\eta-1}})^{n^{\beta'}} |\langle g_0^{(n)}, \chi_\lambda \rangle| |\langle f_n, \chi_\lambda \rangle| \\ &= \sum_{\lambda \in \mathcal{P}_n'} e^{-n^{\beta' + \eta - 1}} |\langle g_0^{(n)}, \chi_\lambda \rangle| |\langle f_n, \chi_\lambda \rangle|. \end{aligned} \quad (6.33)$$

Nun schätzen wir die Anzahl der Summanden ab. Dazu überlegen wir uns, dass das Young-Diagramm $[\lambda]$ eines beliebigen $\lambda \in \mathcal{P}_n$ mit $B_\lambda = s$ aus einem quadratischen $s \times s$ Block von Kästchen besteht, an den pro Reihe und Zeile jeweils höchstens $n - 1$ Kästchen angefügt werden. Um also die Anzahl der Partitionen $\lambda \in \mathcal{P}_n$ mit gegebener Diagonalenlänge sehr grob abzuschätzen, erhalten wir

$$\#\{\lambda \in \mathcal{P}_n : B_\lambda = s\} \leq \underbrace{n \cdot \dots \cdot n}_{2s \text{ Faktoren}} = n^{2s}$$

und weiter vereinfacht mit $t > 0$

$$\#\{\lambda \in \mathcal{P}_n : B_\lambda \leq t\} \leq tn^{2t}.$$

Damit können wir $|\mathcal{P}'_n|$ wie folgt abschätzen:

$$|\mathcal{P}'_n| \leq \#\{\lambda \in \mathcal{P}_n : B_\lambda \leq n^\beta\} \leq n^\beta n^{2n^\beta} = e^{(\beta+2n^\beta)\log n} \leq e^{3n^\beta \log n}.$$

Schließlich fehlen noch die beiden letzten Terme aus (6.33). Für den einen erhalten wir aus der Cauchy-Schwarzschen Ungleichung

$$\langle f_n, \chi_\lambda \rangle \leq \underbrace{\langle f_n, f_n \rangle}_{\leq 1} \underbrace{\langle \chi_\lambda, \chi_\lambda \rangle}_{=1} \leq 1, \quad (6.34)$$

sowie für den anderen

$$\begin{aligned} |\langle g_0^{(n)}, \chi_\lambda \rangle| &\leq \sum_{\gamma \in \mathcal{P}_n} |g_0^{(n)}(\gamma) \chi_\lambda(\gamma) \pi_S^{(n)}(\gamma)| \\ &= \sum_{\gamma \in \mathcal{P}_n} |\mu_0^{(n)}(\gamma) \chi_\lambda(\gamma)| \\ &\leq \sup_{N_\gamma \leq n^\beta} |\chi_\lambda(\gamma)| \end{aligned}$$

wegen $\mu_0^{(n)} \in \mathcal{P}_{n,\beta}$.

Jetzt gilt es nur noch, den Charakter $\chi_\lambda(\gamma)$ geeignet abzuschätzen (für $\lambda, \gamma \in \mathcal{P}_n$ und $N_\gamma \leq n^\beta$). Dies gelingt abermals unter Verwendung der Murnaghan-Nakayama-Formel 6.4, mithilfe derer wir $\chi_\lambda(\gamma)$ als Summe von Charakteren reduzierter Partitionen $\chi_{\lambda^*}(\hat{\gamma})$ schreiben können. Da wir Partitionen aus \mathcal{P}_n betrachten, umfasst die Summe höchstens n Summanden. Diese zerlegen wir wiederum, was so lange fortgesetzt wird, bis als reduzierte Partitionen nur noch die leere Partition \emptyset übrigbleibt; wegen $N_\gamma \leq n^\beta$ ist dies nach höchstens n^β Schritten der Fall. Wir erhalten also

$$|\chi_\lambda(\gamma)| \leq \sum_{i=1}^n |\chi_{\lambda^*}(\hat{\gamma})| \leq \dots \leq n^{n^\beta} \cdot \underbrace{|\chi_{\emptyset}(\emptyset)|}_{=1} = e^{n^\beta \log n},$$

was schließlich

$$|\langle g_0^{(n)}, \chi_\lambda \rangle| \leq \sup_{N_\gamma \leq n^\beta} |\chi_\lambda(\gamma)| \leq e^{n^\beta \log n} \quad \text{für alle } \lambda \in \mathcal{P}'_n \quad (6.35)$$

liefert.

Nachdem nun alle Bestandteile von Gleichung (6.33) geeignet abgeschätzt sind, müssen wir die Ergebnisse nur noch passend zusammensetzen. Es gilt

$$\begin{aligned}
 |\Delta_C^{(n)}(k_n)| &\leq \sum_{\lambda \in \mathcal{P}'_n} \theta_\lambda^{(n)k_n} |\langle g_0^{(n)}, \chi_\lambda \rangle| |\langle f_n, \chi_\lambda \rangle| \\
 &\leq e^{3n^\beta \log n} \cdot e^{-n^{\beta'} + \eta^{-1}} \cdot e^{n^\beta \log n} \cdot 1 \\
 &= e^{-n^\beta (n^{(\beta' - \beta) - (1 - \eta)} - 4 \log n)}.
 \end{aligned}$$

Die Summandenanzahl und der Term $\langle g_0^{(n)}, \chi_\lambda \rangle$ sind zusammen von der Größenordnung $n^{4n^\beta} = e^{4n^\beta \log n}$. Dies kann jedoch durch $\theta_\lambda^{(n)k_n}$ kompensiert werden: Wegen $(\beta' - \beta) - (1 - \eta) > 0$ gilt $n^{(\beta' - \beta) - (1 - \eta)} - 4 \log n \geq \frac{1}{2}$ für hinreichend große n , und wir erhalten schließlich

$$|\Delta_C^{(n)}(k_n)| \leq e^{-\frac{n^\beta}{2}} \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Damit ist der Beweis des Theorems abgeschlossen. □

7 Beweis der Vermutung

Nun stehen alle Hilfsmittel bereit, um Vermutung 1.1 zu beweisen, die wir hier noch einmal neu formulieren wollen (vgl. [7], S. 931):

7.1 Theorem. *Ist $\mu \in \mathcal{M}_1(\Omega_1)$ CCF-invariant, so ist μ die Poisson-Dirichlet-Verteilung \mathcal{PD}_1 .*

Beweis. Um die Eindeutigkeit zu zeigen, müssen wir

$$\mu(A) = \mathcal{PD}_1(A)$$

für alle $A \in \mathcal{B}_{\Omega_1}$ nachweisen. Eine solche Gleichheit der Maße können wir zunächst nur auf dem kleineren Mengensystem $\mathcal{C} \subset \mathcal{B}_{\Omega_1}$ zeigen, da nur dort schnelle DCF⁽ⁿ⁾-Konvergenz vorliegt:

7.2 Lemma. *Ist $\mu \in \mathcal{M}_1(\Omega_1)$ CCF-invariant, so gilt*

$$\mu(C) = \mathcal{PD}_1(C) \quad \text{für alle } C \in \mathcal{C}. \quad (7.1)$$

Um zu zeigen, dass daraus schon die Gleichheit auf allen Mengen folgt, benötigen wir noch ein weiteres Lemma:

7.3 Lemma. *Sei $\Omega'_1 := \{p \in \Omega_1 : \exists \text{ unendlich viele } n \in \mathbb{N} \text{ mit } p_n > \sum_{j>n} p_j\}$. Dann gilt:*

$$(i) \quad \mathcal{PD}_1(\Omega'_1) = 1.$$

(ii) *Durch seine Werte auf $\{C \cap \Omega'_1 : C \in \mathcal{C}\} \subset \mathcal{B}_{\Omega'_1}$ ist ein Maß bereits auf ganz $(\Omega'_1, \mathcal{B}_{\Omega'_1})$ festgelegt.*

$$(iii) \quad \mu(\Omega'_1) = 1.$$

Aus (i) und (iii) folgt sodann, dass sich die Aussage von Lemma 7.2 auch schreiben lässt als

$$\mu(C \cap \Omega'_1) = \mathcal{PD}_1(C \cap \Omega'_1) \quad \text{für alle } C \in \mathcal{C},$$

woraus mit (ii) die Gleichheit der Maße auf $\mathcal{B}_{\Omega'_1}$ folgt. Da sowohl μ als auch \mathcal{PD}_1 keine Masse außerhalb von Ω'_1 besitzen, folgt damit schon die Behauptung des Theorems. \square

Wir bringen nun die Beweise der beiden Lemmata.

