

Kombinatorische Markov-Ketten

Rudolf Grübel

Received: date / Accepted: date

Zusammenfassung Wir behandeln ein aktuelles Teilgebiet der probabilistischen Kombinatorik, die Verwendung von Markov-Ketten bei der Analyse von kombinatorischen Familien. Solche Ketten sind zufällige dynamische Strukturen, die beispielsweise als Output von sequentiellen Algorithmen mit zufälligem Input entstehen. Wir erklären, wie Methoden der diskreten Potentialtheorie über Kompaktifizierungen des Zustandsraums zu starken Grenzwertsätzen und strukturellen Ergebnissen bei zufälligen Permutationen, Zahlenpartitionen, binären Bäumen und Ulam-Harris-Bäumen führen.

Schlüsselwörter Bäume · Doob-Martin-Kompaktifizierung · Partitionen · Permutationen

1 Einführung

Das Thema dieser Arbeit ist an der Schnittstelle der Gebiete Diskrete Mathematik, Stochastik und Theoretische Informatik angesiedelt: Diskrete Mathematik, weil unsere grundlegenden Objekte klassischen kombinatorischen Familien entstammen, Stochastik, weil wir zufällige solche Objekte betrachten, und Theoretische Informatik, weil diese zufälligen diskreten Strukturen durch Algorithmen mit zufälligem Input erzeugt werden.

Wir können unser Vorhaben als Ausflug in die probabilistische Kombinatorik ansehen. Konkret betrachten wir Markov-Ketten $X = (X_n)_{n \in \mathbb{N}}$ mit Werten in einer kombinatorischen Familie \mathbb{F} , wobei X_n Werte in der Menge \mathbb{F}_n aller $x \in \mathbb{F}$ mit Größe (Umfang, Basisparameter) $\phi(x) = n$ annimmt, und wollen den folgenden Sachverhalt erklären und an Beispielen illustrieren: Ausgehend von dem der Kette X zu Grunde liegenden Übergangsmechanismus lässt sich die Familie \mathbb{F} so zu einem Raum $\bar{\mathbb{F}}$ kompaktifizieren (oder vervollständigen), dass X_n mit $n \rightarrow \infty$ fast sicher gegen eine Zufallsgröße X_∞ mit Werten im Rand $\partial\bar{\mathbb{F}}$ konvergiert. Dies ist ein Spezialfall der Doob-Martin-Kompaktifizierung, die sich auf allgemeine transiente Markov-Ketten bezieht. Hier haben wir eine zusätzliche Eigenschaft, die auf eine bestimmte Optimalität der Kompaktifizierung führt.

Die nachfolgenden Abschnitte haben, soweit angebracht, ihren eigenen einführenden Absatz. Hintergrundinformationen zu Kompaktifizierungen etc., aber auch zu den bei den

R. Grübel
Institut für Mathematische Stochastik, Leibniz Universität Hannover, 30167 Hannover, Germany
E-Mail: rgrubel@stochastik.uni-hannover.de

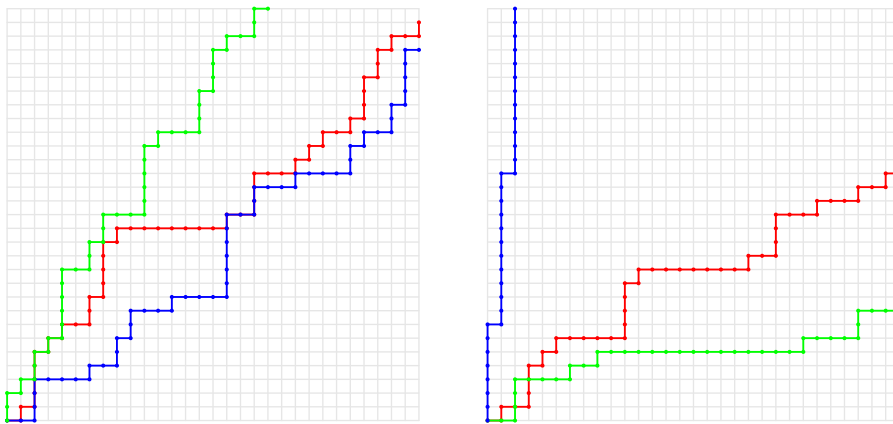


Abb. 1 Drei Pfade der symmetrischen Irrfahrt (links) und der Pólya-Urne (rechts)

Simulationen verwendeten Pseudozufallszahlen, findet man im Anhang. Im Interesse einer glatteren Lesbarkeit haben wir außerdem Literaturverweise und ergänzende Bemerkungen komplett in einen separaten Abschnitt ausgelagert.

2 Die Pólya-Urne

Wir beginnen mit dem Münzwurf, wie es sich gehört, betrachten diesen aber als zufällige Bewegung im rechten oberen Quadranten $\mathbb{N}_0 \times \mathbb{N}_0$ des Zahlengitters: Wir starten zum Zeitpunkt $n = 1$ im Koordinatenursprung $(0, 0)$ und gehen bei Wurf Ergebnis ‘Kopf’ einen Schritt nach rechts, bei ‘Wappen’ einen Schritt nach oben. Die linke Hälfte von Abbildung 1 zeigt drei solche Experimente (Details in Anhang A.3), jeweils ausgeführt bis zu dem Zeitpunkt, zu dem erstmalig eines der beiden Ergebnisse mindestens 30 Mal erschienen ist. Bei einer fairen Münze würden wir erwarten, dass sich, relativ gesehen, alle solchen Pfade der Winkelhalbierenden annähern (starkes Gesetz der großen Zahlen), und wir haben sogar eine Vorstellung von der Größenordnung der Abweichung (Zentraler Grenzwertsatz, hier: Satz von de Moivre-Laplace).

Die Pfade dieser einfachen symmetrischen Nordost-Irrfahrt können auch als Protokoll eines Urnenversuches interpretiert werden: In jedem Durchgang wird in die anfangs leere Urne eine weiße oder eine schwarze Kugel gelegt, jeweils mit Wahrscheinlichkeit $1/2$. Befindet man sich zum Zeitpunkt $n \in \mathbb{N}$ im Zustand $X_n = (i, j)$, so bedeutet dies, dass der Urne bis zu diesem Zeitpunkt i weiße und j schwarze Kugeln hinzugefügt wurden, oder, in der Münzinterpretation, dass man bis dahin i Mal ‘Kopf’ und j Mal ‘Wappen’ geworfen hat.

Im Gegensatz hierzu hängen bei der *Pólya-Urne* die Wahrscheinlichkeiten für einen Schritt nach rechts oder oben von der aktuellen Position ab. Zu Beginn, also bei $n = 1$, hat man eine weiße und eine schwarze Kugel in der Urne; im Schritt von n auf $n + 1$ wird eine der dann vorhandenen $n + 1$ Kugeln zufällig und gleichverteilt ausgewählt, der Urne entnommen und anschließend zusammen mit einer Kugel derselben Farbe zurückgelegt. Die rechte Hälfte von Abbildung 1 zeigt wieder drei solche Experimente, wobei $X_n = (i, j)$ wieder bedeutet, dass bis zum Zeitpunkt n der Urne i weiße und j schwarze Kugeln hinzugefügt wurden.

Die Ergebnisse sehen ‘irgendwie anders’ aus. Im Folgenden werden wir uns hauptsächlich mit Asymptotik beschäftigen, aber eine überschlägige Rechnung können wir bereits jetzt vornehmen: Die Wahrscheinlichkeit dafür, dass die Nordost-Irrfahrt in den ersten 30 Schritten komplett entlang einer der Koordinatenachsen verläuft, ist 2×2^{-30} , eine ziemlich kleine Zahl. Im Stochastik-Universum liegt der Vergleich mit dem Zahlenlotto ‘6 aus 49’ nahe: ‘Sechs Richtige’ ist etwa um den Faktor 38 größer. Bei der Pólya-Urne dagegen erhält man mit der Multiplikationsregel für bedingte Wahrscheinlichkeiten (oder einem Baum-Diagramm) den Wert

$$2 \cdot \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdots \frac{29}{30} = \frac{1}{15},$$

eine vergleichsweise große Zahl im genannten Universum – um den Faktor 2.4 größer als die Wahrscheinlichkeit für eine Doppel-Sechs beim Wurf zweier Würfel. (Ein Leser, der sich an dieser Stelle unterfordert fühlt, mag sich mit der Bestimmung der Verteilung der jeweiligen Austrittsposition in der Situation von Abbildung 1 aufwärmen.)

In beiden Fällen haben wir eine Folge $(X_n)_{n \in \mathbb{N}}$ von Zufallsgrößen, also einen stochastischen Prozess; n ist dabei der Zeitparameter, X_n der Zustand zum Zeitpunkt n . Innerhalb dieser großen Familie gehören beide zu den Markov-Ketten. Wir erinnern an die Grundbegriffe: Ausgangspunkt sind eine endliche oder abzählbar unendliche Menge S und eine Übergangsfunktion auf S , worunter wir eine Abbildung $p : S \times S \rightarrow [0, 1]$ mit der Eigenschaft

$$\sum_{y \in S} p(x, y) = 1 \quad \text{für alle } x \in S \quad (1)$$

verstehen. Außerdem hat man einen Startpunkt $x_1 \in S$ und nennt nun $(X_n)_{n \in \mathbb{N}}$ eine Markov-Kette mit Zustandsraum S , Übergangsfunktion p und Start in x_1 , wenn $P(X_1 = x_1) = 1$ und für alle $n \in \mathbb{N}$, $x_2, \dots, x_n \in S$

$$P(X_n = x_n | X_{n-1} = x_{n-1}, \dots, X_1 = x_1) = P(X_n = x_n | X_{n-1} = x_{n-1}) = p(x_{n-1}, x_n) \quad (2)$$

gilt, vorausgesetzt, die bedingten Wahrscheinlichkeiten sind definiert. In Worten: Der Prozess hat kein Gedächtnis, und die Funktion p gibt an, mit welcher Wahrscheinlichkeit man in einem Schritt vom ersten zum zweiten ihrer Argumente geht.

Nordost-Irrfahrt und Pólya-Urne passen in diesen Rahmen. In beiden Fällen ist $S = \mathbb{N}_0 \times \mathbb{N}_0$ der Zustandsraum, man startet in $x_1 = (0, 0)$, und die Übergangsfunktion ist

$$p((i, j), (i+1, j)) = \theta, \quad p((i, j), (i, j+1)) = 1 - \theta \quad (3)$$

mit $\theta = 1/2$ bei der einfachen symmetrischen Nordost-Irrfahrt, sowie

$$p((i, j), (i+1, j)) = \frac{i+1}{i+j+2}, \quad p((i, j), (i, j+1)) = \frac{j+1}{i+j+2} \quad (4)$$

bei der Pólya-Urne, jeweils für alle $x = (i, j) \in S$. Bemerkenswert ist, dass bei der Nordost-Irrfahrt und bei der Pólya-Urne zu einem Zeitpunkt n nur Zustände $(i, j) \in S$ mit $i+j+1 = n$ möglich sind, der Zeitpunkt also aus dem Zustand rekonstruiert werden kann. Wir sagen dann, dass der Prozess die Raum-Zeit-Eigenschaft hat. Offensichtlich ist jede solche Kette hochgradig transient – jeder Zustand kann höchstens einmal besucht werden, und dann auch nur zu einem bestimmten Zeitpunkt.

Markov-Ketten treten häufig als stochastisches dynamisches System auf, insbesondere auch in den algorithmischen Anwendungen, die wir im Folgenden behandeln werden. Damit ist eine Struktur der Form

$$X_1 := x_1, \quad X_{n+1} := \psi(X_n, U_{n+1}) \quad \text{für alle } n \in \mathbb{N}, \quad (5)$$

gemeint, wobei $(U_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen und identisch verteilten Zufallsgrößen mit Werten in einem Raum E und $\psi : S \times E \rightarrow S$ eine feste Funktion ist. In der Tat kann man *alle* Markov-Ketten auf diese Art darstellen: Schreibt man $\text{unif}(0, 1)$ für die Gleichverteilung auf $(0, 1)$, und ist $(U_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen, $\text{unif}(0, 1)$ -verteilten Zufallsvariablen, so erhält man eine Markov-Kette mit Übergangsfunktion p und Start in x_1 durch (5) – man muss nur $\psi : S \times (0, 1) \rightarrow S$ so wählen, dass das Lebesgue-Maß der Menge $\{u \in (0, 1) : \psi(x, u) = y\}$ den Wert $p(x, y)$ hat, und zwar für alle $x, y \in S$. Bei der Nordost-Irrfahrt mit Übergangsfunktion (3) geht es mit

$$\psi((i, j), u) = \begin{cases} (i+1, j), & \text{falls } u \leq \theta, \\ (i, j+1), & \text{sonst,} \end{cases} \quad (6)$$

bei der Pólya-Urne mit

$$\psi((i, j), u) = \begin{cases} (i+1, j), & \text{falls } u \leq (i+1)/(i+j+2), \\ (i, j+1), & \text{sonst.} \end{cases} \quad (7)$$

Eine praktische Anwendung ergibt sich bei der Simulation solcher Prozesse, da man dann direkt mit einem in der Regel vorhandenen algorithmischen Imitat der Folge $(U_i)_{i \in \mathbb{N}}$ (Pseudozufallszahlen) arbeiten kann; siehe auch Anhang A.3.

Als ersten Schritt einer genaueren Analyse betrachten wir die Verteilungen der einzelnen Variablen $X_n = (X_{n,1}, X_{n,2})$. Bei der Nordost-Irrfahrt ist die Position der ersten Koordinate $X_{n,1}$ nach n Schritten binomialverteilt mit Parametern $n-1$ und θ , woraus sich für $X_{n,1}/(n-1)$ der Erwartungswert θ und die mit $n \rightarrow \infty$ gegen 0 gehende Varianz $\theta(1-\theta)/(n-1)$ ergibt. Wir zeigen nun mit vollständiger Induktion, dass bei der Pólya-Urne die Zufallsgröße $X_{n,1}$ auf der Menge $\{0, \dots, n-1\}$ gleichverteilt ist: Für $n=1$ folgt dies aus den Definitionen. Zerlegt man nach dem Zustand zur Zeit n , so ergibt sich mit der Induktionsvoraussetzung, der Markov-Eigenschaft (2) und mit (4) für $k \in \{1, \dots, n-1\}$

$$\begin{aligned} P(X_{n+1,1} = k) &= P(X_{n+1,1} = k | X_{n,1} = k) P(X_{n,1} = k) \\ &\quad + P(X_{n+1,1} = k | X_{n,1} = k-1) P(X_{n,1} = k-1) \\ &= \frac{n-k}{n+1} \frac{1}{n} + \frac{k}{n+1} \frac{1}{n} \\ &= \frac{1}{n+1}. \end{aligned}$$

Eine naheliegende Modifikation zeigt, dass dies auch für $k=0$ gilt (siehe auch die überschlägige Rechnung auf S.3). Dies komplettiert den Induktionsschritt. Insbesondere geht hier also die Varianz von $X_{n,1}/(n-1)$ nicht gegen 0. Aus unserer Rechnung folgt sogar, dass dieses Verhältnis im Limes eine Gleichverteilung auf dem Einheitsintervall hat, also ‘alle Richtungen gleich wahrscheinlich’ sind.

