

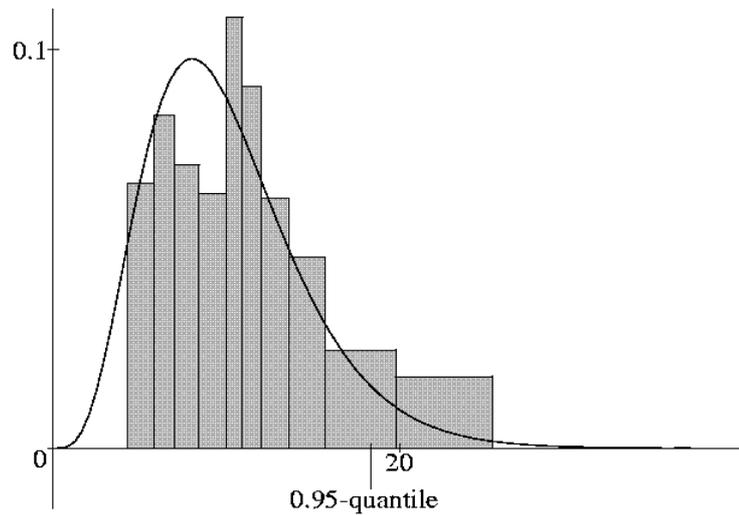


UNIVERSITÄT POTSDAM

Institut für Mathematik

Erzeugung gleichverteilter Stichproben von Lozeng-Teilungen mittels Kopplung von Markovketten

Diplomarbeit von Peter Keller



Mathematische Statistik und
Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Erzeugung gleichverteilter Stichproben
von Lozenge-Teilungen
mittels Kopplung von Markovketten

Diplomarbeit von Peter Keller

Institut für Mathematik
der Universität Potsdam

e-mail: Peter.Keller@mpikg.mpg.de

Preprint 2009/02

März 2009

Impressum

© **Institut für Mathematik Potsdam, März 2009**

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
Am neuen Palais, SansSouci,
14469 Potsdam

Telefon:
Fax: +49-331-977 1500
E-mail: +49-331-977 1578
neisse@math.uni-potsdam.de

ISSN 1613-3307

Universität Potsdam
Mathematisch-Naturwissenschaftliche
Fakultät
Institut für Mathematik

Diplomarbeit

Erzeugung gleichverteilter Stichproben
von Lozeng-Teilungen mittels Kopplung
von Markovketten

vorgelegt von
Peter Keller
Februar 2009

Gutachter
Dr. Pierre Yves Louis
Prof. Dr. Sylvie Roelly

Widmung

Diese Arbeit ist meiner zu früh verstorbenen Mutter, Roswitha Keller, gewidmet.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Definition einer Lozengeteilung	1
1.2	Verbindung zu Modellen der statistischen Physik	3
1.3	Gliederung der Arbeit	5
1.4	Eigenanteil	5
2	Eigenschaften der Lozengeteilungen	7
2.1	vollständige Überdeckbarkeit eines Gebietes	7
2.2	Darstellung von Lozengeteilungen durch deBruijn-Pfade	12
2.3	Anzahl Teilungen konvexer hexagonaler Regionen	14
2.3.1	Lemma von Gessel-Viennot	14
2.3.2	Formel von MacMahon	18
2.4	Wachstumsverhalten der Anzahl Teilungen	22
3	Coupling From The Past (CFTP)	28
3.1	Markovketten als M-Folgen und Notation	28
3.2	Kopplung von Markovketten mit diskretem Zustandsraum	31
3.3	Der Coupling-From-The-Past-Algorithmus	35
3.4	Monotonie und Coupling From The Past	37
4	Simulation von uniform verteilten Lozengeteilungen	40
4.1	Definition der Dynamik als zufällige, iterierte Funktion und Kopp- lung	40
4.2	Monotonie von \mathcal{M}_{loz}	44
4.3	Konnektivität des Zustandsraums	46
4.4	\mathcal{M}_{loz} ist schnell mischend	48
4.5	Verteilung der Koaleszenzzeit	53
5	Programmlisting und Diskussion der Implementierung	54
6	Ausblick	62
6.1	Arctic-Circle-Phänomen	62
6.2	Oktagonale Teilungen	64
A	Anhang	69
B	Eidesstattliche Erklärung	74