Beweis von Lemma 7.3. Zum Beweis von (i) sei daran erinnert, dass die Poisson-Dirichlet-Verteilung durch nichtaufsteigendes Anordnen aus der *GEM*-Verteilung gewonnen werden kann (vgl. Satz 3.14). Sei $(X_n)_{n \geq 1}$ eine *GEM*(1)-verteilte Sequenz, also $X_n := U_n \prod_{i=1}^{n-1} (1 - U_i)$ für $n \geq 1$ mit $U_i \sim U[0, 1]$ (vgl. Definition 3.12). Wir setzen zusätzlich

$$Y_n := \prod_{i=1}^{n-1} (1 - U_i) \quad \text{für } n \geq 1. \quad (7.2)$$

Beachte, dass dann gilt

$$Y_n = \prod_{i=1}^{n-2} (1 - U_i) - \underbrace{U_{n-1} \prod_{i=1}^{n-2} (1 - U_i)}_{=X_{n-1}} = \dots = 1 - \sum_{j < n} X_j = \sum_{j \geq n} X_j$$

sowie $X_n = U_n Y_n$. Wir zeigen nun, dass jede Realisierung von $(X_n)_{n \geq 1}$ nach nichtaufsteigender Anordnung stets auf ein $p \in \Omega'_1$ führt, was die Behauptung beweist.

Hierfür definieren wir Stoppzeiten $(\tau_k)_{k \geq 1}$ mit $\tau_1 := 1$ und

$$\tau_{k+1} := \min \{n > \tau_k : X_n \wedge (Y_n - X_n) < X_j \forall j \leq \tau_k\} \quad \text{für } k \geq 1. \quad (7.3)$$

Interpretieren wir $(X_n)_{n \geq 1}$ wieder als Realisierung des gleichmäßigen Abbrechprozesses (vgl. S. 26), so ist zum Zeitpunkt τ_k mindestens eins der entstehenden Bruchstücke X_n bzw. $Y_n - X_n$ kleiner als alle zuvor abgebrochenen Stücke. Wegen $\sum_{n \geq 1} X_n = 1$ gilt $X_n \searrow 0$ f. s. für $n \rightarrow \infty$ und es folgt, dass jedes τ_k fast sicher endlich ist.

Wir werden nun zeigen:

$$A_k := \left\{ U_{\tau_k} > \frac{1}{2} \right\} = \left\{ X_{\tau_k} > Y_{\tau_k} - X_{\tau_k} \right\}, \quad k \in \mathbb{N} \quad \text{sind unabhängig.} \quad (7.4)$$

$$P(A_k) = \frac{1}{2} \quad \forall k \geq 1. \quad (7.5)$$

Hieraus folgt sofort $U_{\tau_k} > \frac{1}{2}$ unendlich oft f. s. Mit diesem Wissen lässt sich leicht nachweisen, dass die nichtaufsteigende Anordnung der $(X_n)_{n \geq 1}$ eine Partition aus Ω'_1 liefert: Die τ_k mit $U_{\tau_k} > \frac{1}{2}$ führen jeweils gerade zu einem Abschnitt in $p \in \Omega'_1$ mit der gewünschten Eigenschaft. Auf jedem A_k gilt nach Wahl von τ_k

$$\sum_{j > \tau_k} X_j = Y_{\tau_k+1} = Y_{\tau_k} - X_{\tau_k} < X_i \quad \forall i \leq \tau_k,$$

woraus insbesondere folgt, dass bei der nichtaufsteigenden Sortierung von $(X_n)_{n \geq 1}$ Elemente mit Index aus $\{1, \dots, \tau_k\}$ nicht mit Elementen mit Index aus $\{\tau_k + 1, \dots\}$ vertauscht werden. Damit gilt insgesamt

$$p_{\tau_k} = \min_{i \leq \tau_k} X_i > \sum_{j > \tau_k} X_j = \sum_{j > \tau_k} p_j.$$

Um die Unabhängigkeit der A_k sowie $P(A_k) = \frac{1}{2}$ zu beweisen, stellen wir für $n \geq 1$ die Teilungsvariable U_n wie folgt dar:

$$U_n = \begin{cases} V_n, & \text{falls } \eta_n = 1, \\ 1 - V_n, & \text{falls } \eta_n = 0, \end{cases} \quad (7.6)$$

wobei $V_n \sim U[0, 1]$ und $\eta_n \sim B(1, \frac{1}{2})$ voneinander unabhängige Zufallsvariablen sind. Damit können wir das Ereignis A_k schreiben als

$$A_k = (B_k \cap C_k) \cup (B_k^c \cap C_k^c)$$

mit $B_k := \{V_{\tau_k} > \frac{1}{2}\}$ und $C_k := \{\eta_{\tau_k} = 1\}$.

Setzen wir nun $\mathcal{F}_n := \sigma(V_1, \dots, V_n, \eta_0, \dots, \eta_{n-1})$ (mit $\eta_0 := 1$), so sind die oben definierten $(\tau_k)_{k \geq 1}$ Stoppzeiten bezüglich der Filtration $(\mathcal{F}_n)_{n \geq 1}$. Nach Konstruktion

gilt dann $B_k \in \mathcal{F}_{\tau_k}$, während C_k unabhängig von \mathcal{F}_{τ_k} ist für alle $k \in \mathbb{N}$. Wir erhalten damit für beliebiges $D \in \mathcal{F}_{\tau_k}$

$$\begin{aligned} P(D \cap A_k) &= P((D \cap B_k) \cap C_k) + P((D \cap B_k^c) \cap C_k^c) \\ &= P(D \cap B_k) \cdot P(C_k) + P(D \cap B_k^c) \cdot P(C_k^c) \\ &= \frac{1}{2} (P(D \cap B_k) + P(D \cap B_k^c)) \\ &= \frac{1}{2} P(D) \end{aligned}$$

unter Verwendung von $P(C_k) = P(C_k^c) = \frac{1}{2}$. Wählen wir nun $D := \Omega_1$, so erhalten wir $P(A_k) = \frac{1}{2}$, was die zweite Behauptung zeigt.

Für den Beweis der ersten Behauptung sei $k \geq 1$ und $D := \bigcap_{j \in J} A_j$ mit $J \subset \{1, \dots, k-1\}$. Wegen $A_j \in \mathcal{F}_{\tau_{j+1}} \subset \mathcal{F}_{\tau_k}$ für alle $j < k$ gilt $D \in \mathcal{F}_{\tau_k}$, und wir erhalten

$$P\left(\bigcap_{j \in J} A_j \cap A_k\right) = P(A_k) \cdot P\left(\bigcap_{j \in J} A_j\right).$$

Da k beliebig war, lässt sich dies induktiv fortsetzen zu

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \quad \text{für beliebige endliche } I \subset \mathbb{N},$$

womit die Unabhängigkeit gezeigt ist.

Um (ii) zu zeigen, müssen wir nachweisen, dass $\mathcal{C}' := \{C \cap \Omega'_1 : C \in \mathcal{C}\}$ ein \cap -stabiles Erzeugendensystem von $\mathcal{B}_{\Omega'_1}$ ist (vgl. [2], S. 10).

Dazu machen wir einen Umweg über die Topologie auf Ω'_1 (vgl. Kapitel 2.1): $\mathcal{B}_{\Omega'_1}$ ist die von der Topologie auf Ω'_1 erzeugte σ -Algebra, d. h. die kleinste σ -Algebra, die alle offenen Mengen umfasst. Die Topologie von Ω'_1 wiederum ist die von der l_1 -Norm induzierte Topologie \mathfrak{T}_1 , die aus allen Vereinigungen von offenen l_1 -Kugeln besteht (vgl. etwa [19], S. 22). Wir sagen daher auch, die Topologie wird von den offenen l_1 -Kugeln erzeugt, und schreiben $\mathfrak{T}_1 = \mathfrak{T}(B_1)$.

Wir zeigen nun, dass \mathfrak{T}_1 schon von \mathcal{C}' erzeugt wird, d. h. $\mathfrak{T}_1 = \mathfrak{T}(\mathcal{C}')$. Für die Inklusion $\mathfrak{T}_1 \subset \mathfrak{T}(\mathcal{C}')$ müssen wir nachweisen, dass es für $p \in \Omega'_1$ zu jeder offenen l_1 -Kugel $B_\varepsilon(p)$ mit $\varepsilon > 0$ ein $C \in \mathcal{C}$ gibt derart, dass $p \in C \cap \Omega'_1 \subset B_\varepsilon(p)$ gilt

(vgl. [19], S. 23). Hierbei wird sich als entscheidend erweisen, dass aus „ $p \in C$ für ein $C \in \mathcal{C}$ “ nicht nur Aussagen über den Anfang von p folgen, sondern sich dadurch auch der Partitionsrest abschätzen lässt (vgl. S. 91 ff.).