Was passiert auf der Ebene der Zufallsgrößen selbst? Zunächst einmal liefert der Übergangsmechanismus (4)

$$\begin{aligned} E[X_{n+1,1} | X_{n,1} = k] &= (k+1)P(X_{n+1,1} = k+1 | X_{n,1} = k) + kP(X_{n+1,1} = k | X_{n,1} = k) \\ &= k + \frac{k+1}{n+1}. \end{aligned}$$

Bezeichnen wir mit $Y_{n,i} := (X_{n,i} + 1)/(n + 1)$ die relative Anzahl weißer ($i = 1$) bzw. schwarzer ($i = 2$) Kugeln zum Zeitpunkt n , und setzt man $\mathcal{F}_n := \sigma(\{X_1, \dots, X_n\})$, so führt dies mit der Markov-Eigenschaft (2) und den Rechenregeln für bedingte Erwartungswerte auf

$$E[Y_{n+1,1} | \mathcal{F}_n] = Y_{n,1} \quad \text{für alle } n \in \mathbb{N},$$

d.h. $(Y_{n,1}, \mathcal{F}_n)_{n \in \mathbb{N}}$ ist ein Martingal. Da die Y -Variablen nichtnegativ sind, lässt sich der Vorwärtskonvergenzsatz von Doob anwenden, und man erhält, dass $Y_{n,1}$ mit $n \rightarrow \infty$ gegen eine fast sicher endliche Zufallsvariable $Y_{\infty,1}$ konvergiert. (Fast sichere Konvergenz impliziert die bereits bewiesene Konvergenz der Verteilungen, ist aber eine erheblich stärkere Aussage.) Wir haben also

$$Y_n \rightarrow Y_\infty = (Y_{\infty,1}, Y_{\infty,2}) \text{ fast sicher, mit } Y_{\infty,2} = 1 - Y_{\infty,1} \text{ und } \mathcal{L}(Y_{\infty,1}) = \text{unif}(0, 1).$$

(Wir schreiben ganz allgemein $\mathcal{L}(X)$ für die Verteilung einer Zufallsgröße X .) Auch bei der Pólya-Urne erhält man also mit Wahrscheinlichkeit 1 im Limes eine Gerade durch den Koordinatenursprung, nun allerdings mit zufälliger Steigung.

Als nächstes wollen wir die bedingte Verteilung von X unter Y_∞ bestimmen. Wenn man $Y_{\infty,1}$ als Steigung der asymptotischen Geraden interpretiert, so ist die Nordost-Irrfahrt X^θ aus (3) mit $\theta = Y_{\infty,1}$ ein naheliegender Kandidat, eine Vermutung, die auf die Gleichheit zweier Wahrscheinlichkeitsmaße hinausläuft. In der vorliegenden Situation können wir uns beim Nachweis der Gleichheit auf Anfangsstücke der möglichen Pfade beschränken. Diese sind bei Länge n durch die Positionen $I \subset \{2, \dots, n\}$ der Schritte nach rechts (also der Übergänge $(i, j) \rightarrow (i + 1, j)$) festgelegt. Schreiben wir $A(n, I)$ für solche Mengen, so ergibt sich bei der Irrfahrt mit Parameter θ

$$P(X^\theta \in A(n, I)) = \theta^k (1 - \theta)^{n-1-k}, \quad \text{mit } k := \#I.$$

Bei $\text{unif}(0, 1)$ -verteiletem θ wird dies zu

$$\int_0^1 P(X^\theta \in A(n, I)) d\theta = \frac{\Gamma(k+1)\Gamma(n-k)}{\Gamma(n+1)} \quad (8)$$

(Stochastikern ist dieses Integral im Zusammenhang mit der Beta-Verteilung bekannt). Andererseits erhält man bei der Pólya-Urne für die Menge $A(n, I)$ die Wahrscheinlichkeit

$$P(X \in A(n, I)) = \frac{k!(n-1-k)!}{n!}. \quad (9)$$

Um dies einzusehen, bemerken wir zunächst, dass bei beiden Übergangstypen im Schritt von m auf $m + 1$ im Nenner der Faktor $m + 1$ hinzukommt. Hat man $I = \{n_1, \dots, n_k\}$, so liefert die Wahl einer weißen Kugel zum Zeitpunkt n_m im Zähler den Faktor m , was insgesamt für die weißen Kugeln auf $k!$ führt. Bei den schwarzen Kugeln, die zu den insgesamt $n - 1 - k$ Zeitpunkten $\{2, \dots, n\} \setminus I$ hinzugefügt werden, erhält man ganz analog den Faktor $(n - 1 - k)!$. Die bekannte Formel $\Gamma(m + 1) = m!$ für natürliche Zahlen m zeigt, dass bei (8) und (9) dasselbe herauskommt. Insgesamt erhalten wir somit das folgende Resultat.

Satz 1 *Es sei $X = (X_n)_{n \in \mathbb{N}}$ die Pólya-Urne, X^θ die Nordost-Irrfahrt mit Parameter θ . Dann gilt:*

- (a) *Mit $n \rightarrow \infty$ konvergiert $n^{-1}X_n$ fast sicher gegen eine Zufallsgröße Y_∞ mit Werten in der Menge der Paare $(\theta, 1 - \theta)$, $0 \leq \theta \leq 1$.*

- (b) Die Zufallsgröße Y_∞ aus (a) hat dieselbe Verteilung wie die Zufallsgröße $(U, 1 - U)$ mit $\mathcal{L}(U) = \text{unif}(0, 1)$.
- (c) Die bedingte Verteilung von X unter $Y_\infty = (\theta, 1 - \theta)$ stimmt überein mit der Verteilung von X^θ .

Man hat also, wie beim starken Gesetz der großen Zahlen, Konvergenz mit Wahrscheinlichkeit 1 (und spricht von einem starken Grenzwertsatz), der Grenzwert ist aber nun eine Zufallsgröße. Deren Verteilung lässt sich angeben. Bedingt man den Prozess auf einen bestimmten Wert dieser Zufallsgröße, so erhält man wieder eine Markov-Kette, deren Übergangsfunktion sich ebenfalls angeben lässt. Man kann die Pólya-Urne also auch als zweistufiges Experiment interpretieren: Zunächst wird θ zufällig gewählt, mit Verteilung $\text{unif}(0, 1)$, dann lässt man eine Nordost-Irrfahrt mit Parameter θ laufen.

Im weiteren Verlauf der Arbeit werden wir sehen, dass sich Satz 1 als Spezialfall einer allgemeinen Theorie ergibt. Darüberhinaus wird die Theorie einen Zugang zu transienten Markov-Ketten auf kombinatorischen Familien eröffnen.

3 Kombinatorische Familien und Algorithmen

Diskrete Strukturen eines bestimmten Typs fassen wir zu einer kombinatorischen Familie \mathbb{F} zusammen. Solche Strukturen x haben in der Regel einen Basisparameter (oder Größe, Umfang) $\phi(x) \in \mathbb{N}$, es sei $\mathbb{F}_n = \{x \in \mathbb{F} : \phi(x) = n\}$ die Menge aller Objekte der Größe n . In unseren Beispielen gilt stets $a_n := \#\mathbb{F}_n < \infty$ für alle $n \in \mathbb{N}$, insbesondere ist dann \mathbb{F} eine abzählbare Menge. Die Folge $(a_n)_{n \in \mathbb{N}}$ ist ein zentraler Gegenstand der enumerativen Kombinatorik.

Uns geht es um dynamische Strukturen, bei denen Relationen zwischen Objekten unterschiedlicher Größe eine Rolle spielen. Konkret haben wir eine partielle Ordnung ‘ \leq ’ auf \mathbb{F} bezüglich der die Größenfunktion isoton ist, also $\phi(x) \leq \phi(y)$ für alle $x, y \in \mathbb{F}$ mit $x \leq y$ gilt. Darüberhinaus soll zu jedem $y \in \mathbb{F}_{n+1}$ mindestens ein $x \in \mathbb{F}_n$ mit $x \leq y$ existieren, und schließlich soll \mathbb{F}_1 aus nur einem Element bestehen, für das wir gelegentlich e schreiben,

Wir beschränken uns auf vier solche Familien, die wir in diesem Abschnitt vorstellen und dann mit Hilfe von Algorithmen zueinander in Beziehung setzen. Jede dieser Familien war und ist Gegenstand zahlreicher Untersuchungen aus ganz unterschiedlichen Blickwinkeln. Als Resultat gibt es keine allgemein akzeptierte Notation – wir verwenden daher kurzerhand unsere eigene. Aus der Informatik übernehmen wir die Notation $\Sigma^* := \bigcup_{k \in \mathbb{N}_0} \Sigma^k$ für die Menge aller endlichen Worte (Zeichenketten) mit Buchstaben (Zeichen) aus dem Alphabet Σ , wobei Σ^0 aus dem einzigen Element \emptyset , dem ‘leeren Wort’ besteht. Sind $u = (u_1, \dots, u_k)$ und $v = (v_1, \dots, v_l)$ in Σ^* , so nennen wir u einen Präfix von v , wenn $k \leq l$ und $u_i = v_i$ für $i = 1, \dots, k$ gilt. Offensichtlich wird hierdurch eine partielle Ordnung definiert, die der obigen Zusammenhangsbedingung genügt. Wir können also Σ^* selbst als kombinatorische Familie ansehen, wenn wir einem Wort x der Länge k die Größe $\phi(x) = k + 1$ zuordnen. (Solche Verschiebungen können bei der Fehlersuche Kopfschmerzen machen, lassen sich aber wohl nicht ganz vermeiden. Dies betrifft auch die unterschiedlichen Konventionen bzgl. des ersten Zeitpunktes bei einem stochastischen Prozess – soll man bei $n = 0$ oder bei $n = 1$ anfangen?)

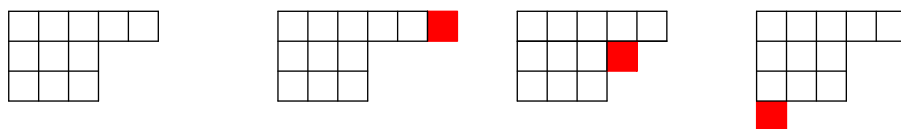


Abb. 2 Die Young-Diagramme zu $\lambda = (5, 3, 3)$ und dessen Nachfolgern $(6, 3, 3)$, $(5, 4, 3)$ und $(5, 3, 3, 1)$

3.1 Permutationen

Bekanntlich nennt man eine bijektive Abbildung $\pi : M \rightarrow M$ eine Permutation der Menge M . Wir schreiben \mathbb{S}_n für die Menge der Permutationen von $\{1, 2, \dots, n\}$ und setzen $\mathbb{S} = \bigcup_{n=1}^{\infty} \mathbb{S}_n$. Der Parameter $\phi(\pi)$ einer Permutation $\pi \in \mathbb{S}$ ist die Anzahl der Elemente von M .

Aus den Grundvorlesungen zur Mathematik ist \mathbb{S} als ‘Mutter aller endlichen Gruppen’ bekannt.¹ In der elementaren Kombinatorik und Stochastik tauchen Permutationen beispielsweise beim Rencontre-Problem auf, bei dem es (in der Grundschulvariante) um die Wahrscheinlichkeit dafür geht, dass ein Kind beim ‘Wichteln’ sein eigenes Geschenk erhält. Zufällig und gleichverteilte Zufallsgrößen mit Werten in \mathbb{S}_{32} bzw. \mathbb{S}_{52} bilden eine Grundlage des prä-elektronischen Unterhaltungswesens. Auch ist $\#\mathbb{S}_n = n!$ wohlbekannt.

Eine Permutation $\pi \in \mathbb{S}_n$ ist eine Abbildung, lässt sich also als Tabelle schreiben, mit den Argumenten $1, \dots, n$ in der ersten und den Werten $\pi(1), \pi(2), \dots, \pi(n)$ in der zweiten Zeile. Dabei kann natürlich die erste Zeile auch weggelassen werden. Andererseits zerlegt die bijektive Abbildung π den Grundraum in Bahnen, die sich durch wiederholtes Anwenden von π ergeben. Dies führt auf die Zykelschreibweise, die beispielsweise durch die Forderung standardisiert werden kann, dass das kleinste Element jeweils als Erstes aufgelistet wird und die Zyklen dann aufsteigend nach ihrem ersten Element angeordnet werden. Zu dem als Tabelle geschriebenen Element

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 5 & 4 & 9 & 6 & 8 & 7 & 1 & 3 \end{pmatrix} \tag{10}$$

von \mathbb{S}_9 erhält man so die Zykelschreibweise $((1, 2, 5, 6, 8)(3, 4, 9)(7))$.² Das siebte Kind hätte also Anlass, enttäuscht zu sein.

Eine naheliegende partielle Ordnung ergibt sich durch ‘Weglassen’: Entsteht $\sigma \in \mathbb{S}_n$ aus $\pi \in \mathbb{S}_{n+1}$ durch Streichen des Wertes $n + 1$ in der reduzierten Tabellenschreibweise, oder der Zykelschreibweise, so setzen wir $\sigma \leq \pi$. Im ersten Fall würde Weglassen des größten Wertes auf $\sigma = (2, 5, 4, 6, 8, 7, 1, 3)$, im zweiten auf $\sigma = ((1, 2, 5, 6, 8)(3, 4)(7))$ führen. Insbesondere sind die so erhaltenen Ordnungen nicht gleich, da man im ersten Fall $\sigma(4) = 6$ und im zweiten $\sigma(4) = 3$ erhält. In beiden Fällen ist aber klar, wie man die Ordnung auf Paare fortsetzt, deren Größenparameter sich um mehr als 1 unterscheiden. Auch gelten in beiden Fällen die zu Beginn von Abschnitt 3 genannten Eigenschaften: Zu einer Permutation $\pi \in \mathbb{S}_{n+1}$ erhält man einen Vorgänger $\sigma \in \mathbb{S}_n$ durch das bereits erwähnte Streichen von $n + 1$ in der jeweiligen Darstellung, und offensichtlich ist $e = (1)$ das einzige Element von \mathbb{S}_1 .

¹ Jede Gruppe mit n Elementen ist isomorph zu einer Untergruppe von \mathbb{S}_n .