1 Einleitung

Aix, die historische Hauptstadt der Provence, glänzt seit dem 16. Jahrhundert durch eine besondere konditorische Spezialität, den Calissons. Man sagt sie seien zur Hochzeit des Herzogs René von Anjou mit Isabella, der Erbtöchter des Herzogs Karl II. von Lothringen, von dessen Koch erfunden worden, um ihr besorgtes Gesicht etwas zu erhellen. Mandelmus von süßen und bitteren Mandeln und kandierte Melonen auf einer Oblate mit Zuckerguss machen diese ovale, fast rautenförmige Süßigkeit zu einer exklusiven Spezialität. Die besondere Form soll der Legende nach auf die Form der damals üblichen Schiffe zurückzuführen sein, vielleicht aber auch auf die Weberschiffchen der Fischer, die damit ihre Netze flickten.



Abbildung 1: Die originalen Calissons aus Aix (Frankreich).

Abstrahiert man die ovale Form der Calissons zu einer Raute, die aus zwei an einer Seite zusammengeklebten, gleichseitigen Dreiecken mit Seitenlänge 1 zusammengesetzt sind, so stellt sich die interessante Frage, wie man eine endliche Menge solcher Calissons in einer sechseckigen, nicht notwendigerweise regelmäßigen, Box verpacken könnte, wieviele unterschiedliche Arrangements existieren und wie man solche zufällig erzeugt (Aufgetaucht sind diese Fragen z.B. in [9]). Diese Fragestellung führt zu den sogenannten *Lozengeteilungen*, die Gegenstand dieser Arbeit sind.

1.1 Definition einer Lozengeteilung

Es bezeichne $\mathcal{H}_{a,b,c}$ eine hexagonale Region des Dreieckgitters¹ \mathcal{D} , wobei das betreffende Hexagon im Uhrzeigersinn links anfangend die Seitenlängen a, b, c, a, b, c habe. Die Menge aller solcher auf diese Weise darstellbarer Hexagons sei mit \mathcal{H} bezeichnet.

Die zentrale Frage dieser Arbeit ist die Frage nach Möglichkeiten zur Herstellung gleichverteilter Stichproben aus der Menge aller Teilungen von $\mathcal{H}_{a,b,c}$.

Die Frage ist nun, warum es nicht genügt oder möglich ist die inverse Verteilungsfunktion zu berechnen und damit uniform verteilte Stichproben zu erzeugen.

¹Es sei das Dreieckgitter die Teilung der Ebene durch gleichseitige Dreiecke.

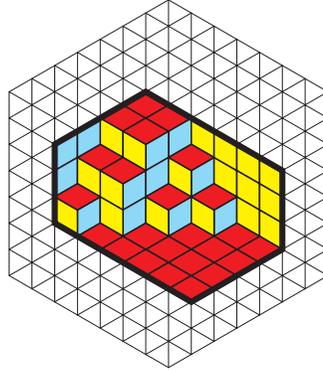


Abbildung 2: Beispielteilung im Hexagon $\mathcal{H}_{3,4,6}$

Die Größe des Zustandsraumes wird dies im Allgemeinen verhindern, denn schon für $\mathcal{H}_{10,10,10}$ ist die Anzahl Teilungen sehr groß, es gibt etwa $9.3 \cdot 10^{33}$ Teilungen.

Könnte man davon je eine Teilung auf einem einzelnen Proton speichern, verbrauchte man ein Drittel *aller* Protonen im Universum.

Da die meisten Computer wesentlich weniger Speicherplätze besitzen und in der Physik und den Anwendungen grössere Teilungen von Interesse sind, ist dies kein gangbarer Weg. Es zeigt sich aber, daß man mit Hilfe von Kopplungen von Markovketten und der Definition einer partiellen Ordnung auf der Menge aller Teilungen mit Hilfe der Kopplung aus der Vergangenheit mit geringem Aufwand das gewünschte erreichen kann. Die Methodik wird in den Abschnitten 2-4 präzisiert.

Im Folgenden bezeichne \mathcal{R} eine Region des Dreieckgitters \mathcal{D} . Eine solche Region kann auch als Menge von gleichseitigen Dreiecken gesehen werden, deren Randpolygon zunächst nicht zwingend konvex oder zusammenhängend sein muss. Ausserdem soll im weiteren unter einem Lozenge eine Raute verstanden werden, die genau zwei Dreiecke des Dreieckgitters überdeckt. Man spricht von einer Teilung T der Region \mathcal{R} , falls alle Dreiecke der Region überlappungsfrei und lückenlos durch Lozenges überdeckt werden. Eine solche Teilung heißt auch *Lozengeteilung*, die Region dann entsprechend lozengeteilbar oder einfach nur teilbar.