Sei also $\varepsilon > 0$ und $p \in \Omega'_1$. Aufgrund der speziellen Form von p können wir ein geeignetes $C = C_{\mathbf{a}, \mathbf{b}}$ konstruieren, so dass p in $C_{\mathbf{a}, \mathbf{b}}$ enthalten ist. Dazu wählen wir ein hinreichend großes $k \in \mathbb{N}$, so dass $q := \sum_{j>k} p_j < \min(p_k, \frac{\varepsilon}{5})$ gilt. Dies ist nach Wahl von Ω'_1 immer möglich, ebenso gilt stets $q > 0$. Wir setzen dann

$$\delta := \frac{q \wedge (p_k - q)}{k + 2}$$

sowie

$$\left. \begin{array}{l} a_i := p_i - \delta \\ b_i := p_i + \delta \end{array} \right\} \quad \text{für } i = 1, \dots, k$$

und behaupten, dass das so konstruierte $(\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1}^k$ in I_k liegt. In der Tat gilt aufgrund der Wahl von δ und wegen $q > 0$

$$\sum_{i \leq k} b_i = \sum_{i \leq k} (p_i + \delta) \leq 1 - q + \frac{k}{k+2} q < 1$$

sowie

$$\begin{aligned} a_k + \sum_{i \leq k} a_i &= (p_k - \delta) + \sum_{i \leq k} (p_i - \delta) \\ &= (p_k + 1 - q) - (k+1)\delta \\ &\geq 1 + \underbrace{(p_k - q)}_{>0} \left(1 - \frac{k+1}{k+2}\right) > 1, \end{aligned}$$

und es folgt $(\mathbf{a}, \mathbf{b}) \in I_k$ (vgl. Gleichung (6.2)).

Für beliebiges $x \in C_{\mathbf{a}, \mathbf{b}}$ gilt nun

$$\sum_{j>k} x_j = 1 - \sum_{j=1}^k x_j \leq 1 - \sum_{j=1}^k (p_j - \delta)$$

$$\begin{aligned}
 &= 1 - \sum_{j=1}^k p_j + k\delta \\
 &\leq q + \frac{k}{k+2}q < 2q
 \end{aligned}$$

und damit

$$\begin{aligned}
 |x - p|_1 &\leq \sum_{j=1}^k |x_j - p_j| + \sum_{j>k} p_j + \sum_{j>k} x_j \\
 &\leq \delta \cdot 2k + q + 2q \\
 &\leq \left(3 + 2\frac{k}{k+2}\right) q < \varepsilon,
 \end{aligned}$$

was zeigt, dass zu jeder offenen l_1 -Kugel $B_\varepsilon(p)$ von p in Ω'_1 ein solches $C \in \mathcal{C}$ existiert, so dass $p \in C \cap \Omega'_1 \subset B_\varepsilon(p)$ gilt. Umgekehrt gibt es für jedes $p \in C \cap \Omega'_1$ mit $C \in \mathcal{C}$ ein $\varepsilon > 0$, so dass $B_\varepsilon(p) \subset C \cap \Omega'_1$ gilt. Damit ist auch $\mathfrak{T}_1 \supset \mathfrak{T}(\mathcal{C}')$ gezeigt, und es gilt insgesamt $\mathfrak{T}_1 = \mathfrak{T}(\mathcal{C}')$.

Offenbar können wir genauso argumentieren, wenn wir zusätzlich verlangen, dass die Endpunkte des Intervalls $(\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1}^k$ alle rational sind. Somit lässt sich zeigen, dass die Topologie von Ω'_1 sogar schon von

$$\mathcal{C}'_{\mathbb{Q}} := \{C \cap \Omega'_1 : C = C_{\mathbf{a}, \mathbf{b}} \text{ mit } (\mathbf{a}, \mathbf{b}) \in I_k \text{ für ein } k \geq 1; \mathbf{a}, \mathbf{b} \in \mathbb{Q}^k\} \subset \mathcal{C}' \quad (7.7)$$

erzeugt wird, d. h. $\mathfrak{T}_1 = \mathfrak{T}(\mathcal{C}'_{\mathbb{Q}})$. Da $\mathcal{C}'_{\mathbb{Q}}$ nach Konstruktion abzählbar ist, ist jede Vereinigung von Mengen aus $\mathcal{C}'_{\mathbb{Q}}$ schon abzählbar, woraus insbesondere

$$\mathfrak{T}(\mathcal{C}'_{\mathbb{Q}}) \subset \sigma(\mathcal{C}'_{\mathbb{Q}})$$

folgt. Damit ergibt sich für die Borel'sche σ -Algebra

$$\mathcal{B}_{\Omega'_1} = \sigma(\mathfrak{T}_1) = \sigma(\mathfrak{T}(\mathcal{C}'_{\mathbb{Q}})) \subset \sigma(\mathcal{C}'_{\mathbb{Q}}) \subset \sigma(\mathcal{C}'),$$

und zusammen mit $\mathcal{C}' \subset \mathfrak{T}(\mathcal{C}') = \mathfrak{T}_1$ erhalten wir schließlich $\mathcal{B}_{\Omega'_1} = \sigma(\mathcal{C}')$.

Es fehlt nun noch der Nachweis, dass \mathcal{C} (und damit auch \mathcal{C}') abgeschlossen ist unter Schnitten. Seien also $(\mathbf{a}_1, \mathbf{b}_1) \in I_j$ und $(\mathbf{a}_2, \mathbf{b}_2) \in I_k$ für $j, k \in \mathbb{N}$ mit $j \leq k$.

Falls $j < k$, so setzen wir $a_{1i} := 0$ und $b_{1i} := 1$ für $i = j+1, \dots, k$. Definieren wir nun (\mathbf{a}, \mathbf{b}) durch $a_i := a_{1i} \vee a_{2i}$ und $b_i := b_{1i} \wedge b_{2i}$ für $i = 1, \dots, k$, so gilt offensichtlich $(\mathbf{a}, \mathbf{b}) \in I_k$ und $C_{\mathbf{a}_1, \mathbf{b}_1} \cap C_{\mathbf{a}_2, \mathbf{b}_2} = C_{\mathbf{a}, \mathbf{b}}$. Damit ist (ii) bewiesen.

Für den Beweis von (iii) sei zunächst angemerkt, dass die Gleichheit der Maße μ und \mathcal{PD}_1 , wie sie in Gleichung (7.1) formuliert ist, sogar schon für alle Mengen aus der von \mathcal{C} erzeugten σ -Algebra gilt, da – wie im Beweis von (ii) gezeigt – \mathcal{C} ein \cap -stabiles Erzeugendensystem ist (vgl. [2], S. 10). Wir haben also

$$\mu(A) = \mathcal{PD}_1(A) \quad \text{für alle } A \in \sigma(\mathcal{C})$$

Zeigen wir nun, dass Ω'_1 selbst schon in $\sigma(\mathcal{C})$ liegt, so folgt mit (i)

$$\mu(\Omega'_1) = \mathcal{PD}_1(\Omega'_1) = 1,$$

und die Behauptung ist bewiesen.

Für den Nachweis von $\Omega'_1 \in \sigma(\mathcal{C})$ stellen wir Ω'_1 wie folgt dar:

$$\begin{aligned} \Omega'_1 &= \bigcap_{n \geq 1} \left\{ p \in \Omega_1 : \exists n \text{ verschiedene } k \in \mathbb{N} \text{ mit } p_k > \sum_{i > k} p_i \right\} \\ &= \bigcap_{n \geq 1} \bigcup_{\substack{I \subset \mathbb{N} \\ |I|=n}} \bigcap_{k \in I} \left\{ p \in \Omega_1 : p_k > \sum_{i > k} p_i \right\}. \end{aligned}$$

Ersetzen wir hierin $\{p \in \Omega_1 : p_k > \sum_{i > k} p_i\}$ durch

$$A_k := \left\{ p \in \Omega_1 : p_k > \sum_{i > k} p_i, \sum_{i=1}^k p_i < 1 \right\} \quad (7.8)$$

für $k \in \mathbb{N}$, so wird die Menge auf der rechten Seite der Gleichung höchstens kleiner. Da jedoch für jedes $p \in \Omega'_1$ immer schon $\sum_{i=1}^k p_i < 1$ für alle $k \in \mathbb{N}$ gilt, herrscht sogar Gleichheit, und wir können Ω'_1 auch schreiben als

$$\Omega'_1 = \bigcap_{n \geq 1} \bigcup_{\substack{I \subset \mathbb{N} \\ |I|=n}} \bigcap_{k \in I} A_k. \quad (7.9)$$

Da die Vereinigung und beide Schnitte höchstens abzählbar sind (beachte dazu, dass insbesondere die Menge aller endlichen Teilmengen von \mathbb{N} abzählbar ist), folgt $\Omega'_1 \in \sigma(\mathcal{C})$, wenn wir zeigen, dass alle A_k in $\sigma(\mathcal{C})$ liegen. Dazu überlegen wir uns für ein beliebiges $k \in \mathbb{N}$, dass wir für ein $p \in A_k$ und $i = 1, \dots, k$ immer hinreichend große $a_i \in \mathbb{Q}$ mit $a_i < p_i$ wählen können, so dass

$$p_k > a_k > 1 - \sum_{i=1}^k a_i > 1 - \sum_{i=1}^k p_i = \sum_{i>k} p_i$$

gilt. Ebenso lassen sich für $i = 1, \dots, k$ hinreichend kleine $b_i \in \mathbb{Q}$ mit $b_i > p_i$ finden, so dass

$$\sum_{i=1}^k p_i < \sum_{i=1}^k b_i < 1$$

erfüllt ist. Nach Wahl der a_i und b_i liegt das Intervall $(\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1}^k$ in I_k .