² Formal ist dies bei der hier im Vordergrund stehenden Sichtweise eine geordnete Menge von geordneten Mengen; in der Algebra ist die kompaktere Schreibweise $(1\ 2\ 5\ 6\ 8)(3\ 4\ 9)(7)$ mit der Interpretation als Produkt von zyklischen Permutationen gebräuchlich.

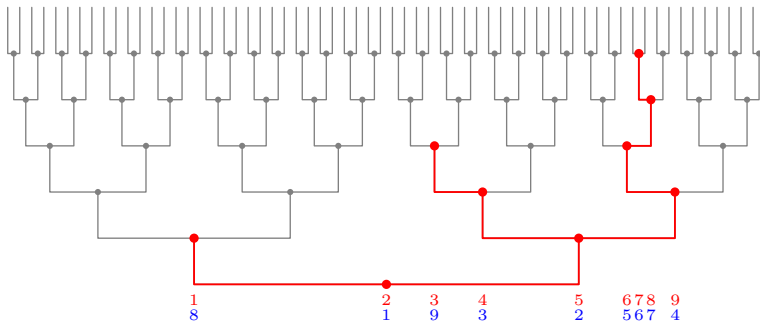


Abb. 3 Ein binärer Baum (rot) mit Bewertung (rot) und Aufnahmereihenfolge (blau) der Knoten

3.2 Zahlenpartitionen

Unter einer Partition einer natürlichen Zahl n verstehen wir eine Zerlegung dieser Zahl in Summanden, wobei die Reihenfolge *nicht* berücksichtigt wird. Jede solche Partition lässt sich eindeutig als antitone Summandenfolge schreiben. Schreiben wir \mathbb{Y} für die gesamte kombinatorische Familie, so besteht die Teilfamilie der Objekte mit Basisparameter n aus

$$\mathbb{Y}_n := \left\{ \lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{N}^* : k \in \mathbb{N}, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k, \sum_{i=1}^k \lambda_i = n \right\}.$$

Man schreibt bei Partitionen λ aus \mathbb{Y}_n auch kurz $\lambda \vdash n$. Die Untersuchung der Partitionszahlen $p(n) = \#\mathbb{Y}_n$ reflektiert ein interessantes Stück Mathematikgeschichte; wir notieren nur, dass $p(n) \sim (4n\sqrt{3})^{-1} \exp(\pi\sqrt{2n/3})$ gilt, die Anzahl also subexponentiell, aber schneller als jede Potenz wächst. (Die Notation $a_n \sim b_n$ bedeutet, dass das Verhältnis der beiden Zahlen mit $n \rightarrow \infty$ gegen 1 geht.)

Zahlenpartitionen lassen sich durch Young-Diagramme visualisieren, die aus vertikal und horizontal angeordneten Quadraten derselben Seitenlänge bestehen. In der ersten Reihe hat man λ_1 solche ‘Kästchen’, linksbündig hierunter λ_2 in der zweiten etc.; Abbildung 2 zeigt ganz links ein Beispiel, mit $\lambda = (5, 3, 3) \in \mathbb{Y}_{11}$.

Die Wahl der partiellen Ordnung auf \mathbb{Y} ist einigermaßen kanonisch und kann auf der Ebene der Diagramme durch Festlegen der möglichen Positionen für das nächste Kästchen erfolgen: Das Bild muss zusammenhängend bleiben, und die Anzahl der Kästchen darf auf dem Weg nach unten nicht zunehmen. Die Diagramme in Abbildung 2 verdeutlichen dies durch rote Einfärbung der möglichen Nachfolgekästchen.

3.3 Binäre Bäume

Wir nennen eine Teilmenge x von $\{0, 1\}^*$ einen binären Baum, wenn x mit jedem u auch alle v enthält, die Präfix von u sind, und schreiben \mathbb{B}_n für die Menge der endlichen Bäume mit n Elementen (Knoten) sowie $\mathbb{B} = \bigcup_{n=1}^{\infty} \mathbb{B}_n$ für die kombinatorische Familie der endlichen binären Bäume. Lässt man Bäume nach oben wachsen, so bietet sich die Visualisierung durch Punkte aus $[0, 1] \times \mathbb{N}_0$ an: Jedem $u = (u_1, \dots, u_k)$ wird der Punkt mit den Koordinaten

$$\beta(u) := \frac{1}{2} + \sum_{j=1}^k \frac{2u_j - 1}{2^{j+1}} \quad (11)$$

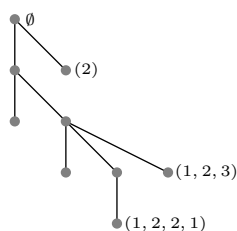


Abb. 4 Ein ebener Baum, mit ausgewählten Knotenadressen

und k zugeordnet. In Abbildung 3 ist dies für

$$x = \{\emptyset, (0), (1), (1, 0), (1, 0, 0), (1, 1), (1, 1, 0), (1, 1, 0, 1), (1, 1, 0, 1, 1)\} \in \mathbb{B}_9$$

ausgeführt, inklusive augenfreundlicher Hilfslinien. Wir schreiben $x \leq y$, wenn $x \subset y$ gilt. Die zugehörigen Anzahlen sind $\#\mathbb{B}_n = C_n$, wobei $C_n = \frac{1}{n+1} \binom{2n}{n}$ die n -te Catalansche Zahl bezeichnet. Die asymptotische Größenordnung $C_n \sim 4^n / \sqrt{\pi n^3}$ ergibt sich leicht mit der Stirling-Formel, sie liegt zwischen der der Permutationen und der der Zahlenpartitionen.

3.4 Ebene Bäume

Unter einem ebenen Baum verstehen wir eine präfix-stabile Teilmenge x von \mathbb{N}^* , die der Bedingung genügt, dass mit $u = (u_1, \dots, u_k)$ im Falle $u_k > 1$ auch $u = (u_1, \dots, u_k - 1)$ ein Knoten des Baumes ist. Wir schreiben \mathbb{E}_n für die Menge der ebenen Bäume mit $n + 1$ Knoten, $\mathbb{E} := \bigcup_{n=1}^{\infty} \mathbb{E}_n$ für die gesamte Familie. Das einzige Element von \mathbb{E}_1 ist $\{\emptyset, (1)\}$. Diese Strukturen werden auch Ulam-Harris-Bäume genannt und tauchen in der Stochastik im Zusammenhang mit Verzweigungsprozessen auf. Abbildung 4 illustriert den Baum

$$x = \{\emptyset, (1), (2), (1, 1), (1, 2), (1, 2, 1), (1, 2, 2), (1, 2, 3), (1, 2, 2, 1)\} \in \mathbb{E}_8.$$

Die Analogie zu Stammbäumen mit eingeschlechtlicher Vermehrung ist offensichtlich, sie ist auch ein Grund für die Darstellung mit einem Wachstum von oben nach unten. Im Gegensatz zu binären Bäumen kann ein Knoten nun mehr als zwei Nachkommen haben, und bei nur einem Nachkommen wird nicht zwischen links und rechts unterschieden. Die kanonische partielle Ordnung ist wieder die Enthaltenseinsrelation. Als Anzahl der Elemente von \mathbb{E}_n ergibt sich abermals die n -te Catalansche Zahl. Diese Übereinstimmung *muss* einen ‘bijektiven Beweis’ haben, und in der Tat wird durch einen als ‘tree rotation’ bekannten Algorithmus eine bijektive Abbildung zwischen \mathbb{E}_n und \mathbb{B}_n definiert.

3.5 Algorithmen

Ordnet man einer Permutation $\pi \in \mathbb{S}$ den Vektor der absteigenden Zyklennlängen zu, so ergibt sich offensichtlich eine Abbildung $\Psi : \mathbb{S} \rightarrow \mathbb{Y}$, die den Basisparameter unverändert lässt: Permutiert π die Zahlen $1, \dots, n$, so ist die Summe der Zyklennlängen n , denn jede Zahl kommt in genau einem Zykel vor. Natürlich ist diese Abbildung nicht injektiv – nur die Identität führt auf $(1, 1, \dots, 1) \vdash n$, alle $(n - 1)!$ zyklischen Permutationen ergeben (n) , die Partition mit nur einem Summanden.

1	3	6	7
2	4		
5	8		
9			

1	2	4	6
3	5		
7	9		
8			

7	5	2	1
4	2		
3	1		
1			

Abb. 5 Zum RSK-Algorithmus und der Hakenformel (siehe Text)

Eine weitere solche Abbildung von \mathbb{S} nach \mathbb{Y} erhält man über den berühmten RSK-Algorithmus (das Akronym steht für die Mathematiker/Informatiker Robinson, Schensted und Knuth). Dieser liefert zu einer Permutation ein Young-Tableau in Standardform, in dem die n Kästchen eines Young-Diagramms mit den Zahlen $1, \dots, n$ in nach rechts und nach unten streng monotoner Anordnung belegt sind. Ausgehend von der zweiten Zeile der Tabellenschreibweise wird dabei zunächst das einzige Kästchen des einzigen Elements von \mathbb{Y}_1 mit $\pi(1)$ belegt. Hat man bereits $\pi(1), \dots, \pi(k)$ in ein Tableau auf der Basis eines Elements y von \mathbb{Y}_k eingefügt, so verfährt man bei Ankunft des nächsten Wertes $\pi(k+1) = i$ wie folgt: Gibt es in der ersten Zeile von y keinen größeren Wert, so wird dieser Zeile rechts ein weiterer Kästchen angefügt und mit i belegt. Ansonsten ersetzt i das kleinste Element j der ersten Zeile von y , das größer als i ist, und man macht rekursiv mit j und dem um die erste Zeile reduzierten Tableau weiter. Landet man schließlich in einer leeren Zeile, so wird unten ein Einzelkästchen angehängt. Zu der Permutation in Abschnitt 3.1 mit zweiter Tabellenzeile $(2, 5, 4, 9, 6, 8, 7, 1, 3)$ ergibt sich so das linke Tableau in Abbildung 5, das mittlere Tableau notiert die Aufnahmereihenfolge der einzelnen Quadrate. Eine Abbildung $\Psi : \mathbb{S} \rightarrow \mathbb{Y}$ erhält man aus dem RSK-Algorithmus durch Weglassen der Bewertungen, in unserem numerischen Beispiel wäre dies $\lambda = (4, 2, 2, 1)$. Bei der durch absteigende Zyklenlängen definierten Abbildung würde man als Partition $\lambda = (5, 3, 1)$ erhalten.

Auch der uns interessierende \mathbb{S} und \mathbb{B} verbindende Algorithmus ist unter einem Akronym bekannt, und auch hier gehen wir den Umweg über eine bewertete Struktur: Der BST-Algorithmus (binary search tree) sortiert die Werte in der zweiten Zeile der Tabellenschreibweise von π von links nach rechts in die Knoten eines binären Baumes ein. Dabei wird zunächst $\pi(1)$ dem einzigen Knoten des einzigen Elements von \mathbb{B}_1 , also dem Wurzelknoten \emptyset , zugeordnet. Hat man bereits $\pi(1), \dots, \pi(k)$ in ein Element x von \mathbb{B}_k eingefügt, so verfährt man bei Ankunft des nächsten Wertes $\pi(k+1) = i$ wie folgt: Bei der Wurzel beginnend vergleicht man i mit der Bewertung j des aktuellen (belegten) Knotens und geht im Falle $i < j$ zum linken, bei $i > j$ zum rechten Nachkommen. Sollte man x auf diese Weise verlassen, so wird der neue Knoten dem Baum x hinzugefügt. Die Permutation aus Abschnitt 3.1 führt so auf den binären Baum in Abbildung 3. Dabei haben wir die Bewertungen nicht direkt an die Knoten geschrieben, sondern unten an deren β -Projektion, mit β wie in (11). Dies erleichtert das Entdecken einer fundamentalen Eigenschaft des BST-Algorithmus: Eine als 'in order traversal' bekannte Reise durch den Baum liefert die Werte in aufsteigend sortierter Form. Zusätzlich sind unten in der Abbildung zu den Knoten noch die Zeitpunkte ihrer Aufnahme in den Baum angegeben (auch hier gibt es etwas zu entdecken).

Um ebene Bäume einbinden zu können, aber auch im Hinblick auf spätere Aussagen, führen wir eine weitere kombinatorische Familie ein:

$$\mathbb{A}_n := \prod_{k=1}^n \{1, \dots, k\}, \quad \mathbb{A} := \bigcup_{n=1}^{\infty} \mathbb{A}_n,$$

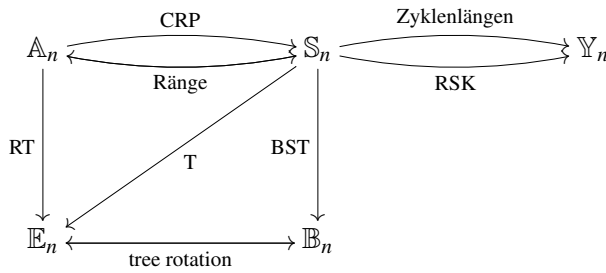


Abb. 6 Algorithmische Verbindungen zwischen den kombinatorischen Familien

mit der kanonischen Bewertungsfunktion und der Präfixordnung. Dabei ist mit dem Produkt von Mengen das kartesische Produkt gemeint. Man sieht leicht, dass $\#\mathbb{A}_n = \#\mathbb{S}_n$ gilt, es gibt also bijektive Abbildungen zwischen \mathbb{A}_n und \mathbb{S}_n . Es gibt sogar zwei, die für unsere Zwecke geeignet sind: Ψ_1 betrachtet den Input-Vektor $a = (a_1, \dots, a_n)$ als Vektor der relativen Ränge der Werte in der zweiten Zeile der Tabelle zu π . Bei der Umkehrabbildung werden absolute Ränge zu relativen Rängen,

$$(\Psi_1^{-1}(\pi))_k := \#\{1 \leq i \leq k : \pi(i) \leq \pi(k)\}.$$

Die Beispielpermutation aus Abschnitt 3.1 erhält man als Bild von $a = (1, 2, 2, 4, 4, 5, 5, 1, 3)$ unter $\Psi_1(a)$. Bei Ψ_2 wird die Zyklusstruktur von $\pi \in \mathbb{S}_n$ schrittweise aus den Komponenten a_1, \dots, a_n von $a \in \mathbb{A}_n$ aufgebaut. Die Prozedur ist als ‘Chinese restaurant process’ (CRP) bekannt: Der erste Kunde setzt sich an den ersten (runden) Tisch. Sind bereits k Kunden da, so sucht sich Kunde $k + 1$ als rechten Sitznachbarn den Kunden mit der Nummer a_{k+1} aus; im Falle $a_{k+1} = k + 1$ macht er einen neuen Tisch auf. Zu $a = (1, 2, 2, 4, 4, 5, 5, 1, 3)$ erhält man so die Permutation (in Zykelschreibweise) $\Psi_2(a) = ((1, 8), (2, 9, 3), (4, 6, 7, 5))$.