Es bezeichne $\mathcal{L}_{\mathcal{R}}$ die Menge aller Lozengeteilungen der Region \mathcal{R} . Kapitel 2 wird eine naive Charakterisierung dieser Regionen liefern, im Sinne der Überdeckbarkeit durch Lozenges, denn es ist nicht sofort klar, welche Regionen überhaupt lozengeteilbar sind.

Abbildung (2) zeigt exemplarisch eine Region mit einer Lozengeteilung. Die Farben indizieren dabei die drei möglichen Ausrichtungen der Raute.

1.2 Verbindung zu Modellen der statistischen Physik

Die Lozengeteilungen sind zu einer Reihe interessanter Probleme aus der Physik und Polymerchemie äquivalent, die folgenden Abschnitte orientieren sich stark an [26], die Abbildungen dieses wurden den Abbildungen aus den Referenzen nachempfunden.

Etwas allgemeiner kann man zunächst endliche, planare Graphen betrachten. In der Chemie kann ein solcher Graph unter bestimmten Regularitätsbedingungen als Grenzfläche eines Kristalls angesehen werden. Wichtige Beispiele solcher Graphen sind etwa der Dreieckgraph, das kubische Gitter und der Bienenwabengraph, wie sie in Abb. 3 angedeutet werden.

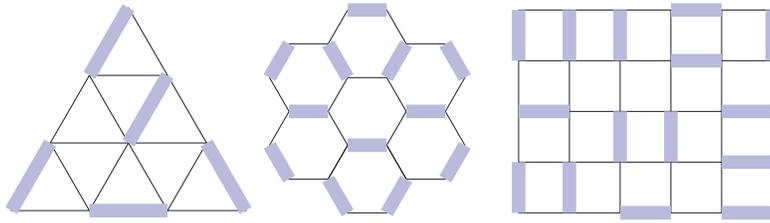


Abbildung 3: Dreieckgitter, Bienenwabengraph und kubischer Graph, jeweils mit perfect matching (die Dimere werden durch die blauen Balken kodiert).

Bringt man nun eine solche Grenzfläche in eine Umgebung mit Dimeren² ein, so werden sich die einzelnen Dimere an je zwei Knoten der Kristallstruktur anheften. Man nennt diesen Vorgang *Adsorption*.

Es ist ein klassisches Thema der Kombinatorik nach der Anzahl möglicher dichter Packungen von Dimeren zu fragen, d.h. Packungen ohne einzelne freie Knoten. Diese Frage wurde schon in den 1960er Jahren des letzten Jahrhunderts für planare Graphen unabhängig von Kasteleyn (siehe etwa [25] oder [24]) und Temperley/Fisher (siehe [19] und [20]) bewiesen. In der Literatur wird bei dichten Packungen oft von *perfect matchings* gesprochen.

Fixiert man nun eine endliche Teilregion eines solchen Gitters, kann man zu den selben Fragen nach Teilbarkeit und Anzahl der Teilungen kommen, sowie dem Einfluß der Randbedingungen. Die Verbindung zu den zufälligen Teilungen wird klar, wenn man den Dimeren Polygone zuordnet, die entsprechend der Gitterstruktur aneinandergelegt werden können, siehe dazu Abbildung 4.

Es gibt verschiedene Typen solcher Teilungen. Die bekanntesten sind sicherlich die Dominoteilungen (siehe z.B. [27]) und die Rhombusteilungen.

Eine mögliche Interpretation solcher Teilungen ist, die Möglichkeit die Struktur von Quasikristallen zu beschreiben.

Diese Quasikristalle weisen einige merkwürdige Eigenschaften auf, wie etwa die spontane lokale Reorganisation der Molekularstruktur, und können anhand der

²Ein Dimer ist eine Verbindung aus zwei gleichartigen Molekülen oder Atomen.