Umgekehrt gilt für jedes $p \in \Omega_1$ mit $p_i \in (a_i, b_i)$ für ein $(\mathbf{a}, \mathbf{b}) \in I_k$ nach Wahl von I_k stets

$$p_k > a_k > 1 - \sum_{i=1}^k a_i > 1 - \sum_{i=1}^k p_i = \sum_{i>k} p_i \quad \text{sowie} \quad \sum_{i=1}^k p_i < \sum_{i=1}^k b_i < 1.$$

Somit haben wir

$$A_k = \bigcup_{\substack{(\mathbf{a}, \mathbf{b}) \in I_k \\ \mathbf{a}, \mathbf{b} \in \mathbb{Q}^k}} \{p \in \Omega_1 : p_i \in (a_i, b_i) \text{ für } i = 1, \dots, k\} \equiv \bigcup_{\substack{(\mathbf{a}, \mathbf{b}) \in I_k \\ \mathbf{a}, \mathbf{b} \in \mathbb{Q}^k}} C_{\mathbf{a}, \mathbf{b}}.$$

Da diese Vereinigung wiederum abzählbar ist, liegt mit den $C_{\mathbf{a}, \mathbf{b}}$ auch A_k in $\sigma(\mathcal{C})$. Da k beliebig war, gilt dies für alle $k \in \mathbb{N}$. \square

Beweis von Lemma 7.2. An dieser Stelle fließen nun unsere gesammelten Vorarbeiten ein. Seien dazu $\alpha, \beta, \gamma \in \mathbb{R}$ beliebig mit $\frac{2}{5} < \alpha < \beta < \gamma < \frac{1}{2}$. Da μ nach Annahme CCF-invariant ist, gilt die A-priori-Momentenabschätzung 2.1 und wir erhalten $\int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) < \infty$, d. h. μ legt nicht so viel Masse auf Partitionen mit vielen kleinen Abschnitten. Unter diesen Voraussetzungen können wir Theorem 4.1 zur Anwendung bringen und erhalten für jedes $n \in \mathbb{N}$ einen Wahrscheinlichkeitsraum $\Omega_{cde}^{(n)}$

mit Wahrscheinlichkeitsmaß $Q^{(n)} = Q_\mu^{(n)}$, auf dem ein Prozess definiert ist, aus dem wir eine CCF-Markov-Kette $(p(k))_{k \geq 0}$ mit (stationärer) Anfangsverteilung μ und eine DCF⁽ⁿ⁾-Markov-Kette $(\ell^{(n)}(k))_{k \geq 0}$ erhalten. Der diskrete Prozess $(\ell^{(n)}(k))_{k \geq 0}$ approximiert jeweils $(p(k))_{k \geq 0}$ derart, dass zwischen beiden Prozessen bis zum Zeitpunkt $\lfloor n^\gamma \rfloor$ Kopplung besteht. Überdies gilt $\lim_{n \rightarrow \infty} Q_\mu^{(n)}[N_{\ell^{(n)}(0)} \geq n^\beta] = 0$, wir haben also eine gewisse Kontrolle über die Anzahl der Abschnitte in der Startpartition von $(\ell^{(n)}(k))_{k \geq 0}$. Es lässt sich zeigen, dass dies genügt, um Theorem 6.1 anzuwenden; auf \mathcal{C} liegt somit schnelle Konvergenz des DCF⁽ⁿ⁾-Prozesses gegen $\pi_S^{(n)}(n \cdot)$ vor. Wenn wir schließlich noch $\pi_S^{(n)}(n \cdot) \xrightarrow{w} \mathcal{PD}_1$ für $n \rightarrow \infty$ verwenden (vgl. S. 5), lässt sich die Behauptung zeigen.

Um das beschriebene Vorgehen umzusetzen, sind allerdings noch einige technische Einschränkungen nötig: Die Konvergenzaussage über die DCF-CCF-Kopplung ist nicht dergestalt, dass damit die Gleichheit $\mu(C) = \mathcal{PD}_1(C)$ für $C \in \mathcal{C}$ direkt gezeigt werden kann – Gleichung (4.3) macht nur eine Aussage über die *erwartete* l_1 -Abweichung. Wir behelfen uns, indem wir die Menge C ein wenig vergrößern bzw. verkleinern: Mit $(\mathbf{a}, \mathbf{b}) \in I_k$ für ein $k \in \mathbb{N}$ gilt auch $((1 + \varepsilon)\mathbf{a}, (1 - \varepsilon)\mathbf{b}) \in I_k$ für alle ε in einer geeigneten Umgebung von 0. Gilt für (\mathbf{a}, \mathbf{b}) überdies $C := C_{\mathbf{a}, \mathbf{b}} \neq \emptyset$, so gilt auch $C^{(\varepsilon)} := C_{(1+\varepsilon)\mathbf{a}, (1-\varepsilon)\mathbf{b}} \neq \emptyset$ für alle ε in einer geeigneten Umgebung von 0. Für diese Mengen werden wir vermöge des Weges über Kopplung und DCF⁽ⁿ⁾-Konvergenz eine Ungleichung der Form

$$\mathcal{PD}_1(C^{(-\varepsilon)}) \geq \mu(C) \geq \mathcal{PD}_1(C^{(\varepsilon)}) \quad (7.10)$$

für alle hinreichend kleinen $\varepsilon > 0$ angeben können. Lassen wir hierin nun ε gegen null gehen (setze $\varepsilon := 1/n$ und $n \rightarrow \infty$), so bilden die $C^{(\varepsilon)}$ eine konvergente isotone Mengenfolge $C^{(\varepsilon)} \uparrow \overset{\circ}{C}$ (die $C^{(-\varepsilon)}$ bilden entsprechend eine antitone Mengenfolge $C^{(-\varepsilon)} \downarrow \bar{C}$). Wie man sich leicht überlegt, können wir ein nichtleeres $C = C_{\mathbf{a}, \mathbf{b}}$ mit $(\mathbf{a}, \mathbf{b}) = (a_i, b_i)_{i=1, \dots, k}$ auch schreiben als $C = C_{\tilde{\mathbf{a}}, \tilde{\mathbf{b}}}$ mit $\tilde{a}_i := \max_{i \leq j \leq k} a_j$ und $\tilde{b}_i := \min_{1 \leq j \leq i} b_j$ für $i = 1, \dots, k$. Damit gilt für den Rand von C offenbar

$$\partial C = \{p \in \Omega_1 \mid (p_i)_{i=1, \dots, k} = (\tilde{b}_i)_{i=1, \dots, k} \vee (p_i)_{i=1, \dots, k} = (\tilde{a}_i)_{i=1, \dots, k}\},$$

und es lässt sich unter Rückgriff auf die *GEM*-Verteilung (vgl. S. 30) zeigen, dass

∂C eine \mathcal{PD}_1 -Nullmenge ist. Somit erhalten wir (vgl. [2], S. 5)

$$\mathcal{PD}_1(C^{(\varepsilon)}) \rightarrow \mathcal{PD}_1(\overset{\circ}{C}) = \mathcal{PD}_1(C) \quad \text{bzw.} \quad \mathcal{PD}_1(C^{(-\varepsilon)}) \rightarrow \mathcal{PD}_1(\bar{C}) = \mathcal{PD}_1(C)$$

für $\varepsilon \searrow 0$, und Gleichung (7.10) liefert uns mit $\varepsilon \searrow 0$ die Behauptung.