Die Elemente von \mathbb{A} liefern auch eine Konstruktionsvorschrift für einen ebenen Baum, wobei wir wieder den Umweg über eine bewertete Struktur gehen (und nach erfolgter Konstruktion die Werte weglassen). Wie im binären Fall wächst der Baum knotenweise, wobei die Bewertungen die Reihenfolge festhalten, mit der die Knoten aufgenommen werden. Für den schrittweisen Aufbau von $x \in \mathbb{E}_n$ aus $a = (a_1, \dots, a_n) \in \mathbb{A}_n$ verfahren wir nun wie folgt: An den bereits vorhandenen Wurzelknoten (mit Wert 0) hängen wir einen Knoten mit Wert 1 als ersten Nachkommen an. Haben wir bereits k Knoten eingefügt, mit den Bewertungen $1, \dots, k$, so wird der nächste Knoten an den Knoten mit Bewertung (Ankunftszeit) $a_{k+1} - 1$ als neuer Nachkomme (in unserem Bild: rechts) angehängt. Auf diese Art bewertete ebene Bäume sind auch als rekursive Bäume bekannt; wir schreiben kurz RT für die resultierende Abbildung von \mathbb{A}_n nach \mathbb{E}_n . Mit $a = (1, 2, 2, 4, 1, 4, 7, 4)$ erhält man den Baum in Abbildung 4.

In Abbildung 6 sind die algorithmischen Querverbindungen zusammengefasst. Nur T haben wir noch nicht besprochen: Ausgehend von der zweiten Tabellenzeile zu π erhält man x , indem Knoten i an den am weitesten rechts stehenden Knoten $k \leq i$ mit $\pi(k) < \pi(i)$ als neuer Nachkomme angehängt wird. Gibt es kein solches i , so wählt man den Wurzelknoten als direkten Vorfahren. Man überprüft leicht, dass alle diese Abbildungen isoton sind, also die partiellen Ordnungen der Familien erhalten; bei \mathbb{S} muss natürlich die passende der beiden in Abschnitt 3.1 beschriebenen partiellen Ordnungen gewählt werden.

4 Kombinatorische Markov-Ketten

Wir führen nun die beiden titelgebenden Konzepte zusammen: Ist \mathbb{F} eine kombinatorische Familie mit Größenfunktion ϕ und partieller Ordnung ' \leq ' und ist $X = (X_n)_{n \in \mathbb{N}}$ eine Markov-Kette mit Zustandsraum \mathbb{F} , so nennen wir X eine kombinatorische Markov-Kette zur Familie \mathbb{F} , wenn $P(X_n \in \mathbb{F}_n) = 1$ gilt für alle $n \in \mathbb{N}$, und man außerdem für alle $n \in \mathbb{N}$, $x \in \mathbb{F}_n$, $y \in \mathbb{F}_{n+1}$

$$P(X_{n+1} = y | X_n = x) > 0 \implies x \leq y \quad (12)$$

hat. Wir werden außerdem eine schwache Form der Irreduzibilität, nämlich

$$P(X_n = x) > 0 \quad \text{für alle } x \in \mathbb{F}_n \quad (13)$$

voraussetzen. Dieses Konzept ist natürlich nicht auf die in Abschnitt 3 betrachteten Familien beschränkt. In der Tat kann man sogar die Pólya-Urne und die Nordost-Irrfahrten in diesen Rahmen einbetten, wenn man $\mathbb{F} = \mathbb{N}_0 \times \mathbb{N}_0$ mit $\phi(i, j) = i + j + 1$ wählt und $(i, j) \leq (k, l)$ im Falle $i \leq l$ und $j \leq k$ schreibt.

Eine dynamische Sichtweise auf kombinatorische Familien ist bereits durch die Algorithmen aus Abschnitt 3.5 vorgegeben: RSK, BST, RT, CRP und die Rangübergänge arbeiten alle sequentiell die Komponenten des Eingabevektors ab. Dort ergab sich auch \mathbb{A} als Basisfamilie. Für \mathbb{A} ist die Konstruktion einer Markov-Kette im Stil eines dynamischen Systems der Form (5) besonders einfach: Man hängt an den aktuellen Vektor (a_1, \dots, a_n) die Komponente $a_{n+1} = \lceil (n+1)U_{n+1} \rceil$ an, wobei wieder $(U_i)_{i \in \mathbb{N}}$ eine Folge von unabhängigen, $\text{unif}(0, 1)$ -verteilten Zufallsvariablen ist. Mit dieser Konstruktion laufen die beiden in Abschnitt 3.5 genannten Abbildungen von \mathbb{A} auf \mathbb{S} auf zwei Verfahren zum induktiven Kartenmischen hinaus: Im rangbasierten Fall wird die $(n+1)$ -te Karte in eine zufällig und gleichverteilte Lücke (das schließt die beiden Positionen ganz links und ganz rechts von der aktuellen Zeile ein) gelegt, bei der zyklenorientierten Methode wird jede der n bereits gelegten Karten mit Wahrscheinlichkeit $1/(n+1)$ der rechte Nachbar, mit der verbleibenden Wahrscheinlichkeit $1/(n+1)$ beginnt man einen neuen Zykel (siehe auch CRP, S. 11).

Beide Verfahren liefern eine bijektive, monotone und größentreue Abbildung Ψ von \mathbb{A} nach \mathbb{S} . Es ist leicht einzusehen, dass man mit einer solchen Abbildung aus einer kombinatorischen Markov-Kette $X = (X_n)_{n \in \mathbb{N}}$ zur einen Familie durch $Y_n := \Psi(X_n)$ eine kombinatorische Markov-Kette $Y = (Y_n)_{n \in \mathbb{N}}$ zur anderen Familie erhält.

Ist Ψ nicht injektiv, so kann beim Übergang von X auf Y aber durchaus die Markov-Eigenschaft verlorengehen. Betrachtet man die in Abschnitt 3.5 beschriebenen Abbildungen unter diesem Gesichtspunkt, so stellt sich folgendes heraus: Der BST-Algorithmus führt auf eine kombinatorische Markov-Kette zu \mathbb{B} , mit Übergangswahrscheinlichkeiten

$$p(x, y) = \begin{cases} 1/(1 + \phi(x)), & \text{wenn } x \leq y \text{ und } \phi(y) = \phi(x) + 1, \\ 0, & \text{sonst.} \end{cases} \quad (14)$$

In Worten: Der nächste Baum wird zufällig und gleichverteilt aus der Menge der Möglichkeiten ausgesucht. Auch der durch RT induzierte Übergang von \mathbb{A} auf \mathbb{E} führt auf eine Markov-Kette mit der Uniformitätseigenschaft (14).

Bei der kombinatorischen Familie \mathbb{Y} folgt aus $\lambda \leq \mu$ mit $\lambda = (\lambda_1, \dots, \lambda_k) \in \mathbb{Y}_n$, $\mu = (\mu_1, \dots, \mu_l) \in \mathbb{Y}_{n+1}$, dass entweder $\mu = (\lambda_1, \dots, \lambda_k, 1)$ gilt oder ein $j \in \{1, \dots, k\}$ existiert

mit $\mu = (\lambda_1, \dots, \lambda_{j-1}, \lambda_j + 1, \lambda_{j+1}, \dots, \lambda_k)$. Bei der Zyklen-Kette erhält man als Übergangswahrscheinlichkeiten

$$p(\lambda, \mu) = \begin{cases} 1/(\phi(\lambda) + 1), & \text{wenn } \mu = (\lambda_1, \dots, \lambda_k, 1), \\ \lambda_j \cdot \#\{i \geq j : \lambda_i = \lambda_j\} / (\phi(\lambda) + 1), & \text{wenn } \mu = (\lambda_1, \dots, \lambda_j + 1, \dots, \lambda_k) \\ 0, & \text{sonst.} \end{cases} \quad (15)$$

Für die in Abbildung 2 angegebenen Nachfolger ergeben sich so, von links nach rechts, die Übergangswahrscheinlichkeiten $5/12$, $1/2$ und $1/12$.

Im Abschnitt 2 haben wir für die Nordost-Irrfahrt und die Pólya-Urne die Verteilung von X_n explizit angeben können. Entsprechende Formeln existieren auch für die kombinatorischen Markov-Ketten dieses Abschnitts; sie laufen auf das Zählen der Permutationen $\pi \in \mathbb{S}_n$ hinaus, die auf das Objekt $x \in \mathbb{F}_n$ führen. Bei der BST-Kette erhält man

$$P(X_n = x) = \prod_{u \in x} \frac{1}{\#\{v \in x : u \leq v\}}, \quad (16)$$

bei der Zyklenkette

$$P(X_n = \pi) = \prod_{j=1}^{\infty} \frac{1}{j^{c_j(\pi)} c_j(\pi)!}, \quad (17)$$

wobei $c_j(\pi)$ die Anzahl der Zyklen der Länge j in π bezeichnet. Bei der RSK-Kette wird man auf die Frage geführt, wieviele Bewertungen es zu einer vorgegebenen Partition λ gibt, die aus dieser ein Tableau in Standardform machen. Die Antwort gibt die berühmte Hakenformel, für die auch Nicht-Stochastiker gern einen probabilistischen Beweis heranziehen: Der Haken zu einem Quadrat in Zeile i und Spalte j in λ besteht aus dem Quadrat selbst sowie den darunter und den rechts daneben liegenden, $h_\lambda(i, j)$ ist die zugehörige Länge (Anzahl der Quadrate). Mit

$$H(\lambda) := \prod_{(i,j) \in \lambda} h_\lambda(i, j), \quad (18)$$

wird die gesuchte Anzahl der Bewertungen dann zu $\phi(\lambda)! / H(\lambda)$. In Abbildung 5 sind ganz rechts bei einem Element von \mathbb{Y}_9 die Hakenlängen eingetragen. Mit H lassen sich nun die Übergangswahrscheinlichkeiten und die Randverteilungen zur RSK-Kette angeben:

$$p(\lambda, \mu) = \frac{H(\lambda)}{H(\mu)}, \quad P(X_n = \lambda) = \frac{\phi(\lambda)!}{H(\lambda)^2}. \quad (19)$$

In der linken Formel muss natürlich $\lambda \leq \mu$ und $\phi(\mu) = \phi(\lambda) + 1$ vorausgesetzt werden. In der Situation von Abbildung 2 ergeben sich so, von links nach rechts, die Übergangswahrscheinlichkeiten $5/24$, $4/15$ und $21/40$. Die rechte Formel in (19) ist eine unmittelbare Konsequenz der fundamentalen Eigenschaft des RSK-Algorithmus, dass durch diesen \mathbb{S}_n auf die Menge der Paare von Young-Tableaus (oder Young-Tableaux) zu jeweils demselben Young-Diagramm aus \mathbb{Y}_n bijektiv abgebildet wird.

5 Die Doob-Martin-Kompaktifizierung

Nachdem wir in den letzten beiden Abschnitten kombinatorische Familien und Markov-Ketten mit Werten in diesen Familien eingeführt haben, wenden wir uns nun der zentralen Frage zu: Kann man den Zustandsraum so erweitern, dass die Zufallsgrößen einen

(fast sicheren) Grenzwert haben? Man kann, sogar ganz allgemein bei transienten Markov-Ketten. Bei kombinatorischen Markov-Ketten ergeben sich einige Vereinfachungen sowie eine zusätzliche Aussage.

Wir betrachten zunächst die abstrakte Konstruktion, arbeiten dann die Details für die in Abschnitt 2 eingeführte Pólya-Urne aus, und geben schließlich die zentralen Resultate der Theorie an.

5.1 Erweiterung des Zustandsraums

Es sei nun $X = (X_n)_{n \in \mathbb{N}}$ eine kombinatorische Markov-Kette zu einer kombinatorischen Familie \mathbb{F} . Der zugehörige Martin-Kern ist die durch

$$K(x, y) = \begin{cases} \frac{P(X_n = y | X_m = x)}{P(X_n = y)}, & \text{falls } x \in \mathbb{F}_m, y \in \mathbb{F}_n, n \geq m, \\ 0, & \text{sonst,} \end{cases}$$

definierte Abbildung $K : \mathbb{F} \times \mathbb{F} \rightarrow \mathbb{R}_+$ (hier ist (13) nützlich). Wir versehen \mathbb{F} mit der diskreten Topologie; insbesondere sind dann alle Abbildungen $f : \mathbb{F} \rightarrow \mathbb{R}$ stetig. Wegen

$$K(x, y) = \frac{P(X_n = y, X_m = x)}{P(X_n = y)P(X_m = x)} \leq \frac{1}{P(X_m = x)} =: c(x) < \infty \quad \text{für alle } y \in \mathbb{F} \quad (20)$$

kann das System der Funktionen $y \mapsto K(x, y)/c(x)$, $x \in \mathbb{F}$, als Ausgangspunkt für die in Anhang A.1 beschriebene Konstruktion dienen (die dortige Trennungsbedingung (27) ist leicht zu verifizieren). Man erhält so die *Doob-Martin-Kompaktifizierung* $\bar{\mathbb{F}}$ von \mathbb{F} und bezeichnet $\partial\mathbb{F} := \bar{\mathbb{F}} \setminus \mathbb{F}$ als den *Martin-Rand*.

Diese Erweiterung des Zustandsraumes ergibt sich auch über eine geeignete metrische Struktur. Bekanntlich lässt sich die diskrete Topologie auf \mathbb{F} durch d_0 metrisieren, wobei $d_0(x, y) = 1$ im Falle $x \neq y$, und natürlich $d_0(x, x) = 0$. Sei $w : \mathbb{F} \rightarrow (0, \infty)$ eine Funktion mit der Eigenschaft $\sum_{x \in \mathbb{F}} w(x) < \infty$. Mit d_0 und w definieren wir eine weitere Metrik d auf \mathbb{F} durch

$$d(y, z) := \sum_{x \in \mathbb{F}} w(x) (c(x) + 1)^{-1} (|K(x, y) - K(x, z)| + d_0(y, z)),$$

mit $c(x)$ gemäß (20). Nun verfährt man wie in Anhang A.2 beschrieben. Die dortige Beschränktheitsbedingung (28) ist leicht nachzuweisen, man erhält also insgesamt einen kompakten metrischen Raum.

Für das weitere Vorgehen ist wichtig, dass der Martin-Kern stetig von $S \times S$ auf $S \times \bar{S}$ fortgesetzt werden kann; wir bezeichnen auch die Fortsetzung mit K . In der Tat ist Konvergenz einer Folge $(y_n)_{n \in \mathbb{N}}$ in dem vervollständigten Raum äquivalent dazu, dass sie eine der beiden folgenden Bedingungen erfüllt:

- (i) die Folge ‘friert ein’, d.h. es gibt ein n_0 mit $y_n = y_{n_0}$ für alle $n \geq n_0$,
- (ii) für jedes $x \in \mathbb{F}$ konvergiert $K(x, y_n)$ mit $n \rightarrow \infty$ in \mathbb{R} .