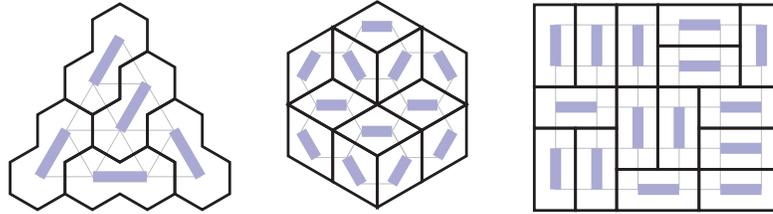


Abbildung 4: Die Abbildung zeigt die den perfect matchings der vorigen Abbildung zugeordneten Teilungen.

Generierung von Teilungen am Computer simuliert und analysiert werden. Für eine Einführung siehe [3],[43] und [44].

Ein anderes interessantes Modell sind die perfect matchings auf dem sogenannten Fisher-Graphen, es ist äquivalent zum 2-Spin-Ising-Modell auf einer fixierten Region von \mathbb{Z}^2 . Die zugehörige Bijektion wird in [26] auf Seite 3 genauer beschrieben. Ein Beispiel für die Interpretation des Isingmodells als zufällige Teilung wird in Abbildung 5 gegeben.

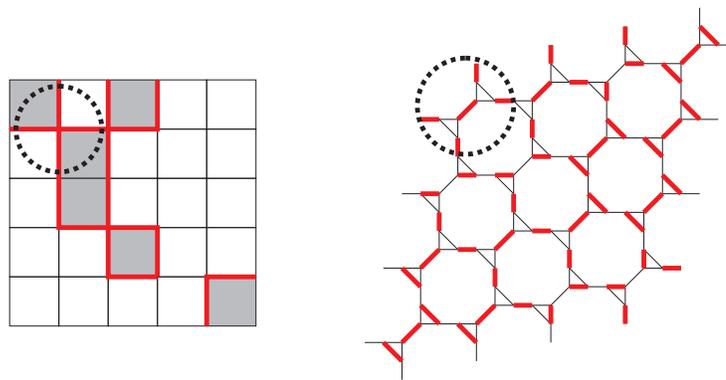


Abbildung 5: Interpretation des Isingmodells (links) als perfect matching auf dem Fisher-Graphen (rechts). Rechts sind die Dimere rot markiert, die Spins des Isingmodells werden durch die beiden Farben grau und weiß dargestellt, die gestrichelten Kreise geben ein Beispiel für die Interpretation von Kanten zwischen gegensätzlich geladenen Spins und Teilen des Fisher-Graphen mit perfect matching.

Interessant ist es dann, gleichverteilte Stichproben³ solcher Teilungen zu generieren, um gewisse typische Eigenschaften wie etwa die Entropie oder die Struktur des Stichprobenraums statistisch auswerten zu können.

³Das Auftauchen jeder Stichproben sollte gleiche Wahrscheinlichkeit haben.

Die wesentlichen Schwierigkeiten liegen dabei zumeist bei der oft enormen Größe des Stichprobenraums.

1.3 Gliederung der Arbeit

Abschnitt 1 führte bereits nach der Definition der Lozengeteilungen einige physikalische Motivationen ein.

Abschnitt 2 befasst sich mit elementaren Eigenschaften der Lozengeteilungen. Darunter zählen zunächst die Frage nach der Überdeckbarkeit beliebig geformter Regionen und die Darstellung von Lozengeteilungen durch Pfade im euklidischen Gitter. Diese Darstellung erlaubt es, mit Hilfe des berühmten Lemmas von Gessel-Viennot, die Anzahl möglicher Teilungen einer hexagonalen Region auf die Berechnung einer Determinante zu reduzieren. Abschliessend wird das Wachstum der Menge aller möglichen Teilungen exemplarisch für regelmäßige Hexagons untersucht. Dies liefert eine Rechtfertigung für das Erzeugen zufälliger Teilungen mittels der noch vorzustellenden zufälligen Dynamik.

Abschnitt 3 führt nach der Klärung einiger Notation und der Niederschrift wichtiger, noch zu verwendender Resultate, zur Kopplung aus der Vergangenheit (CFTP).⁴ Mit Hilfe von zwei gekoppelten, ansonsten aber identischen, ergodischen Markovketten lassen sich mit diesem Algorithmus exakt nach der Gleichverteilung der Originaldynamik verteilte Stichproben erzeugen. Kann der Zustandsraum der Originalkette zusätzlich noch mit einer partiellen Ordnung versehen werden, so lässt sich die Erzeugung der Stichproben wesentlich vereinfachen.