Wir stellen nun die gewünschte Ungleichung (7.10) auf, zunächst den rechten Teil. Sei dazu $C = C_{\mathbf{a}, \mathbf{b}} \in \mathcal{C}$ nichtleer mit $(\mathbf{a}, \mathbf{b}) \in I_k$ für ein $k \geq 1$ und $\varepsilon > 0$ hinreichend klein, so dass $C^{(\pm\varepsilon)}$ ebenfalls nicht leer ist. Wir schreiben μ als $Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in \cdot)$ und stellen die Differenz zu \mathcal{PD}_1 mithilfe des approximierenden DCF $^{(n)}$ -Prozesses $(\ell^{(n)}(k))_{k \geq 0}$ sowie der DCF $^{(n)}$ -Gleichgewichtsverteilung $\pi_S^{(n)}(n \cdot)$ dar. Somit erhalten wir für beliebiges $n \in \mathbb{N}$ unter zusätzlicher Verwendung der umgekehrten Dreiecksungleichung

$$\begin{aligned} \mu(C) - \mathcal{PD}_1(C^{(\varepsilon)}) &= Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) - \mathcal{PD}_1(C^{(\varepsilon)}) \\ &\geq Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) - Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\right) \\ &\quad - \left| Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\right) - \pi_S^{(n)}(nC^{(\varepsilon)}) \right| \\ &\quad + \pi_S^{(n)}(nC^{(\varepsilon)}) - \mathcal{PD}_1(C^{(\varepsilon)}) \\ &=: D_1^{(\varepsilon)}(n) - D_2^{(\varepsilon)}(n) + D_3^{(\varepsilon)}(n). \end{aligned} \tag{7.11}$$

Wir zeigen nun, dass der Limes inferior der $D_1^{(\varepsilon)}(n)$ größer gleich null ist, während die anderen beiden Terme für $n \rightarrow \infty$ verschwinden.

Dazu schätzen wir $D_1^{(\varepsilon)}(n)$ unter Zuhilfenahme einer gleichmäßigen Schranke $\varepsilon' := \varepsilon \min\{a_i : 1 \leq i \leq k\} > 0$ ab und nutzen aus, dass zum Zeitpunkt $\lfloor n^\gamma \rfloor$ $(p(k))_{k \geq 0}$ und $(\ell^{(n)}(k))_{k \geq 0}$ noch gekoppelt sind:

$$\begin{aligned} D_1^{(\varepsilon)}(n) &= - \left(-Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) + Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\right) \right) \\ &\geq -Q^{(n)}\left(\left\{p(\lfloor n^\gamma \rfloor) \notin C\right\} \cap \left\{\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(\varepsilon)}\right\}\right) \\ &\geq -Q^{(n)}\left(\left|p(\lfloor n^\gamma \rfloor) - \frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor)\right|_1 > \varepsilon'\right). \end{aligned}$$

Hierauf wenden wir die Markov-Ungleichung (vgl. [2], S. 72) an, was

$$D_1^{(\varepsilon)}(n) \geq -\frac{1}{\varepsilon'} E_{Q^{(n)}} \left[\left| p(\lfloor n^\gamma \rfloor) - \frac{1}{n} \ell^{(n)}(\lfloor n^\gamma \rfloor) \right|_1 \right]$$

liefert. Gemäß Theorem 4.1 verschwindet die rechte Seite für $n \rightarrow \infty$ und es folgt $\liminf_{n \rightarrow \infty} D_1^{(\varepsilon)}(n) \geq 0$.

Um $D_2^{(\varepsilon)}(n)$ abzuschätzen, machen wir von der schnellen DCF⁽ⁿ⁾-Konvergenz Gebrauch. Das entsprechende Theorem 6.1 kann jedoch nur verwendet werden, wenn der Anfangszustand des Prozesses mit Sicherheit in $\mathcal{P}_{n,\beta}$ liegt, was für $(\ell^{(n)}(k))_{k \geq 0}$ wegen $\lim_{n \rightarrow \infty} Q^{(n)}[N_{\ell^{(n)}(0)} \geq n^\beta] = 0$ jedoch nur *annähernd* der Fall ist. In dieser Situation ist aber die Definition eines neuen Wahrscheinlichkeitsmaßes sinnvoll, das nach Konstruktion schon in $\mathcal{M}_1(\mathcal{P}_{n,\beta})$ liegt; wir setzen

$$\tilde{Q}^{(n)}(\cdot) := Q^{(n)}(\cdot | \ell^{(n)}(0) \in \mathcal{P}_{n,\beta}). \quad (7.12)$$

Offenbar ist $(\ell^{(n)}(k))_{k \geq 0}$ auch unter $\tilde{Q}^{(n)}$ eine DCF⁽ⁿ⁾-Kette.

Es gilt nun

$$\begin{aligned} 0 \leq D_2^{(\varepsilon)}(n) &\leq |\tilde{Q}^{(n)}(\ell^{(n)}(\lfloor n^\gamma \rfloor) \in nC^{(\varepsilon)}) - \pi_S^{(n)}(nC^{(\varepsilon)})| \\ &\quad + |\tilde{Q}^{(n)}(\ell^{(n)}(\lfloor n^\gamma \rfloor) \in nC^{(\varepsilon)}) - Q^{(n)}(\ell^{(n)}(\lfloor n^\gamma \rfloor) \in nC^{(\varepsilon)})| \\ &\leq |\tilde{Q}^{(n)}(\ell^{(n)}(\lfloor n^\gamma \rfloor) \in nC^{(\varepsilon)}) - \pi_S^{(n)}(nC^{(\varepsilon)})| + \|\tilde{Q}^{(n)} - Q^{(n)}\| \end{aligned} \quad (7.13)$$

für alle $n \in \mathbb{N}$, wobei der erste Term gemäß Theorem 6.1 (mit $\beta \equiv \beta$ und $k_n \equiv \lfloor n^\gamma \rfloor$) für $n \rightarrow \infty$ verschwindet.

Um zu zeigen, dass auch der zweite Term verschwindet, benötigen wir folgendes Lemma:

7.4 Lemma. *Sei (Ω, \mathcal{A}, Q) ein Wahrscheinlichkeitsraum und sei $B \in \mathcal{A}$ mit $Q(B) > 0$. Dann gilt für $Q'(\cdot) := Q(\cdot | B)$*

$$\|Q - Q'\| \leq Q(B^c). \quad (7.14)$$

Beweis. $Q - Q'$ stellt ein signiertes Maß auf (Ω, \mathcal{A}) dar, das sich nach der Hahn-Jordan-Zerlegung in einen positiven und einen negativen Teil zerlegen lässt (vgl. [3], S. 188).

Wir behaupten, dass $(Q - Q')(A) \geq 0$ für alle $A \subset B^c$ gilt sowie $(Q - Q')(A) \leq 0$ für alle $A \subset B$. Für den Nachweis sei $A \in \mathcal{A}$; wir erhalten

$$(Q - Q')(A) = Q(A) - \frac{Q(A \cap B)}{Q(B)} = \begin{cases} Q(A) - \frac{Q(A)}{Q(B)} \leq 0 & \text{falls } A \subset B, \\ Q(A) \geq 0 & \text{falls } A \subset B^c. \end{cases}$$

Somit gilt

$$Q - Q' = (Q - Q')(\cdot \cap B^c) - (Q' - Q)(\cdot \cap B),$$

und es folgt

$$\|Q - Q'\| = \max \{ (Q - Q')(B^c), (Q' - Q)(B) \}.$$

Wegen $(Q - Q')(B^c) = Q(B^c)$ sowie $(Q' - Q)(B) = 1 - Q(B) = Q(B^c)$ ist damit die Behauptung gezeigt. \square

Mit dem Lemma gilt sodann

$$\|\tilde{Q}^{(n)} - Q^{(n)}\| \leq Q^{(n)}(\ell^{(n)}(0) \notin \mathcal{P}_{n,\beta}),$$

und wir erhalten wegen $Q^{(n)}(\ell^{(n)}(0) \notin \mathcal{P}_{n,\beta}) = Q^{(n)}[N_{\ell^{(n)}(0)} \geq n^\beta] \rightarrow 0$ für $n \rightarrow \infty$ gemäß Theorem 4.1 schließlich

$$D_2^{(\varepsilon)}(n) \rightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Für $D_3^{(\varepsilon)}(n)$ nutzen wir $\pi_S^{(n)}(n \cdot) \xrightarrow{w} \mathcal{PD}_1$ für $n \rightarrow \infty$ aus (vgl. S. 5). Wegen $\mathcal{PD}_1(\partial C^{(\varepsilon)}) = 0$ (vgl. S. 119) folgt mit dem Portmanteau-Theorem (vgl. [2], S. 186)

$$\pi_S^{(n)}(nC^{(\varepsilon)}) \rightarrow \mathcal{PD}_1(C^{(\varepsilon)}) \quad \text{für } n \rightarrow \infty,$$

und wir erhalten $\lim_{n \rightarrow \infty} D_3^{(\varepsilon)}(n) = 0$. Damit ergibt sich insgesamt

$$\mu(C) - \mathcal{PD}_1(C^{(\varepsilon)}) \geq \liminf_{n \rightarrow \infty} D_1^{(\varepsilon)}(n) - \lim_{n \rightarrow \infty} D_2^{(\varepsilon)}(n) + \lim_{n \rightarrow \infty} D_3^{(\varepsilon)}(n) \geq 0.$$