Bei kombinatorischen Markov-Ketten hat (i) stets die Wahrscheinlichkeit 0, denn es gilt $P(X_n \in \mathbb{F}_n) = 1$ für alle $n \in \mathbb{N}$.

5.2 Ein Spezialfall

Es sei $X = (X_n)_{n \in \mathbb{N}}$ die Pólya-Urne aus Abschnitt 2, also $X_n = (X_{n,1}, X_{n,2})$, wobei $X_{n,i}$ die Anzahl der bis zum Zeitpunkt n hinzugefügten weißen ($i = 1$) bzw. schwarzen ($i = 2$) Kugeln bezeichnet. Aus Abschnitt 4 wissen wir, dass dieser Prozess sich als kombinatorische Markov-Kette zur Familie $\mathbb{F} = \mathbb{N}_0 \times \mathbb{N}_0$ auffassen lässt.

Wie sieht der zugehörige Martin-Kern aus? Zur Zeit n befinden sich $n + 1$ Kugeln in der Urne, $X_{n,1} + 1$ von diesen sind weiß. Um von (i, j) zu (k, l) zu kommen, was nur bei $i \leq k$ und $j \leq l$ möglich ist, benötigt man $k - i$ Schritte nach rechts und $l - j$ Schritte nach oben. Die folgende Argumentation ist uns bereits vertraut: Mit der Multiplikationsregel für bedingte Wahrscheinlichkeiten erhält man für jeden einzelnen solchen Pfad ein Produkt von Brüchen, wobei im Nenner $(i + j + 2) \cdots (k + l + 1)$ sowie im Zähler $(i + 1) \cdots (k - 1)k$ für die Schritte nach rechts und $(j + 1) \cdots (l - 1)l$ für die Schritte nach oben steht. Es gibt $\binom{k+l-i-j}{k-i}$ Möglichkeiten, die Schritte nach rechts auf die möglichen Positionen zu verteilen. Dies liefert

$$P(X_{k+l+1} = (k, l) | X_{i+j+1} = (i, j)) = \binom{k+l-i-j}{k-i} \frac{(i+1) \cdots (k-1)k(j+1) \cdots (l-1)l}{(i+j+2) \cdots (k+l+1)}$$

und damit

$$K((i, j), (k, l)) = \frac{P(X_{k+l+1} = (k, l) | X_{i+j+1} = (i, j))}{P(X_{k+l+1} = (k, l))} = \frac{\binom{k+l-i-j}{k-i} (i+j+1)!}{\binom{k+l}{k} i! j!}.$$

Der Quotient der Binomialkoeffizienten wird zu

$$\frac{(k+l-i-j)!k!l!}{(k+l)!(k-i)!(l-j)!} = \frac{k(k-1) \cdots (k-i+1)}{(k+l) \cdots (k+l-i+1)} \cdot \frac{l(l-1) \cdots (l-j+1)}{(k+l-i) \cdots (k+l-i-j+1)}.$$

Im linken Bruch hat man oben und unten i Faktoren, im rechten j . Man macht sich leicht klar, dass dieser Ausdruck für eine nicht einfrierende Folge $((k_n, l_n))_{n \in \mathbb{N}}$ von Zuständen dann für alle $i, j \in \mathbb{N}_0$ konvergiert, wenn $k_n/(k_n + l_n)$ einen Grenzwert $\alpha \in [0, 1]$ hat. Wegen

$$K((1, 0), (k, l)) = \frac{2k}{k+l}$$

ist diese Bedingung auch notwendig. Man kann $\bar{\mathbb{F}}$ also mit $\mathbb{F} \cup [0, 1]$ identifizieren und erhält als stetige Fortsetzung des Martin-Kerns

$$K((i, j), \alpha) = \frac{(i+j+1)!}{i! j!} \alpha^i (1-\alpha)^j, \quad 0 \leq \alpha \leq 1. \quad (21)$$

Die Kompaktifizierung von \mathbb{F} besteht also aus der Hinzunahme von Punkten α , die als Grenzwert von solchen Folgen erscheinen, bei denen der Anteil der x -Koordinate an der Koordinatensumme gegen α konvergiert.

5.3 Zwei zentrale Resultate

Die Übergangsfunktion p einer Markov-Kette $X = (X_n)_{n \in \mathbb{N}}$ mit Zustandsraum \mathbb{F} definiert einen linearen Operator auf dem Raum aller Abbildungen $f: \mathbb{F} \rightarrow \mathbb{R}$ durch

$$(pf)(x) := \sum_{y \in \mathbb{F}} p(x,y)f(y) \quad \text{für alle } x \in \mathbb{F}. \quad (22)$$

Da $p(x,y) > 0$ bei $x \in \mathbb{F}_n$ nur für y aus der endlichen Menge \mathbb{F}_{n+1} möglich ist, sind die in (22) auftauchenden Summen stets endlich. Das ‘overloading’ bei dem Symbol p lässt sich durch die Beobachtung rechtfertigen, dass sich $p(x,y)$ ergibt, wenn man den Operator auf die Indikatorfunktion der Menge $\{y\}$ anwendet. Die Eigenfunktionen zum Eigenwert 1, also die Funktionen $h: \mathbb{F} \rightarrow \mathbb{R}$, für die $ph = h$ gilt, nennt man harmonisch.

Die fast sichere Konvergenz der Pólya-Urne in der in Abschnitt 5.2 erarbeiteten Kompaktifizierung ergab sich direkt aus dem Konvergenzsatz für nicht-negative Martingale, konkret mit $h(X_n) = (X_{n,1} + 1)/(X_{n,1} + X_{n,2} + 1)$. Die Funktion h fiel dabei, wie man so sagt, vom Himmel. Eine erste wichtige Beobachtung auf dem Weg zu einem systematischeren Vorgehen ist die bereits durch die Symbolwahl angedeutete Beziehung: Ist h harmonisch, so erhält man mit der Markov-Eigenschaft für alle $x \in \mathbb{F}_n$

$$1_{\{x\}}(X_n) E[h(X_{n+1}) | \mathcal{F}_n] = \sum_{y \in \mathbb{F}} h(y)p(x,y) = h(x),$$

also $E[h(X_{n+1}) | \mathcal{F}_n] = h(X_n)$, und damit ein Martingal $(h(X_n), \mathcal{F}_n)_{n \in \mathbb{N}}$; umgekehrt ergibt sich aus dieser Gleichung, dass h harmonisch ist.

Offensichtlich bilden die harmonischen Funktionen einen linearen Raum, die nichtnegativen harmonischen Funktionen einen Kegel hierin, und diejenigen nichtnegativen harmonischen Funktionen h , die im Sinne von $h(e) = 1$ normiert sind, eine konvexe Menge. Die nächste wichtige Beobachtung besteht darin, dass die durch den erweiterten Martin-Kern definierten Funktionen $x \mapsto K(x, \alpha)$, $\alpha \in \partial\mathbb{F}$, alle in der letztgenannten Menge liegen. Insbesondere definiert jedes Wahrscheinlichkeitsmaß ν auf der Doob-Martin-Kompaktifizierung, das auf den Martin-Rand konzentriert ist, für das also $\nu(\partial\mathbb{F}) = 1$ gilt, durch

$$h(x) := \int_{\partial\mathbb{F}} K(x, \alpha) \nu(d\alpha) \quad \text{für alle } x \in \mathbb{F}$$

eine nichtnegative normierte harmonische Funktion. Das folgende erste fundamentale Resultat besagt, dass man *alle* solchen Funktionen auf diese Art erhält.

Satz 2 *Es sei \mathcal{H}_1 die Menge der nichtnegativen normierten harmonischen Funktionen einer kombinatorischen Markov-Kette mit Martin-Rand $\partial\mathbb{F}$ und erweitertem Martin-Kern K . Dann existiert zu jedem $h \in \mathcal{H}_1$ ein Wahrscheinlichkeitsmaß ν_h auf $\partial\mathbb{F}$ derart, dass gilt*

$$h(x) = \int_{\partial\mathbb{F}} K(x, \alpha) \nu_h(d\alpha) \quad \text{für alle } x \in \mathbb{F}. \quad (23)$$

Das zweite für uns zentrale Resultat behandelt die Konvergenz der Markov-Kette sowie strukturelle Aspekte, die mit dem Grenzwert zusammenhängen. Letztere machen maßgeblich von einer weiteren fundamentalen Idee von Doob Gebrauch: Ist h eine positive harmonische Funktion, so erhält man, wie man leicht nachrechnet, durch

$$p^{(h)}(x,y) := \frac{p(x,y)h(y)}{h(x)} \quad \text{für alle } x,y \in \mathbb{F}$$

eine neue Übergangsfunktion auf \mathbb{F} . Es sei $X^{(h)} = (X_n^{(h)})_{n \in \mathbb{N}}$ die zugehörige Markov-Kette, wieder mit Start in e . Offensichtlich ist mit X auch $X^{(h)}$ eine kombinatorische Markov-Kette zu \mathbb{F} . Der Martin-Kern K wird beim Übergang von X auf $X^{(h)}$ zu $K^{(h)}(x, y) = K(x, y)/h(x)$, also haben beide Markov-Ketten dieselbe Doob-Martin-Kompaktifizierung.

Wieder ist ein Blick auf die Pólya-Urne instruktiv: Mit der (extremalen) harmonischen Funktion $K(\cdot, \alpha)$ aus (21) ergibt sich

$$p^{(h)}((i, j), (i+1, j)) = \frac{\frac{i+1}{i+j+2} \frac{(i+j+2)!}{(i+1)!j!} \alpha^{i+1} (1-\alpha)^j}{\frac{(i+j+1)!}{i!j!} \alpha^i (1-\alpha)^j} = \alpha.$$

Man erhält also als h -transformierte Markov-Kette bei $h = K(\cdot, \alpha)$ die Nordost-Irrfahrt mit Parameter $\theta = \alpha$. Das folgende zweite fundamentale Resultat ist nun die angekündigte Verallgemeinerung von Satz 1.

Satz 3 *Es sei $X = (X_n)_{n \in \mathbb{N}}$ eine kombinatorische Markov-Kette mit Martin-Rand $\partial\mathbb{F}$. Dann gilt:*

- (a) *Mit $n \rightarrow \infty$ konvergiert X_n fast sicher gegen eine $\partial\mathbb{F}$ -wertige Zufallsgröße X_∞ .*
- (b) *Das Wahrscheinlichkeitsmaß $\nu_1 := \mathcal{L}(X_\infty)$ stellt die harmonische Funktion $h \equiv 1$ im Sinne von (23) dar.*
- (c) *Die bedingte Verteilung von X unter $X_\infty = \alpha$ stimmt überein mit der Verteilung der h -transformierten Markov-Kette $X^{(h)}$ mit $h = K(\cdot, \alpha)$.*

Es bleibt die Diskussion der Optimalität dieser Konstruktion.

5.4 ... und eine Ergänzung im Raum-Zeit-Fall

Ginge es nur um eine Kompaktifizierung, in der X_n mit $n \rightarrow \infty$ fast sicher konvergiert, so wäre man mit der Einpunktkompaktifizierung schneller am Ziel: Man fügt dem Zustandsraum E einen Punkt e_∞ hinzu, nennt die Komplemente endlicher Teilmengen von E offen, und erhält $X_n \rightarrow e_\infty$ mit Wahrscheinlichkeit 1, da eine transiente Markov-Kette jede endliche Teilmenge von E irgendwann ‘für immer’ verlässt. Diese Kompaktifizierung ist natürlich nicht besonders interessant.

Es ist klar, dass eine Kompaktifizierung unter der Maßgabe der genannten Konvergenz umso besser ist, je ‘feiner’ sie ist. Gibt es eine in diesem Sinn optimale Erweiterung? Für die Anhang A.1 skizzierte (volle) Stone-Čech-Kompaktifizierung ist eine solche Universalität bekannt, sie kann über Faktorisierungen definiert werden. In Analogie dazu, und auch im Hinblick auf Anwendungen, betrachten wir Funktionale $Y_n = \Phi(X_n)$, $\Phi : \mathbb{F} \rightarrow \mathbb{R}$, der kombinatorischen Markov-Kette $(X_n)_{n \in \mathbb{N}}$. Lässt sich Konvergenz von Y_n gegen Y_∞ , sofern gegeben und wie immer in diesem Abschnitt fast sicher und mit $n \rightarrow \infty$, über die Konvergenz $X_n \rightarrow X_\infty$ in der Doob-Martin-Kompaktifizierung in dem Sinn darstellen, dass Y_∞ als Funktion von X_∞ geschrieben werden kann? Dies ist nach einem maßtheoretischen Faktorisierungssatz immer dann der Fall, wenn X_∞ die terminale σ -Algebra

$$\mathcal{T} := \bigcap_{n=1}^{\infty} \sigma(\{X_m : m \geq n\})$$

zu der Folge $X = (X_n)_{n \in \mathbb{N}}$ erzeugt, denn Y_∞ ist \mathcal{T} -messbar und damit, da die Kompaktifizierung einen polnischen Raum liefert, eine Funktion von X_∞ (modulo Nullmengen).

Der folgende Satz ist nun die Ergänzung, aus der sich bei kombinatorischen Markov-Ketten die Optimalität der Doob-Martin-Kompaktifizierung im genannten Sinn ergibt.

Satz 4 *In der Situation von Satz 3 erzeugt X_∞ die terminale σ -Algebra \mathcal{T} zu X .*

Wir beschreiben kurz, wie dies mit der Raum-Zeit-Eigenschaft zusammenhängt. Ganz allgemein lässt sich eine Verbindung zwischen den beschränkten, \mathcal{T} -messbaren Zufallsgrößen Z und den Raum-Zeit-harmonischen Funktionen

$$h : \mathbb{N} \times \mathbb{F} \rightarrow \mathbb{R}, \quad h(n+1, x) = \sum_{y \in \mathbb{F}} p(x, y) h(n, y) \quad \text{für alle } n \in \mathbb{N}, x \in \mathbb{F}$$

mit Hilfe der Markov-Eigenschaft herstellen durch

$$Z \mapsto h, \quad h(n, x) := E[Z | \mathcal{F}_n] \quad \text{für alle } n \in \mathbb{N}, x \in \mathbb{F}.$$

In der Gegenrichtung kann man Z aus h zurückgewinnen (formal muss man zu Äquivalenzklassen von Zufallsgrößen übergehen, die außerhalb einer Nullmenge übereinstimmen), indem man verwendet, dass $(h(n, X_n), \mathcal{F}_n)_{n \in \mathbb{N}}$ ein beschränktes Martingal ist, das nach Konstruktion den Grenzwert Z hat. Bei kombinatorischen Ketten ist nun n eine Funktion von $x \in \mathbb{F}$, d.h. jede beschränkte Raum-Zeit-harmonische Funktion entspricht genau einer beschränkten harmonischen Funktion. Andererseits erhält man eine solche Darstellung der beschränkten harmonischen Funktionen durch Satz 2 zusammen mit Teil (c) von Satz 3, nun als Raum der messbaren und fast sicher beschränkten reellwertigen Abbildungen auf dem Rand $\partial\mathbb{F}$, durch $h \mapsto \Phi$, wobei Φ eine Dichte von ν_h bzgl. ν_1 ist. Dies liefert $Z = \Phi(X_\infty)$, wieder modulo Nullmengen. Im Vergleich zu der allgemeinen, auf einem maßtheoretischen Existenzsatz beruhenden Argumentation ist dies ein geradezu konstruktives Verfahren.