Abschnitt 4 wendet die bis dahin dargestellten Ergebnisse auf die Lozengeteilungen an. Nach Einführung der *plane partitions*, die sich als hilfreiche zweite Repräsentation der Lozengeteilungen erweisen werden, wird die Markovkette \mathcal{M}_{loz} definiert. Danach wird die Menge der *plane partitions* mit einer natürlichen partiellen Ordnung versehen werden, die zusammen mit der Ergodizität von \mathcal{M}_{loz} die Anwendung des verbesserten CFTP-Algorithmus erlaubt. Abschliessend wird noch mit Hilfe eines Submartingal-Argumentes und einer Abschätzung der Mischungszeit durch die Koaleszenzzeit gezeigt, daß \mathcal{M}_{loz} schnell mischend ist. Als letzter Teil folgen Untersuchungen zur Verteilung der in den vorigen Argumenten verwendeten Koaleszenzzeit.

Abschnitt 5 widmet sich der Vorstellung der Implementation der vorgestellten Dynamik und geht kurz auf die wesentlichen Aspekte der Programmierung und der verwendeten Software ein. Ein kurzes Programmlisting schliesst sich an.

Abschnitt 6 gibt darüber hinaus Einblicke über mögliche Verallgemeinerungen auf Rhombus- und Rhomboederteilungen und die Schwierigkeiten bei der Erzeugung selbiger. Darüberhinaus werden kurz Aspekte der aktuellen Forschung vorgestellt, wie etwa das „arctic circle“ Phänomen und artverwandte Probleme.

⁴CFTP=Coupling From The Past

1.4 Eigenanteil

Während der Bearbeitung des Themas sind an einigen Stellen Fragen aufgetaucht, die zum Teil beantwortet werden konnten. So wurde etwa im Abschnitt 2 Lemma 2.1.1 gezeigt, daß die einzigen überdeckbaren konvexen Regionen gerade Hexagons sind, deren parallele Seiten gleich lang sind. Dies lässt zu, für beliebig geformte Regionen ein Teilbarkeitskriterium im Sinne der völligen Überdeckung der Region durch Lozenges zu formulieren. Abschnitt 2.4 beinhaltet dann einige Betrachtungen zum Wachstum der Anzahl der Lozengeteilungen. Mit Hilfe der Formel von MacMahon konnte das exponentielle Wachstum der Anzahl Teilungen gezeigt werden, insbesondere Satz 2.4.1 ergibt eine obere und untere Schranke. Abschnitt 4 weicht dann in den Betrachtungen von der üblichen Verwendung der deBruijn-Pfade ab und führt die so genannten plane partitions ein. Diese Sichtweise erlaubt die Eigenschaften der noch vorzustellenden Dynamik \mathcal{M}_{loz} aus [33] genauer zu untersuchen. Im Rahmen dessen konnte auch eine minimale Argumentationslücke in Lemma 11 der Referenz geschlossen werden. Abschnitt 5 beinhaltet dann die Beschreibung der Implementierung von \mathcal{M}_{loz} , die direkt auf der Idee der Kodierung der Lozengeteilungen durch plane partitions basiert. Jene Implementierung ist auch die Grundlage für die in Abschnitt 4.5 durchgeführte Untersuchung der Verteilung der Koaleszenzzeiten. Es stellt sich heraus, daß die gewonnenen Daten für eine Gumbelverteilung der Koaleszenzzeiten sprechen.

Außerdem wurden an vielen Stellen, insbesondere in Lemma 2.3.3, die Berechnungen ausführlicher als in den Quellen dargestellt.

Sämtliche Abbildungen (bis auf Abbildung 1) sind mit CorelDraw oder dem entwickelten Programm vom Autoren selbst erstellt worden. Viele der Abbildungen sind in ähnlicher Form in den entsprechenden Referenzen zu finden.

2 Eigenschaften der Lozengeteilungen

Im Folgenden werden einige Eigenschaften der Lozengeteilungen näher betrachtet, die insbesondere für die Simulation von gleichverteilten Stichproben wichtig sind. Zunächst werden Regionen des Dreieckgitters auf Überdeckbarkeit ohne Lücken oder Überschneidungen durch Lozenges geprüft.