Für die umgekehrte Ungleichung $\mathcal{PD}_1(C^{(-\varepsilon)}) \geq \mu(C)$ erhalten wir analog

$$\begin{aligned}
\mathcal{PD}_1(C^{(-\varepsilon)}) - \mu(C) &= \mathcal{PD}_1(C^{(-\varepsilon)}) - Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) \\
&\geq \mathcal{PD}_1(C^{(-\varepsilon)}) - \pi_S^{(n)}(nC^{(-\varepsilon)}) \\
&\quad - \left| \pi_S^{(n)}(nC^{(-\varepsilon)}) - Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(-\varepsilon)}\right) \right| \\
&\quad + Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(-\varepsilon)}\right) - Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) \\
&\equiv -D_3^{(-\varepsilon)}(n) - D_2^{(-\varepsilon)}(n) - D_1^{(-\varepsilon)}(n). \tag{7.15}
\end{aligned}$$

Für die Terme $D_2^{(-\varepsilon)}(n)$ und $D_3^{(-\varepsilon)}(n)$ gelten dieselben Überlegungen wie für $D_2^{(\varepsilon)}(n)$ und $D_3^{(\varepsilon)}(n)$. Für $D_1^{(-\varepsilon)}(n)$ erhalten wir mit $\varepsilon'' := \varepsilon \min\{b_i : 1 \leq i \leq k\} > 0$ ähnlich wie oben

$$\begin{aligned}
D_1^{(-\varepsilon)}(n) &= Q^{(n)}(p(\lfloor n^\gamma \rfloor) \in C) - Q^{(n)}\left(\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \in C^{(-\varepsilon)}\right) \\
&\leq Q^{(n)}\left(\{p(\lfloor n^\gamma \rfloor) \in C\} \cap \left\{\frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor) \notin C^{(-\varepsilon)}\right\}\right) \\
&\leq Q^{(n)}\left(\left|p(\lfloor n^\gamma \rfloor) - \frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor)\right|_1 > \varepsilon''\right) \\
&\leq \frac{1}{\varepsilon''} E_{Q^{(n)}}\left[\left|p(\lfloor n^\gamma \rfloor) - \frac{1}{n}\ell^{(n)}(\lfloor n^\gamma \rfloor)\right|_1\right]. \tag{7.16}
\end{aligned}$$

Für $n \rightarrow \infty$ verschwindet wiederum die rechte Seite der Gleichung, und wir erhalten $\limsup_{n \rightarrow \infty} D_1^{(-\varepsilon)}(n) \leq 0$. Somit gilt

$$\mathcal{PD}_1(C^{(-\varepsilon)}) - \mu(C) \geq -\lim_{n \rightarrow \infty} D_3^{(-\varepsilon)}(n) - \lim_{n \rightarrow \infty} D_2^{(-\varepsilon)}(n) - \limsup_{n \rightarrow \infty} D_1^{(-\varepsilon)}(n) \geq 0,$$

und insgesamt folgt

$$\mathcal{PD}_1(C^{(-\varepsilon)}) \geq \mu(C) \geq \mathcal{PD}_1(C^{(\varepsilon)}),$$

womit Lemma 7.2 bewiesen ist. \square

A Anhang

A.1 Beweis der A-priori-Momentenabschätzung

Wir greifen noch einmal die Proposition 2.1 auf, um sie anschließend zu beweisen (vgl. [7], S. 935 f.):

A.1 Proposition. *Sei $\mu \in \mathcal{M}_1(\Omega_1)$ CCF-invariant. Dann gilt für alle $\alpha > 2/5$*

$$\int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) < \infty. \quad (\text{A.1})$$

Beweis. Um den CCF-Übergang in einer Partition $p \in \Omega_1$ darzustellen, zerlegen wir $(0, 1]$ in disjunkte Intervalle $J_n := (2^{-n-1}, 2^{-n}]$, $n \geq 0$, und definieren für $n \geq 1$ Zufallsvariablen $W_n: \Omega_1 \rightarrow \mathbb{R}$ vermöge

$$W_n(p) := \sum_{i \geq 1} p_i \mathbf{1}_{p_i > 2^{-n}}. \quad (\text{A.2})$$

Ein solches W_n gibt die Gesamtmasse aller Abschnitte mit einer Länge $> 2^{-n}$ in p an. Ist nun $\mu \in \mathcal{M}_1(\Omega_1)$ CCF-invariant, so gilt

$$\int_{\Omega_1} \widehat{\mathbb{M}}(p, \cdot) \mu(dp) = \mu(\cdot),$$

wobei $\widehat{\mathbb{M}}$ den CCF-Übergangskern $\widehat{\mathbb{M}}: \Omega_1 \times \mathcal{B}_{\Omega_1} \rightarrow [0, 1]$ bezeichnet. Übertragen auf die Zufallsvariablen W_n liefert dies

$$\begin{aligned} \int_{\Omega_1} W_n(p) \mu(dp) &= \int_{\Omega_1} W_n(p') \mu(dp') \\ &= \int_{\Omega_1} W_n(p') \int_{\Omega_1} \widehat{\mathbb{M}}(p, dp') \mu(dp) \\ &= \int_{\Omega_1} \int_{\Omega_1} W_n(p') \widehat{\mathbb{M}}(p, dp') \mu(dp). \end{aligned}$$

Dies können wir wegen $W_n(p) = \int_{\Omega_1} W_n(p) \widehat{\mathbb{M}}(p, dp')$ auch schreiben als

$$0 = \int_{\Omega_1} W_n(p) \mu(dp) - \int_{\Omega_1} \int_{\Omega_1} W_n(p') \widehat{\mathbb{M}}(p, dp') \mu(dp)$$

$$= \int_{\Omega_1} \int_{\Omega_1} W_n(p) - W_n(p') \widehat{\mathbb{M}}(p, dp') \mu(dp). \quad (\text{A.3})$$

Die Differenz $W_n(p) - W_n(p')$ zerlegen wir nun in einen positiven und einen negativen Teil:

$$W_n(p) - W_n(p') := \Delta_+(p, p') - \Delta_-(p, p') \quad \text{mit } \Delta_+, \Delta_- \geq 0. \quad (\text{A.4})$$

Damit können wir Gleichung (A.3) weiter umformen zu

$$\begin{aligned} 0 &= \int_{\Omega_1} \int_{\Omega_1} \Delta_+(p, p') - \Delta_-(p, p') \widehat{\mathbb{M}}(p, dp') \mu(dp) \\ &= \int_{\Omega_1} \left(\int_{\Omega_1} \Delta_+(p, p') \widehat{\mathbb{M}}(p, dp') - \int_{\Omega_1} \Delta_-(p, p') \widehat{\mathbb{M}}(p, dp') \right) \mu(dp). \end{aligned}$$

Setzen wir nun noch $E\Delta_{\pm}(p) := \int_{\Omega_1} \Delta_{\pm}(p, p') \widehat{\mathbb{M}}(p, dp')$, so haben wir schließlich

$$\int_{\Omega_1} E\Delta_+(p) - E\Delta_-(p) \mu(dp) = 0. \quad (\text{A.5})$$

Um damit die gewünschte Abschätzung (A.1) zeigen zu können, müssen wir $E\Delta_+$ und $E\Delta_-$ berechnen bzw. abschätzen. Dazu überlegen wir uns für ein festes $n \in \mathbb{N}$, wie sich ein CCF-Übergang auf die Zufallsvariable W_n auswirkt. Bei einem Merge-Übergang kann W_n höchstens wachsen. Dies ist der Fall, wenn ein neuer Abschnitt mit einer Länge $> 2^{-n}$ entsteht, und mindestens einer der beiden Ausgangsabschnitte über eine Länge $\leq 2^{-n}$ verfügte. Andernfalls bleibt W_n konstant. Bei einem Split-Übergang hingegen kann W_n höchstens abnehmen: Wird ein Abschnitt mit einer Länge $> 2^{-n}$ aufgeteilt, so verringert sich W_n um die Länge eines bzw. beider Bruchstücke, so diese Werte kleiner gleich 2^{-n} sind.

Damit folgt, dass sich Δ_+ nur aus Merge-Übergängen speisen kann, während Δ_- nur von Split-Übergängen herrühren kann. Wir können daher Δ_+ nach unten abschätzen, indem wir alle Zuwächse, die sich aus der Vereinigung von Abschnitten $p_i \in J_n$ mit Abschnitten $p_j > 2^{-n-1}$ ergibt, berechnen; andere Übergänge, die ggf. auch zu einem Zuwachs von W_n führen, bleiben unberücksichtigt. Bei den genannten Übergängen entsteht ein neuer Abschnitt mit einer Länge $p_i + p_j > 2^{-n}$, und W_n wächst mindestens um p_i (nach Annahme wurde p_j unter Umständen schon zuvor erfasst). Die Wahrscheinlichkeit eines solchen Merge-Ereignisses beträgt gerade $p_i p_j$.