6 Ergebnisse

Wir wollen nun die im vorangegangenen Abschnitt skizzierte Theorie auf die in Abschnitt 4 beschriebenen kombinatorischen Markov-Ketten anwenden. Ganz allgemein gilt, dass eine solche Anwendung, insbesondere die Bestimmung des Martin-Kerns, schwierig sein kann, dass man das Resultat meistens auch direkter erhalten kann, und dass die Kompaktifizierung möglicherweise ein wenig brauchbares oder schlichtweg uninteressantes Resultat liefert. Dessen ungeachtet führt die allgemeine Theorie auf eine generelle Perspektive, und kann häufig als Ausgangspunkt für spezielle Überlegungen dienen.

Zunächst betrachten wir in Abschnitt 6.1 Ketten, bei denen sich frühere Zustände aus dem aktuellen Zustand rekonstruieren lassen. Dies betrifft insbesondere die Prozesse zu \mathbb{S} und \mathbb{A} . Die in Abschnitt 5.1 beschriebene Vorgehensweise kann in dieser Situation recht leicht umgesetzt werden. Anschließend behandeln wir die Prozesse zu \mathbb{Y} , \mathbb{B} und \mathbb{E} , die alle diese Eigenschaft nicht haben; beispielsweise lässt sich dem n -ten Suchbaum i.d.R. nicht entnehmen, welcher Knoten als letzter hinzugenommen wurde. In den Abschnitten 6.2-6.4 werden nur einige exemplarische Resultate aufgeführt, auf Beweise müssen wir leider verzichten. Natürlich sind die betreffenden Vervollständigungen nur bis auf Homöomorphie eindeutig. Wir geben jeweils einen für Anwendungen geeigneten Repräsentanten an. Einen Teil der Resultate illustrieren wir mit simulierten Pfaden und verweisen zur Datenherkunft wieder auf Anhang A.3.

6.1 Ketten mit vollständigem Gedächtnis

Es sei wieder \mathbb{F} eine kombinatorische Markov-Kette mit Größenfunktion ϕ , partieller Ordnung \leq und einzigem Element e der Größe 1. Wir betrachten den gerichteten Graphen mit Knotenmenge \mathbb{F} und Kanten

$$E(\mathbb{F}) := \{(x, y) \in \mathbb{F} \times \mathbb{F} : x \leq y, \phi(y) = \phi(x) + 1\}$$

und nennen eine Sequenz $(y_1, \dots, y_k) \in \mathbb{F}^k$ mit $(y_i, y_{i+1}) \in E(\mathbb{F})$ für $i = 1, \dots, k-1$ einen Pfad von y_1 nach y_k . Einige der in Abschnitt 3 eingeführten kombinatorischen Familien führen so auf einen Baum mit Wurzelknoten e : Die Kanten zeigen von der Wurzel weg, und es gibt zu jedem $x \in \mathbb{F}$ genau einen Pfad von e nach x .

Ist nun $X = (X_n)_{n \in \mathbb{N}}$ eine kombinatorische Markov-Kette zu einer solchen Familie, so lassen sich aus einem Zustand nicht nur der zugehörige Zeitpunkt, sondern wegen (12) auch alle früheren Zustände ablesen. Wir sagen dann, dass der Prozess vollständiges Gedächtnis hat. Offensichtlich ist dies stets bei den Familien Σ^* (siehe S.6), \mathbb{A} und \mathbb{S} der Fall; auch bei den anderen Familien führen geeignete Bewertungen auf diese Situation.

Für Ketten mit vollständigem Gedächtnis lässt sich die Doob-Martin-Kompaktifizierung leicht bestimmen, sie stimmt mit einer aus der Graphentheorie bekannten Kompaktifizierung überein. Zunächst einmal reduzieren sich Übergangswahrscheinlichkeiten auf simple Produkte, d.h. für alle $n \in \mathbb{N}$, $x_n \in \mathbb{F}_n$ gilt

$$P(X_n = x_n) = \prod_{i=1}^{n-1} p(x_i, x_{i+1}),$$

wobei x_i den i -ten Wert auf dem eindeutigen Pfad von $x_1 = e$ nach x_n bezeichnet. Für den Martin-Kern erhält man so für alle $x_m \in \mathbb{F}_m$, $x_n \in \mathbb{F}_n$ mit $m < n$

$$\begin{aligned} K(x_m, x_n) &= \frac{P(X_m = x_m, X_n = x_n)}{P(X_m = x_m)P(X_n = x_n)} \\ &= \frac{\prod_{i=1}^{n-1} p(x_i, x_{i+1})}{\prod_{i=1}^{m-1} p(x_i, x_{i+1}) \prod_{i=1}^{n-1} p(x_i, x_{i+1})} \\ &= \frac{1}{\prod_{i=1}^{m-1} p(x_i, x_{i+1})}, \end{aligned}$$

wenn x_m auf dem Pfad von e nach x_n liegt; ist dies nicht der Fall, so gilt $K(x_m, x_n) = 0$. Aus dieser sehr einfachen Abhängigkeit vom zweiten Argument folgert man leicht, dass $(K(x, y_n))_{n \in \mathbb{N}}$ genau dann für alle $x \in \mathbb{F}$ gegen einen positiven Grenzwert konvergiert (sogar in einem sehr starken Sinn), wenn die Folge $(y_n)_{n \in \mathbb{N}}$ schließlich in einem Strahl des Baumes enthalten ist. Ein solcher besteht im vorliegenden Fall aus einer Folge $(z_n)_{n \in \mathbb{N}}$ mit $z_n \in \mathbb{F}_n$, $z_n \leq z_{n+1}$ für alle $n \in \mathbb{N}$. Insbesondere ist die Doob-Martin-Kompaktifizierung des Zustandsraums \mathbb{F} der Markov-Kette gleich der Endenkompaktifizierung des Baumes $(\mathbb{F}, E(\mathbb{F}))$.

Bei einer Markov-Kette kommen im Vergleich zur graphentheoretischen Situation zwei Fragen hinzu: Wie sieht die Verteilung von X_∞ aus, und was ergibt sich als bedingte Verteilung von X unter $X_\infty = y \in \partial\mathbb{F}$? Auch diese beiden Fragen haben bei Ketten mit vollständigem Gedächtnis einfache Antworten. Zunächst einmal macht man sich leicht klar, dass durch

$$Q(\{y \in \partial\mathbb{F} : x_n \leq y\}) = \prod_{i=1}^{n-1} p(x_i, x_{i+1})$$

mit x_1, \dots, x_n wie oben und der in naheliegender Weise auf den Rand fortgesetzten partiellen Ordnung ein Wahrscheinlichkeitsmaß auf diesem Rand definiert wird, und dass $Q = \mathcal{L}(X_\infty)$ gilt. Als bedingte Verteilung von X unter $X_\infty = y \in \partial\mathbb{F}$, $y = (y_n)_{n \in \mathbb{N}}$, ergibt sich die deterministische Bewegung entlang dieses Strahls, mit Übergangswahrscheinlichkeiten

$$p_y(y_n, y_{n+1}) = 1 \quad \text{für alle } n \in \mathbb{N}.$$

Also: ‘Der Limes ist die Folge’, und ein vollständiges Gedächtnis hat auch die rechts durch X_∞ abgeschlossene Kette.

In den nachfolgenden Abschnitten werden wir interessantere Grenzwerte erhalten, aber die einzelnen Schritte nicht mehr ausführen können.

6.2 Asymptotik zur Zyklenkette

Es sei

$$\text{WS}_\infty := \left\{ (\alpha_i)_{i \in \mathbb{N}} \in [0, 1]^{\mathbb{N}} : \sum_{i \in \mathbb{N}} \alpha_i \leq 1 \right\}$$

die Menge der Folgen nichtnegativer Zahlen mit Summe kleiner oder gleich 1. Jede solche Folge definiert durch

$$Q(\{i\}) = \alpha_i \quad \text{für alle } i \in \mathbb{N}, \quad Q(\{\infty\}) = 1 - \sum_{i \in \mathbb{N}} \alpha_i,$$

ein Wahrscheinlichkeitsmaß Q auf $\mathbb{N}_\infty := \mathbb{N} \cup \{\infty\}$; man kann WS_∞ also als unendlichdimensionalen Wahrscheinlichkeitssimplex auffassen. Ausgehend von einer Folge $(U_i)_{i \in \mathbb{N}}$ von unabhängigen, $\text{unif}(0, 1)$ -verteilten Zufallsvariablen definieren wir nun eine Zufallsgröße $Z = (Z_i)_{i \in \mathbb{N}}$ mit Werten in WS_∞ durch

$$Z_1 := U_1, \quad Z_i := U_i \prod_{j=1}^{i-1} (1 - U_j) \quad \text{für alle } i > 1.$$

Diese Konstruktion ist auch als ‘stick breaking’ bekannt: U_1 ist das linke Bruchstück, das rechte Teilstück wird wieder an einer zufällig und gleichverteilten Stelle durchgebrochen etc.; man erhält so ein zufälliges Wahrscheinlichkeitsmaß Q auf \mathbb{N}_∞ , das dem Wert $\{\infty\}$ mit Wahrscheinlichkeit 1 den Wert 0 zuordnet.

Jedes Element von WS_∞ lässt sich absteigend sortieren; auf diese Art erhalten wir eine Abbildung von WS_∞ nach WS_∞ , die wir mit Ψ^\downarrow bezeichnen.

Satz 5 *Es sei $(X_n)_{n \in \mathbb{N}}$ die kombinatorische Markov-Kette zu \mathbb{Y} mit den durch (15) gegebenen Übergangswahrscheinlichkeiten. Dann gilt:*

- (a) *Der Martin-Rand $\partial_{\text{mk}} \mathbb{Y}$ ist das Bild von WS_∞ unter Ψ^\downarrow , wobei Konvergenz einer Folge $(\lambda(n))_{n \in \mathbb{N}}$ gegen $\alpha \in \partial_{\text{mk}} \mathbb{Y}$ bedeutet, dass gilt:*

$$\lim_{n \rightarrow \infty} \frac{\lambda_i(n)}{n} = \alpha_i \quad \text{für alle } i \in \mathbb{N}.$$

- (b) *Die Verteilung des fast sicheren Grenzwertes X_∞ der Markov-Kette stimmt überein mit der Verteilung von $\Psi^\downarrow(Z)$.*

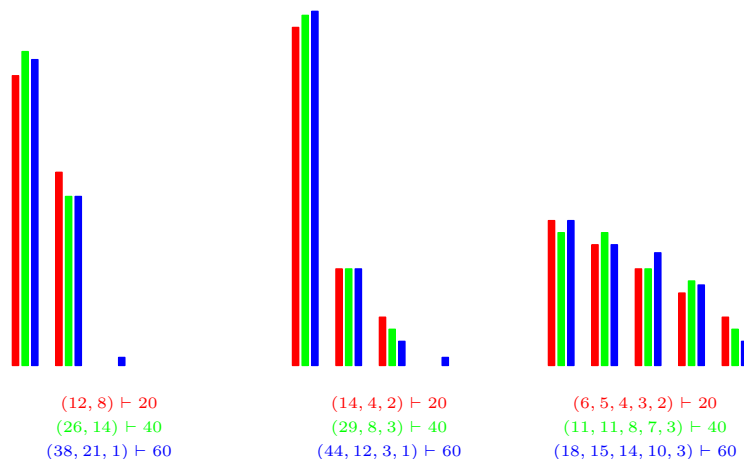


Abb. 7 Drei Pfade der Zyklenkette zu den Zeitpunkten $n = 20$ (grün), $n = 40$ (rot) und $n = 60$ (blau)

Dividiert man also die Zyklenlängen durch n , die Summe der Zyklenlängen, so erhält man fast sichere Konvergenz gegen einen nicht konstanten Grenzwert; Abbildung 7 illustriert dies. Dabei hat beispielsweise der längste auf diese Art standardisierte Zykel dieselbe Verteilung wie das längste Stück in der stick-breaking-Prozedur. Man kann zeigen, dass diese Zufallsgröße Erwartungswert $0.6243\dots$ hat, d.h. im Mittel enthält der längste Zykel einer zufällig und gleichverteilten Permutation fast zwei Drittel der Elemente der (großen) Grundmenge.

6.3 Asymptotik zur BST-Kette

Im Abschnitt 3.3 haben wir binäre Bäume als Teilmengen von $\mathbb{V} = \{0, 1\}^*$ betrachtet. Zu einem Graphen wird \mathbb{V} durch Hinzunahme der Kanten (\bar{u}, u) , $u \in \mathbb{V}$, $u \neq \emptyset$, wobei $\bar{u} = (u_1, \dots, u_{k-1})$ den direkten Vorfahren zu $u = (u_1, \dots, u_k)$ bezeichnet. In der Tat erhält man so einen Graphen ohne Kreise, also einen Baum im Sinne der Graphentheorie, wobei das leere Wort \emptyset als Wurzelknoten fungiert.

Ordnet man den Kanten (\bar{u}, u) das Gewicht $2^{-|u|}$ zu, so erhält man eine Metrik $d_{\mathbb{V}}$ auf \mathbb{V} , indem man zur Ermittlung von $d_{\mathbb{V}}(u, v)$ die Kantenlängen auf dem eindeutigen Weg von u nach v zusammenzählt. Die zugehörige Vervollständigung $(\bar{\mathbb{V}}, d_{\mathbb{V}})$ ist ein kompakter und separabler topologischer Raum, der mit der uns aus Abschnitt 6.1 bekannten Endenkompaktifizierung von \mathbb{V} übereinstimmt. Mit $\partial\mathbb{V} = \bar{\mathbb{V}} \setminus \mathbb{V}$ bezeichnen wir wieder den Rand der Kompaktifizierung. Schließlich sei noch $\mathcal{M}(S)$, für einen beliebigen separablen metrischen Raum (S, d) , die Menge der Wahrscheinlichkeitsmaße auf der σ -Algebra der Borel-Mengen von S . Mit der Topologie der schwachen Konvergenz, in der $\lim_{n \rightarrow \infty} \mu_n = \mu$ für

$$\lim_{n \rightarrow \infty} \int f d\mu_n = \int f d\mu \quad \text{für alle stetigen und beschränkten } f : S \rightarrow \mathbb{R}$$

steht, wird $\mathcal{M}(S)$ zu einem abermals separablen topologischen Raum. Die Bäume x aus \mathbb{B} können wir als Elemente von $\mathcal{M}(\bar{\mathbb{V}})$ auffassen, wenn wir jedem x die Gleichverteilung auf x zuordnen.