Damit ergibt sich

$$\begin{aligned}
E\Delta_+(p) &= \int_{\Omega_1} \Delta_+(p, p') \widehat{\mathbb{M}}(p, dp') \\
&\geq \sum_{i \neq j} p_i \mathbf{1}_{p_i \in J_n} \mathbf{1}_{p_j > 2^{-n-1}} \cdot p_i p_j \\
&= \left(\sum_{i \geq 1} p_i^2 \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \right) - \sum_{i \geq 1} p_i^3 \mathbf{1}_{p_i \in J_n} \\
&\geq 2^{-2n-2} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \right) - 2^{-2n} \sum_{i \geq 1} p_i \mathbf{1}_{p_i \in J_n}. \quad (\text{A.6})
\end{aligned}$$

Δ_- hingegen können wir direkt berechnen: Wird ein Abschnitt $p_i > 2^{-n}$ aufgeteilt, so entstehen daraus die Bruchstücke Xp_i und $(1-X)p_i$ mit $X \sim U[0, 1]$. W_n verringert sich genau dann, wenn mindestens einer dieser beiden neuen Abschnitte kleiner gleich 2^{-n} ist, und zwar genau um die Länge der entsprechenden Abschnitte. Die Wahrscheinlichkeit für das Aufteilen von p_i ist gerade p_i^2 , es gilt also

$$\begin{aligned}
E\Delta_-(p) &= \int_{\Omega_1} \Delta_-(p, p') \widehat{\mathbb{M}}(p, dp') \\
&= \sum_{i \geq 1} \left(\int_0^1 x p_i \mathbf{1}_{x p_i \leq 2^{-n} < p_i} \cdot p_i^2 dx + \int_0^1 (1-x) p_i \mathbf{1}_{(1-x)p_i \leq 2^{-n} < p_i} \cdot p_i^2 dx \right).
\end{aligned}$$

Mit $\int_0^1 x f(x) dx = \int_0^1 (1-x) f(1-x) dx$ vereinfacht sich dies zu

$$\begin{aligned}
E\Delta_-(p) &= 2 \sum_{i \geq 1} \int_0^1 x p_i \mathbf{1}_{x p_i \leq 2^{-n} < p_i} \cdot p_i^2 dx \\
&= 2 \sum_{i \geq 1} p_i^3 \mathbf{1}_{2^{-n} < p_i} \int_0^1 x \mathbf{1}_{x \leq 2^{-n}/p_i} dx \\
&= 2 \sum_{i \geq 1} p_i^3 \mathbf{1}_{2^{-n} < p_i} \left[\frac{x^2}{2} \right]_0^{2^{-n}/p_i} = 2^{-2n} \sum_{i \geq 1} p_i \mathbf{1}_{2^{-n} < p_i}. \quad (\text{A.7})
\end{aligned}$$

Für $E\Delta_+(p) - E\Delta_-(p)$ erhalten wir somit unter Ausnutzung von $\sum_i p_i \mathbf{1}_{p_i \in J_n} + \sum_j p_j \mathbf{1}_{p_j > 2^{-n}} = \sum_i p_i \mathbf{1}_{p_i > 2^{-n-1}}$

$$\begin{aligned}
E\Delta_+(p) - E\Delta_-(p) &\geq 2^{-2n} \left(\frac{1}{4} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} \mathbf{1}_{p_j > 2^{-n-1}} \right) - \sum_{i \geq 1} p_i \mathbf{1}_{2^{-n-1} < p_i} \right) \\
&\geq 2^{-2n} \left(\frac{1}{4} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} \mathbf{1}_{p_j > 2^{-n-1}} \right) - 1 \right). \tag{A.8}
\end{aligned}$$

Für die letzte Ungleichung wurde $\sum_{i \geq 1} p_i \mathbf{1}_{p_i > 2^{-n-1}} \leq \sum_{i \geq 1} p_i = 1$ benutzt.

Bei Integration über Ω_1 erhalten wir daraus mit Gleichung (A.5)

$$0 \geq \int_{\Omega_1} 2^{-2n} \left(\frac{1}{4} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} \mathbf{1}_{p_j > 2^{-n-1}} \right) - 1 \right) \mu(dp),$$

was sich unter Verwendung von $\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} > 2^{-n-1} \sum_{j \geq 1} \mathbf{1}_{p_j \in J_n}$ umformen lässt zu

$$\begin{aligned}
4 &\geq \int_{\Omega_1} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \right) \mu(dp) \tag{A.9} \\
&\geq 2^{-n-1} \int_{\Omega_1} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right)^2 \mu(dp) \\
&\geq 2^{-n-1} \left(\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) \right)^2.
\end{aligned}$$

Im letzten Schritt wurde die Jensen'sche Ungleichung $\phi(EX) \leq E\phi(X)$ für $\phi(X) = X^2$ konvex (vgl. [2], S. 73) verwendet. Das Ergebnis können wir weiter umformulieren zu

$$\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) \leq 2\sqrt{2^{n+1}} = 2^{\frac{n}{2} + \frac{3}{2}}.$$

Da $n \geq 1$ beliebig war, gilt also

$$\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) = O(2^{\frac{n}{2}}) \quad \text{für } n \rightarrow \infty.$$

Dies liefert schließlich mit monotoner Konvergenz

$$\int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) = \int_{\Omega_1} \sum_{i \geq 1} \sum_{n \geq 0} p_i^\alpha \mathbf{1}_{p_i \in J_n} \mu(dp)$$

$$\begin{aligned}
&\leq \sum_{n \geq 0} 2^{\alpha n} \int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) \\
&\leq c \sum_{n \geq 0} 2^{n(\frac{1}{2} - \alpha)} < \infty
\end{aligned} \tag{A.10}$$

für jedes $\alpha > 1/2$ und für ein geeignetes $c > 0$, was die Proposition für alle $\alpha > 1/2$ beweist.

Dieses Ergebnis werden wir nun wie gefordert auf alle $2/5 < \alpha < 1$ ausdehnen. Aus Gleichung (A.9) erhalten wir für ein beliebiges $0 < \beta < 1$

$$\begin{aligned}
4 &\geq \int_{\Omega_1} \left(\sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \right) \left(\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \right) \mathbf{1}_{\{\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} > 2^{-n\beta}\}} \mu(dp) \\
&\geq 2^{-n\beta} \int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mathbf{1}_{\{\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} > 2^{-n\beta}\}} \mu(dp) \\
&= 2^{-n\beta} \left(\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) - \int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mathbf{1}_{\{\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \leq 2^{-n\beta}\}} \mu(dp) \right)
\end{aligned}$$

unter Verwendung von $\mathbf{1}_A = 1 - \mathbf{1}_{A^c}$. Wegen

$$2^{-n\beta} \geq \sum_j p_j \mathbf{1}_{p_j > 2^{-n-1}} \geq 2^{-n-1} \sum_j \mathbf{1}_{p_j > 2^{-n-1}} \geq 2^{-n-1} \sum_j \mathbf{1}_{p_j \in J_n}$$

lässt sich dies weiter umformulieren zu

$$\begin{aligned}
4 &\geq 2^{-n\beta} \left(\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) - \frac{2^{-n\beta}}{2^{-n-1}} \mu \left[\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \leq 2^{-n\beta} \right] \right) \\
&\geq 2^{-n\beta} \left(\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) - 2^{-n\beta+n+1} \mu[p_j \leq 2^{-n\beta} \forall j \geq 1] \right).
\end{aligned} \tag{A.11}$$

Im letzten Schritt wurde benutzt, dass aus $\sum_{j \geq 1} p_j \mathbf{1}_{p_j > 2^{-n-1}} \leq 2^{-n\beta}$ schon $p_j \leq 2^{-n\beta}$ für alle $j \geq 1$ folgt.

Um die Wahrscheinlichkeit für das Ereignis $\{p_j \leq 2^{-n\beta} \forall j \geq 1\}$ abzuschätzen, können wir Gleichung (A.10) verwenden und erhalten

$$\infty > \int_{\Omega_1} \sum_{i \geq 1} p_i^\alpha \mu(dp) \geq \int_{\Omega_1} \left(\sum_{i \geq 1} p_i^\alpha \right) \mathbf{1}_{\{p_i \leq 2^{-n\beta} \forall i \geq 1\}} \mu(dp) \tag{A.12}$$

für alle $1/2 < \alpha < 1$ und $n \geq 1$. Auf dem Ereignis $\{p_i \leq 2^{-n\beta} \forall i \geq 1\}$ gilt unter Verwendung der Jensen'schen Ungleichung $\phi(EX) \leq E\phi(X)$ für $\phi(X) = X^{\alpha-1}$ konvex (vgl. [2], S. 73)

$$\sum_{i \geq 1} p_i^\alpha = \sum_{i \geq 1} p_i \cdot p_i^{\alpha-1} \geq \left(\sum_{i \geq 1} p_i \cdot p_i \right)^{\alpha-1} \geq \left(2^{-n\beta} \sum_{i \geq 1} p_i \right)^{\alpha-1} \geq 2^{n\beta(1-\alpha)},$$

und wir erhalten aus Gleichung (A.12)

$$\infty > 2^{n\beta(1-\alpha)} \mu[p_i \leq 2^{-n\beta} \forall i \geq 1] \quad \text{für alle } n \geq 1.$$

Somit gilt für alle $1/2 < \alpha < 1$

$$\mu[p_i \leq 2^{-n\beta} \forall i \geq 1] = O(2^{-n\beta(1-\alpha)}) \quad \text{für } n \rightarrow \infty.$$

Formen wir nun Gleichung (A.11) passend um zu

$$\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) \leq 4 \cdot 2^{n\beta} + 2^{-n\beta+n+1} \mu[p_j \leq 2^{-n\beta} \forall j \geq 1]$$

und setzen das obige Ergebnis ein, so erhalten wir

$$\begin{aligned} \int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) &= O(2^{n\beta} + 2^{(-n\beta+n)-n\beta(1-\alpha)}) \\ &= O(2^{n \max\{\beta, 1-\beta(2-\alpha)\}}) \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Da dieses Ergebnis für alle $0 < \beta < 1$ und $1/2 < \alpha < 1$ gilt, können diese Parameter geeignet gewählt werden. Es lässt sich durch einfache Rechnung zeigen, dass $\beta \geq 1 - \beta(2 - \alpha)$ äquivalent ist zu $\beta \geq (3 - \alpha)^{-1}$. Wegen $\inf_{1/2 < \alpha < 1} (3 - \alpha)^{-1} = 2/5$ können wir daher für jedes $\beta > 2/5$ ein $\alpha \in (\frac{1}{2}, 1)$ wählen, so dass $\max\{\beta, 1 - \beta(2 - \alpha)\} \equiv \beta$ gilt. Somit haben wir für alle $\beta > 2/5$

$$\int_{\Omega_1} \sum_{i \geq 1} \mathbf{1}_{p_i \in J_n} \mu(dp) = O(2^{\beta n}) \quad \text{für } n \rightarrow \infty.$$

Daraus ergibt sich schließlich wie oben (vgl. Gleichung (A.10))

$$\begin{aligned} \int_{\Omega_1} \sum_{i \geq 1} p_i^{\tilde{\alpha}} \mu(dp) &= \int_{\Omega_1} \sum_{i \geq 1} \sum_{n \geq 0} p_i^{\tilde{\alpha}} \mathbf{1}_{p_i \in J_n} \mu(dp) \\ &\leq c \sum_{n \geq 0} 2^{n(\beta - \tilde{\alpha})} < \infty \end{aligned}$$

für alle $\tilde{\alpha} > \beta$ und geeignetes $c > 0$. Damit ist die Proposition für alle $2/5 < \tilde{\alpha} < 1$ bewiesen. \square

A.2 Charaktere und Darstellungen der S_{10}

Partition	d_ρ	$\chi_\rho(\tau)$	Partition	d_ρ	$\chi_\rho(\tau)$
(10)	1	1	(4,3,3)	210	14
(9,1)	9	7	(4,3,2,1)	768	0
(8,2)	35	21	(4,3,1,1,1)	525	-35
(8,1,1)	36	20	(4,2,2,2)	300	-20
(7,3)	75	35	(4,2,2,1,1)	567	-63
(7,2,1)	160	64	(4,2,1,1,1,1)	350	-70
(7,1,1,1)	84	28	(4,1,1,1,1,1,1)	84	-28
(6,4)	90	34	(3,3,3,1)	210	-14
(6,3,1)	315	91	(3,3,2,2)	252	-28
(6,2,2)	225	55	(3,3,2,1,1)	450	-70
(6,2,1,1)	350	70	(3,3,1,1,1,1)	225	-55
(6,1,1,1,1)	126	14	(3,2,2,2,1)	288	-64
(5,5)	42	14	(3,2,2,1,1,1)	315	-91
(5,4,1)	288	64	(3,2,1,1,1,1,1)	160	-64
(5,3,2)	450	70	(3,1,1,1,1,1,1,1)	36	-20
(5,3,1,1)	567	63	(2,2,2,2,2)	42	-14
(5,2,2,1)	525	35	(2,2,2,2,1,1)	90	-34
(5,2,1,1,1)	448	0	(2,2,2,1,1,1,1)	75	-35
(5,1,1,1,1,1)	126	-14	(2,2,1,1,1,1,1,1)	35	-21
(4,4,2)	252	28	(2,1,1,1,1,1,1,1,1)	9	-7
(4,4,1,1)	300	20	(1,1,1,1,1,1,1,1,1,1)	1	-7

Tabelle A.1: Irreduzible Darstellungen der S_{10} : Dimension d_ρ und Wert des Charakters $\chi_\rho(\tau)$ auf der Äquivalenzklasse der Transpositionen τ (aus [14], S. 354)

Literaturverzeichnis

- [1] ALSMEYER, G.: *Stochastische Prozesse*. 2. Auflage. Universität Münster, 2002 (Skripten zur Mathematischen Statistik Nr. 33)
- [2] ALSMEYER, G.: *Wahrscheinlichkeitstheorie*. 3. Auflage. Universität Münster, 2003 (Skripten zur Mathematischen Statistik Nr. 30)
- [3] ALT, H. W.: *Lineare Funktionalanalysis: eine anwendungsorientierte Einführung*. 3. Auflage. Berlin; Heidelberg; New York : Springer, 1999
- [4] BERTOIN, J.: *Random Fragmentation and Coagulation Processes*. 1. Auflage. Cambridge : Cambridge University Press, 2006 (Cambridge Studies in Advanced Mathematics Vol. 102)
- [5] BOSCH, S.: *Algebra*. 4. Auflage. Berlin; Heidelberg; New York : Springer, 2004
- [6] DIACONIS, P.: *Group Representations in Probability and Statistics*. Hayward : Institute of Mathematical Statistics, 1988 (Institute of Mathematical Statistics: Lecture Notes, Monograph Series Vol. 11)
- [7] DIACONIS, P. ; MAYER-WOLF, E. ; ZEITOUNI, O. ; ZERNER, M. P. W.: The Poisson-Dirichlet Law is the Unique Invariant Distribution for Uniform Split-Merge Transformations. In: *The Annals of Probability* 32 (2004), Nr. 1B, S. 915–938
- [8] DONNELLY, P. ; JOYCE, P.: Continuity and Weak Convergence of Ranked and Size-Biased Permutations on the Infinite Simplex. In: *Stochastic Processes and their Applications* 31 (1989), S. 89–103
- [9] FISCHER, G.: *Lineare Algebra*. 11. Auflage. Braunschweig; Wiesbaden : Friedr. Vieweg & Sohn, 1997

- [10] FLATTO, L. ; ODLYZKO, A. M. ; WALES, D. B.: Random Shuffles and Group Representations. In: *The Annals of Probability* 13 (1985), Nr. 1, S. 154–178
- [11] GELMAN, A.: *Bayesian Data Analysis*. 1. Auflage. London : Chapman and Hall, 1995
- [12] GNEDIN, A. ; KEROV, S.: A Characterization of GEM Distributions. In: *Combinatorics, Probability and Computing* 10 (2001), S. 213–217
- [13] HOLST, L.: *The Poisson-Dirichlet Distribution and its Relatives Revisited*. <http://www.math.kth.se/matstat/fofu/reports/PoiDir.pdf>, Abruf: 29.12.2006. Royal Institute of Technology, Stockholm, Sweden
- [14] JAMES, G. D. ; KERBER, A.: *The Representation Theory of the Symmetric Group*. Reading, Mass. : Addison-Wesley, 1981 (Encyclopedia of Mathematics and its Applications Vol. 16)
- [15] KINGMAN, J. F. C.: Random Discrete Distributions. In: *Journal of the Royal Statistical Society. Series B* 37 (1975), S. 1–15
- [16] KINGMAN, J. F. C.: The Population Structure Associated with the Ewens Sampling Formula. In: *Theoretical Population Biology* 11 (1977), S. 274–283
- [17] KINGMAN, J. F. C.: *Poisson Processes*. Oxford : Oxford Science Publications, 1993 (Oxford Studies in Probability Vol. 3)
- [18] PITMAN, J.: Poisson-Dirichlet and GEM Invariant Distributions for Split-and-Merge Transformations of an Interval Partition. In: *Combinatorics, Probability and Computing* 11 (2002), S. 501–514
- [19] QUERENBURG, B. v.: *Mengentheoretische Topologie*. 3. Auflage. Berlin; Heidelberg; New York : Springer, 2001
- [20] VERSHIK, A. M. ; SHMIDT, A. A.: Limit Measures Arising in the Asymptotic Theory of Symmetric Groups. I. In: *Theory of Probability and its Applications* 22 (1977), Nr. 1, S. 70–85