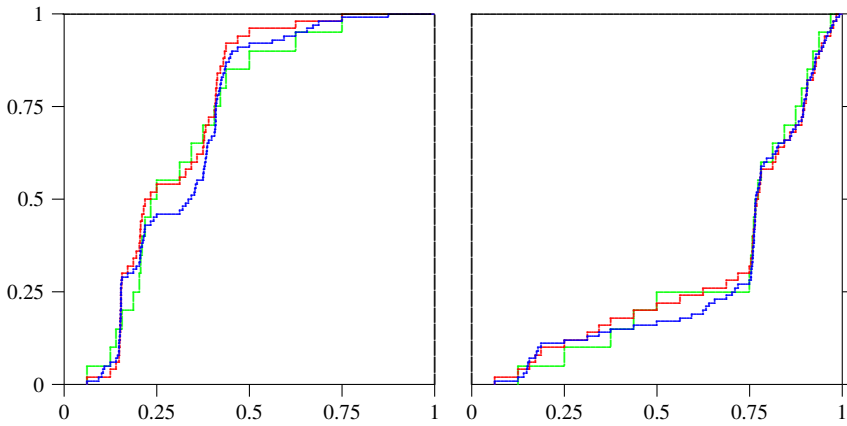


Abb. 8 Werte der Knotenverteilungsfunktion zu zwei Realisierungen der BST-Kette zu den Zeitpunkten $n = 20$ (grün), $n = 50$ (rot) und $n = 100$ (blau)

Satz 6 Es sei $X = (X_n)_{n \in \mathbb{N}}$ die kombinatorische Markov-Kette zu \mathbb{B} mit den durch (14) gegebenen Übergangswahrscheinlichkeiten. Dann ist $\mathcal{M}(\partial\mathbb{V})$ mit $\mathbb{V} = \{0,1\}^*$, versehen mit der Topologie der schwachen Konvergenz, der Martin-Rand der Markov-Kette X , und eine Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in \mathbb{B}_n$ für alle $n \in \mathbb{N}$ konvergiert genau dann gegen $\mu \in \mathcal{M}(\partial\mathbb{V})$, wenn gilt:

$$\lim_{n \rightarrow \infty} \frac{1}{n} \#\{v \in x_n : u \leq v\} = \mu(\{v \in \partial\mathbb{V} : u \leq v\}) \quad \text{für alle } u \in \mathbb{V}.$$

Abbildung 8 illustriert dies. Dabei haben wir nicht die Bäume selbst gezeichnet, sondern die jeweilige ‘Knotenverteilungsfunktion’: Diese ordnet jeder binär-rationalen Zahl mit zugehörigem Knoten $u \in \{0,1\}^*$ die relative Anzahl der Knoten des Baumes zu, deren Wert unter der in (11) definierten Funktion β den von u nicht übersteigt. Man sieht beispielsweise, dass bei den Bäumen im linken Beispiel etwa drei Viertel der Knoten im linken Teilbaum liegen, wogegen der entsprechende Wert in den rechts gezeigten Bäumen weniger als ein Viertel ist.

6.4 Asymptotik zur RSK-Kette

Wir wollen die Gemeinsamkeiten und Unterschiede zu den in den vorangegangenen beiden Unterabschnitten besprochenen Fällen herausstellen und betrachten hierzu noch einmal die Zyklenkette, nun unter dem für die Baumketten entwickelten Gesichtspunkt: Auch hier gibt es wieder einen zugrunde liegenden Digraphen \mathbb{D} , jetzt mit Knotenmenge $\mathbb{V} := \mathbb{N} \times \mathbb{N}$ und Kanten

$$E(\mathbb{N} \times \mathbb{N}) := \{(i, j), (i+1, j) : i, j \in \mathbb{N}\} \cup \{(i, j), (i, j+1) : i, j \in \mathbb{N}\}.$$

Allerdings erhält man mit dieser Konstruktion keinen Baum mehr; beispielsweise führen die Pfade $((1,1), (1,2), (2,2))$ und $((1,1), (2,1), (2,2))$ beide von $(1,1)$ nach $(2,2)$. An die Stelle der Präfixstabilität tritt bei den Elementen von \mathbb{D} nun die Forderung, dass eine Teilmenge von \mathbb{V} mit jedem $v \in \mathbb{V}$ auch die Knoten aller Pfade von der ‘Wurzel’ $(1,1)$ nach

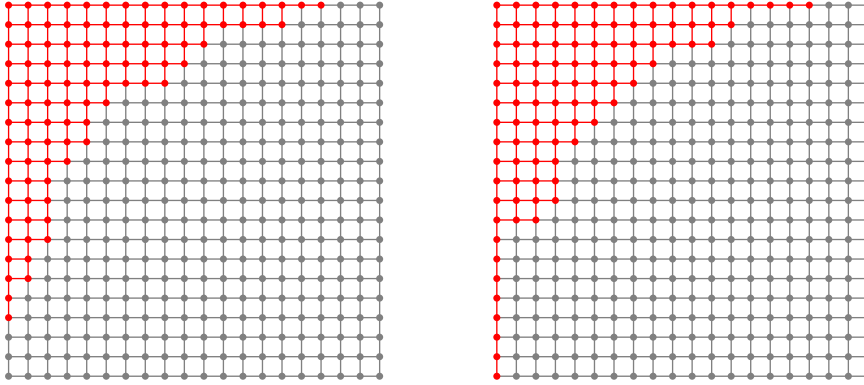


Abb. 9 Zwei Werte der RSK-Kette zum Zeitpunkt $n = 100$ (rot)

v enthält. Sowohl die Zyklenkette als auch die RSK-Kette werden so zu zufällig wachsenden pfadstabilen Teilmengen des zugrunde liegenden Digraphen, wobei in jedem Schritt ein weiterer Knoten eingemeindet wird. Abbildung 9 zeigt zwei simulierte Zustände der RSK-Kette zum Zeitpunkt $n = 100$; man vergleiche dies mit Abbildung 3.

Unter dem Gesichtspunkt des sich auffüllenden Digraphen lässt sich der in Satz 5 erhaltene Rand der Zyklenkette als Teilmenge der Menge der Wahrscheinlichkeitsmaße auf dem Rand einer geeignet gewählten Kompaktifizierung von \mathbb{D} interpretieren: Eine Folge in \mathbb{V} konvergiere, wenn die zweite Koordinate einfriert. Als Rand erhält man $\partial\mathbb{D} = \{\infty\} \times (\mathbb{N} \cup \{\infty\})$, und die Wahrscheinlichkeitsmaße aus $\partial_{\text{zyk}}\mathbb{Y}$ werden durch ihre antitone Massenfunktion $i \rightarrow \alpha_i$, $i \in \mathbb{N}$, beschrieben, wobei $1 - \sum_{i \in \mathbb{N}} \alpha_i$ die Wahrscheinlichkeit von $\{\infty\}$ ist.

Ein augenfälliger Unterschied zwischen Zyklenkette und RSK-Kette ist das Verhalten bei Vertauschung von x - und y -Achse. Formal können wir jedem $\lambda \in \mathbb{Y}$ eine konjugierte Partition $\lambda' \in \mathbb{Y}$ zuordnen durch

$$\lambda'_i := \#\{j \in \mathbb{N} : \lambda_j \geq i\},$$

für alle $i \in \mathbb{N}$, für die die betreffende Menge nicht leer ist. Anschaulich erhält man λ' aus λ durch Spiegelung an der Geraden $t \mapsto (t, -t)$, oder auch als kumulative Summe zum Zähltyp aus (17). Zu der Partition $\lambda = (5, 3, 3)$ aus Abbildung 2 ergibt sich so $\lambda' = (3, 3, 3, 1, 1)$. Man macht sich leicht klar, dass $H(\lambda) = H(\lambda')$ gilt, siehe (18), insbesondere sind die Randverteilungen der RSK-Kette invariant unter dieser Konjugation. Dies muss sich natürlich in den hierauf aufbauenden Strukturen wiederfinden.

Auf der Basis dieser Überlegungen liegt es nahe, als passenden Rand von \mathbb{V} die Menge

$$\partial\mathbb{V} = (\{\infty\} \times \mathbb{N}) \cup (\mathbb{N} \times \{\infty\}) \cup \{(\infty, \infty)\} \quad (24)$$

und als RSK-Rand von \mathbb{Y} wieder eine Menge von Wahrscheinlichkeitsmaßen auf $\partial\mathbb{V}$ zu vermuten. Konkret sei $\partial_{\text{RSK}}\mathbb{Y}$ die Menge der Paare (α, β) von antitonen Folgen $\alpha = (\alpha_i)_{i \in \mathbb{N}}$, $\beta = (\beta_i)_{i \in \mathbb{N}}$ nichtnegativer reeller Zahlen mit $\sum_{i=1}^{\infty} \alpha_i + \sum_{i=1}^{\infty} \beta_i \leq 1$. Analog zu der obigen Interpretation bei Satz 5 betrachten wir α_i als Wahrscheinlichkeit von $\{(\infty, i)\}$, β_i als Wahrscheinlichkeit von $\{(i, \infty)\}$ und $1 - \sum_{i=1}^{\infty} \alpha_i - \sum_{i=1}^{\infty} \beta_i$ als Wahrscheinlichkeit von $\{(\infty, \infty)\}$.

Satz 7 Es sei $X = (X_n)_{n \in \mathbb{N}}$ die kombinatorische Markov-Kette zu \mathbb{Y} mit den durch (19) gegebenen Übergangswahrscheinlichkeiten. Dann ist $\partial_{\text{RSK}} \mathbb{Y}$ der zugehörige Martin-Rand, und eine Folge $(\lambda(n))_{n \in \mathbb{N}}$ mit $\lambda(n) \in \mathbb{Y}_n$ für alle $n \in \mathbb{N}$ konvergiert genau dann gegen $(\alpha, \beta) \in \partial_{\text{RSK}} \mathbb{Y}$, wenn gilt:

$$\lim_{n \rightarrow \infty} \frac{\lambda_i(n)}{n} = \alpha_i \text{ und } \lim_{n \rightarrow \infty} \frac{\lambda'_i(n)}{n} = \beta_i, \text{ für alle } i \in \mathbb{N}. \quad (25)$$

Es stellt sich heraus, dass in dieser Topologie X_n mit $n \rightarrow \infty$ gegen das Element $\alpha = \beta \equiv 0$ des Randes konvergiert. Die Vervollständigung des Zustandsraums liefert bei der RSK-Kette also nur die Trivialität der terminalen σ -Algebra. Anschaulich bedeutet (25), dass wir den Rand der roten Menge in Abbildung 9 entlang von horizontalen und vertikalen Geraden betrachten. Einen interessanteren fast sicheren Grenzwert erhält man, wenn man stattdessen die Geraden $t \mapsto (ut, -(1-u)t)$, $0 \leq u \leq 1$, verwendet und mit $n^{-1/2}$ statt n^{-1} normiert.

7 Bemerkungen und Literaturhinweise

Die Verbindung zwischen Potentialtheorie und Markov-Prozessen ist eines der großen klassischen Themen der Stochastik, sie hat die allgemeine Theorie der stochastischen Prozesse erheblich beeinflusst und ganz wesentlich vorangetrieben. Einer der Hauptakteure war der amerikanische Mathematiker Joseph Leo Doob (1910-2004), [Doo84] ist das Standardwerk zu diesem Thema. Auf Doob geht auch die diskrete Version zurück [Doo59]. Ein guter Einstieg ist, neben dieser Originalarbeit und [Dyn69], der einführende Artikel [Saw97].

Zu Markov-Ketten gibt es eine Reihe von klassischen Lehrbüchern, häufig mit dem Schwerpunkt auf rekurrenten Ketten; siehe auch die Arbeit [JKM04] in den *Semesterberichten*. In dem kürzlich erschienenen Lehrbuch [Woe09] werden auch transiente Markov-Ketten und deren Ränder im Detail behandelt. Darüberhinaus werden in der Monographie [Woe00] des gleichen Autors Markov-Ketten und ihre Ränder als Hilfsmittel bei der Analyse von Gruppen und Graphen verwendet.

Das Modell, das als ‘Pólya-Urne’ zum festen Bestandteil der Stochastikausbildung geworden ist, geht auf [EP23] zurück (dort wird der zweite Autor als Urheber der theoretischen Überlegungen genannt). In [BK64] wurde der zugehörige Martin-Rand bestimmt.

Das Standardwerk zur enumerativen Kombinatorik ist [Sta97, Sta99], im zweiten Band findet man Details zum RSK-Algorithmus. In [Pit06] geht es um kombinatorische stochastische Prozesse, insbesondere werden dort der CRP und seine Varianten behandelt. Zum BST-Algorithmus verweisen wir auf [Knu73] und [Mah92]. Auch zur Stochastik von Suchbäumen gibt es einen einführenden Artikel [Grü06] in den *Semesterberichten*.

Die Doob-Martin-Kompaktifizierung zu den binären Bäumen ist eines der Resultate in [EGW12]. Dort wird ein struktureller Aspekt ausgearbeitet, der beispielsweise die BST-Kette mit der Pólya-Urne verbindet. Auch für Resultate zu Ulam-Harris-Bäumen verweisen wir auf diese Arbeit.

Der originäre Beitrag zur Zyklenkette, motiviert durch Anwendungen in der Populationsgenetik, ist [Kin78]. Dort, wie auch in anderen in diesem Abschnitt genannten Arbeiten, ist der Ausgangspunkt nicht wie hier eine Übergangsfunktion p zum Schritt von n nach $n+1$, sondern eine ‘Rückwärtsübergangsfunktion’ \bar{p} . Bei einer kombinatorischen Markov-Kette $(X_n)_{n \in \mathbb{N}}$ ergibt sich diese durch

$$\bar{p}(y, x) := P(X_n = x | X_{n+1} = y) = \frac{P(X_n = x)}{P(X_{n+1} = y)} p(x, y).$$

(Es ist instruktiv, \bar{p} für die Nordost-Irrfahrt und die Pólya-Urne auszurechnen.) Schreibt man Q_n für die Verteilung von X_n , so erhält man für die Folge $(Q_n)_{n \in \mathbb{N}}$ die Konsistenzbedingung

$$Q_n(\{x\}) = \sum_{y \in \mathbb{F}_{n+1}} Q_{n+1}(\{y\}) \bar{p}(y, x) \quad \text{für alle } x \in \mathbb{F}_n. \quad (26)$$

Mit \bar{p} als Ausgangspunkt wird die Menge der Lösungsfolgen $(Q_n)_{n \in \mathbb{N}}$ zu (26) interessant. Es stellt sich heraus, dass diese Menge ein Choquet-Simplex ist, und dass die Menge der zugehörigen Extrempunkte im Wesentlichen übereinstimmt mit dem Martin-Rand zu p . In diesem Sinn entspricht Satz 5 dem Hauptresultat in [Kin78].

Bei der Zyklenkette und der BST-Kette lassen sich auch (im Stil von Teil (c) von Satz 1) die jeweiligen bedingten Markov-Ketten angeben: Man erhält im ersten Fall eine als ‘Kingman’s paintbox’ bekannte Konstruktion [Pit06, Chapter 2], im BST-Fall ergibt sich ein Zusammenhang zu einem unter dem Kürzel ‘DST’ (digital search tree) bekannten Algorithmus [EGW12, Section 5].

Wir haben Zahlenpartitionen aus der Sicht der probabilistischen Kombinatorik behandelt, was natürlich nicht dem historischen Ablauf entspricht. Der Originalbeweis zu Satz 7 in [Tho64] löst ein Problem aus der Darstellungstheorie der (unendlichen) symmetrischen Gruppe, und diese Verbindung zur Algebra ist nach wie vor von großer Aktualität. Das Buch [Ker03] enthält eine exzellente Einführung in diesen Ideenkreis.

Zwei subjektive Anmerkungen mögen diesen stark ‘kompaktifizierten’ Überblick abschließen. Zum einen ist die strukturelle Ähnlichkeit zwischen den hier betrachteten Prozessen bemerkenswert, speziell bei der Sichtweise der Prozesse als mengenwertige zufällige Mechanismen, die einen gewurzelten Graphen knotenweise auffüllen. Die Arbeit [EGW12] enthält eine Reihe von Resultaten zu einem bestimmten solchen Auffüllmechanismus. Zum anderen stellt sich die Frage nach Nutzen und Anwendbarkeit. In diesem Zusammenhang haben Konvergenzresultate auf der Ebene der diskreten Strukturen selbst ein erhebliches Potential, da sie als gemeinsamer Hintergrund für die in der Literatur in der Regel fallweise vorgenommene Analyse von Funktionalen dieser Strukturen dienen können; [Grü12] enthält Beispiele. Nichts ist praktischer als eine gute Theorie!

Danksagung Frau Prof. C. Bessenrodt, Herrn Dipl.-Math. K. Hagemann und Herrn Prof. W. Woess danke ich für ihre konstruktive Kritik und die Hinweise auf Fehler. Anlass für diese Arbeit war ein Vortrag bei einer Veranstaltung für Lehrer während der Stochastiktage 2012 in Mainz: Ich danke den Organisatoren für die Einladung, und nachträglich den Zuhörern für die angenehme Atmosphäre, ihre Geduld und ihr Interesse.

A Anhang

Wir erinnern in den beiden folgenden Abschnitten an zwei wichtige Konzepte aus der (mengentheoretischen) Topologie. Größtenteils ist dies Material aus dem Mathematik-Grundstudium; eine Wiederholung, ohne Beweise, aber mit Beispielen, sollte den Haupttext zugänglicher machen. Im Interesse der Nachvollziehbarkeit der Experimente, die den Illustrationen zu Grunde liegen, geben wir auch noch an, wie wir die Zahlen für unserer Simulationen erhalten haben.

A.1 Kompaktifizierung

Ausgangspunkt ist ein topologischer Raum (S, \mathcal{U}) , also eine Menge S mit einem System $\mathcal{U} \subset \mathcal{P}(S)$ von offenen Teilmengen. Man nennt ein Paar $(\phi, (S', \mathcal{U}'))$, bestehend aus einem weiteren topologischen Raum (S', \mathcal{U}') und einer injektiven und stetigen Abbildung $\phi : S \rightarrow S'$, eine Kompaktifizierung von (S, \mathcal{U}) , wenn das Bild von S unter ϕ dicht liegt in S' und unter der Spurtopologie homöomorph zum Ausgangsraum ist, und

wenn, natürlich, (S', \mathcal{U}') ein kompakter topologischer Raum ist. Das klassische Beispiel ist der Übergang von $S = \mathbb{N}_0$, versehen mit der diskreten Topologie $\mathcal{U} = \mathcal{P}(S)$, zu $S' := \mathbb{N}_0 \cup \{\infty\}$ mit $\mathcal{U}' := \mathcal{U} \cup \{A \cup \{\infty\} : A \subset \mathbb{N}_0, \#(\mathbb{N}_0 \setminus A) < \infty\}$ und $\phi(n) = n$. Dies ist ein Spezialfall der auch zu Beginn von Abschnitt 5.4 erwähnten Alexandroffschen Einpunktkompaktifizierung; sie liefert einen kompakten Raum, da der Ausgangsraum lokalkompakt ist.

Bei der Stone-Čech-Kompaktifizierung hat man als weiteren Baustein eine Familie $\mathcal{F} = \{f_i : i \in I\}$ von stetigen Funktionen $f_i : S \rightarrow [0, 1]$, die die Punkte von S trennt, d.h.

$$\forall x, y \in S \exists i \in I : f_i(x) \neq f_i(y). \quad (27)$$

Dann lässt sich S (im hier interessierenden diskreten Fall) durch

$$\phi : S \rightarrow [0, 1]^{\mathcal{F}}, \quad x \mapsto (f \mapsto f(x)),$$

injektiv in die Menge aller Abbildungen von S in $[0, 1]$ einbetten. Versieht man diese Menge mit der Produkttopologie, also der größten Topologie, unter der diese Einsetzabbildungen stetig sind, so erhält man nach dem Satz von Tychonov einen kompakten Raum, mit dem Abschluss von $\phi(S)$ hierin also eine Kompaktifizierung. In dieser Kompaktifizierung lassen sich die Funktionen $f_i, i \in I$, stetig fortsetzen. In Vorlesungen zur mengentheoretischen Topologie betrachtet man meistens das System *aller* stetigen Funktionen, und (27) wird dann zu einer Bedingung an den Ausgangsraum; es ist bekannt, dass diese ‘volle’ Stone-Čech-Kompaktifizierung zu recht unhandlichen Räumen führen kann.

A.2 Vervollständigung

Es sei (S, d) ein metrischer Raum. Bekanntlich heißt $(x_n)_{n \in \mathbb{N}} \in S^{\mathbb{N}}$ eine Cauchy-Folge, wenn

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n, m \geq n_0 : d(x_n, x_m) < \varepsilon$$

gilt. Auf der Menge der Cauchy-Folgen erhält man durch

$$(x_n)_{n \in \mathbb{N}} \sim (y_n)_{n \in \mathbb{N}} :\iff \lim_{n \rightarrow \infty} d(x_n, y_n) = 0$$

eine Äquivalenzrelation. In die Menge \bar{S} der zugehörigen Äquivalenzklassen lässt sich S injektiv einbetten durch $x \mapsto (x_n)_{n \in \mathbb{N}}$, wobei $x_n = x$ für alle $n \in \mathbb{N}$. Die Metrik d kann auf \bar{S} fortgesetzt werden; nach Konstruktion ist dieser neue metrische Raum vollständig. Ist der Ausgangsraum totalbeschränkt, d.h. gilt

$$\forall \varepsilon > 0 \exists S_\varepsilon \subset S, \#S_\varepsilon < \infty \forall x \in S \exists y \in S_\varepsilon : d(x, y) < \varepsilon, \quad (28)$$

so ist die Vervollständigung als topologischer Raum kompakt. Das klassische Beispiel für diese Prozedur ist der Übergang von den rationalen Zahlen $S = \mathbb{Q}$, versehen mit dem euklidischen Abstand, zu den reellen Zahlen $\bar{S} = \mathbb{R}$; Bedingung (28) ist in dieser Situation allerdings nicht erfüllt.

Wir schließen uns den in der Mathematik gebräuchlichen Identifizierungsgewohnheiten an und benutzen in der Regel dasselbe Symbol für das Bild $y = \phi(x)$ in der Einbettung und das Argument x aus dem Ausgangsraum (eine Gewohnheit, die man in der praktischen Informatik, konkret beim Programmieren, wieder mühsam entlernen muss).

Bei topologischen Räumen (S, \mathcal{U}) verwenden wir im Zusammenhang mit maßtheoretischen Aussagen stets die Borelsche σ -Algebra $\sigma(\mathcal{U})$.

A.3 Die Daten

Bekanntlich sagt ein Bild mehr als 1000 Worte, und auch im Zusammenhang mit zufälligen diskreten Strukturen können Visualisierungen sehr hilfreich sein. Natürlich ist in einem konkreten Bild nichts mehr zufällig; das zugrunde liegende Zufallsexperiment wurde vom Autor ausgeführt (oder auch nicht).

Universell verwendbar ist ein Zufallsexperiment, das beliebig oft unabhängig wiederholt werden kann und dabei in jedem Durchgang die Realisierung u einer auf dem Einheitsintervall gleichverteilten und von früheren Wiederholungen unabhängigen Zufallsvariablen U liefert. Um solche Daten zu erhalten, haben wir die Nachkommastellen in der Dezimalentwicklung von π in Zehnerblöcke zusammengefasst und mit 10^{-10} multipliziert. Die Abbildungen 1 und 7 basieren für die jeweils drei Einzelexperimente auf den so erhaltenen

Werten $u_1, \dots, u_{60}, u_{61}, \dots, u_{120}$ und u_{121}, \dots, u_{180} , bei den Abbildungen 8 und 9 haben wir im linken Teil die Werte u_{2k-1} und rechts u_{2k} , jeweils mit $k = 1, \dots, 100$, verwendet. Es ist immer noch nicht bekannt, ob π eine normale Zahl ist, ob also beispielsweise die relativen Häufigkeiten der zehn Ziffern in der Dezimaldarstellung von π gegen den Limes $1/10$ konvergieren. Die Behauptung, dass unsere Daten richtige Zufallszahlen seien, wäre also genausowenig gerechtfertigt wie beispielsweise beim üblicherweise verwendeten Output eines linearen Kongruenzgenerators – und sowieso völlig sinnlos.

Beim Übergang von der Folge der Pseudozufallszahlen zur jeweiligen Markov-Kette wurden die Darstellungen (6) und (7) benutzt, sowie die Algorithmen aus Abschnitt 3.5.

Literatur

- [BK64] David Blackwell and David Kendall, *The Martin boundary of Pólya's urn scheme, and an application to stochastic population growth*, J. Appl. Probability **1** (1964), 284–296. MR 0176518 (31 #790)
- [Doo59] J. L. Doob, *Discrete potential theory and boundaries*, J. Math. Mech. **8** (1959), 433–458; erratum 993. MR MR0107098 (21 #5825)
- [Doo84] ———, *Classical potential theory and its probabilistic counterpart*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 262, Springer-Verlag, New York, 1984. MR 731258 (85k:31001)
- [Dyn69] E. B. Dynkin, *The boundary theory of Markov processes (discrete case)*, Uspehi Mat. Nauk **24** (1969), no. 2 (146), 3–42. MR 0245096 (39 #6408)
- [EGW12] Steven N. Evans, Rudolf Grübel, and Anton Wakolbinger, *Trickle-down processes and their boundaries*, Electronic Journal of Probability **17** (2012), Paper no.1, 1–58.
- [EP23] P. Eggenberger and G. Pólya, *Über die Statistik verketteter Vorgänge*, Z. Ang. Math. Mech. **3** (1923), no. 4, 279–289.
- [Grü06] Rudolf Grübel, *Zufällige binäre Bäume: von der average-case Analyse zur Verteilungsasymptotik*, Math. Semesterber. **53** (2006), no. 2, 210–230. MR 2251038
- [Grü12] ———, *Search trees: metric aspects and strong limit theorems*, Tech. report, Leibniz Universität Hannover, 2012.
- [JKM04] Klaus Janssen, Hanns Klinger, and Reinhold Meise, *Markovketten: theoretische Grundlagen, Beispiele und Simulationen mit MAPLE*, Math. Semesterber. **51** (2004), no. 1, 69–93. MR 2076622 (2005f:60001)
- [Ker03] S. V. Kerov, *Asymptotic representation theory of the symmetric group and its applications in analysis*, Translations of Mathematical Monographs, vol. 219, American Mathematical Society, Providence, RI, 2003, Translated from the Russian manuscript by N. V. Tsilevich, With a foreword by A. Vershik and comments by G. Olshanski. MR 1984868 (2005b:20021)
- [Kin78] J. F. C. Kingman, *The representation of partition structures*, J. London Math. Soc. (2) **18** (1978), no. 2, 374–380. MR 509954 (80a:05018)
- [Knu73] Donald E. Knuth, *The art of computer programming. Volume 3, Sorting and searching*, Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1973. MR 0445948 (56 #4281)
- [Mah92] Hosam M. Mahmoud, *Evolution of random search trees*, John Wiley & Sons Inc., New York, 1992. MR 1140708 (93f:68045)
- [Pit06] J. Pitman, *Combinatorial stochastic processes*, Lecture Notes in Mathematics, vol. 1875, Springer-Verlag, Berlin, 2006, Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR MR2245368 (2008c:60001)
- [Saw97] Stanley A. Sawyer, *Martin boundaries and random walks*, Harmonic functions on trees and buildings (New York, 1995), Contemp. Math., vol. 206, Amer. Math. Soc., Providence, RI, 1997, pp. 17–44. MR MR1463727 (98k:60127)
- [Sta97] Richard P. Stanley, *Enumerative Combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 1997, With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original. MR 1442260 (98a:05001)
- [Sta99] ———, *Enumerative Combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999. MR 1676282 (2000k:05026)
- [Tho64] Elmar Thoma, *Die unzerlegbaren, positiv-definiten Klassenfunktionen der abzählbar unendlichen, symmetrischen Gruppe*, Math. Z. **85** (1964), 40–61. MR 0173169 (30 #3382)
- [Woe00] Wolfgang Woess, *Random walks on infinite graphs and groups*, Cambridge University Press, Cambridge, 2000. MR 1743100 (2001k:60006)
- [Woe09] ———, *Denumerable Markov chains. Generating functions, boundary theory, random walks on trees*, European Mathematical Society (EMS), Zürich, 2009. MR 2548569