Danach wird mittels der noch zu definierenden deBruijn-Pfade eine Repräsentation der Lozengeteilungen als Pfade im gerichteten euklidischen Gitter eingeführt, um mit Hilfe des von Ira Gessel und Gérard Viennot in [22], und in ähnlicher Form von Bernt Lindström in [31] für Matroide, gezeigten Lemmas, die Anzahl der möglichen Teilungen zu berechnen.

Diese Anzahl lässt sich für beliebige hexagonale Regionen mit Seitenlängen a, b, c, a, b, c durch die schöne Formel

$$\prod_{i=1}^a \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2}$$

berechnen.

Die Formel geht auf die Arbeiten von MacMahon (siehe [35], Abschnitt 492 und [34]) zurück, in denen er sich mit *plane partitions* (siehe Kapitel 4) beschäftigt, die wiederum bijektiv auf die Lozengeteilungen abgebildet werden können.

MacMahon selbst gibt die Generatorfunktion⁵ dieser Anzahl und P. Stanley findet in [40] schliesslich die obige Form.

2.1 vollständige Überdeckbarkeit eines Gebietes

Scheinbar trivial ist die Frage nach der vollständigen Überdeckbarkeit einer nicht notwendig konvexen, einfach zusammenhängenden Teilregion $\mathcal{R} \subset \mathcal{D}$ durch Lozenges. Einige starke (algebraisch motivierte) Teilbarkeitskriterien wurden z.B. von Conway in [41] gegeben, Thurston gibt in [42] einen Algorithmus mit dem entschieden werden kann, ob eine zusammenhängende Region des Dreieckgitters mit Lozenges auslegbar ist.

Die von Thurston diskutierte Idee von Conway und Lagarias eine Gruppe zur Beschreibung von Regionen zu benutzen, basiert auf der Idee jedem Gebiet im Dreieckgitter ein Wort über dem Alphabet $\{a, b, c, a^{-1}, b^{-1}, c^{-1}\}$ zuzuordnen. Dabei ordnet man zunächst den drei Kanten verschiedener Orientierung Buchstaben zu, siehe Abb. 6.

Jedes Wort über diesem Alphabet ist dann als „Wegbeschreibung“ zu verstehen, ähnlich der Beschreibung für eine Turtlegraphik; so ist etwa das Wort $bc b^{-1} c^{-1}$ als Lozenge zu verstehen, analog $aca^{-1}c^{-1}$ und $aba^{-1}b^{-1}$.

Mit Hilfe dieser Kodierung kann nun die Gruppe

$$L = \langle a, b, c | bcb^{-1}c^{-1} = aba^{-1}b^{-1} = cac^{-1}a^{-1} = \varepsilon \rangle$$

⁵Eine Generatorfunktion ist in der Kombinatorik eine formale Potenzreihe $\sum_{n \geq 0} a_n t^n$, deren Koeffizienten a_n im allgemeinen die Lösung eines Abzählproblems in Abhängigkeit des Parameters n darstellen.

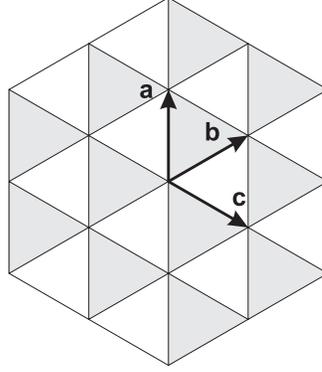


Abbildung 6: Man kann auf dem Dreieckgitter eine Art Koordinatensystem durch die drei Richtungen a , b und c definieren und damit beliebige Randpolygone beschreiben.

definiert werden, wobei ε das triviale, d.h. das leere, Wort sei. Es ist das neutrale Element der Gruppe. Man bemerkt nun, daß die Teilbarkeit einer Region impliziert, daß das randbeschreibende Wort trivial ist, die Umkehrung gilt jedoch nicht. In Abbildung 8 wird eine Region dargestellt, deren Randpolygon durch das Wort

$$accb^{-1}a^{-1}c^{-1}c^{-1}b = \varepsilon$$

beschrieben wird, aber keine Teilung besitzt.

Dies gibt Anlass für eine Reihe von Beobachtungen, die z.B. in [12] diskutiert werden.

Im Folgenden wird eine Herangehensweise gewählt, die die Schwierigkeiten bei der Überdeckbarkeit deutlich machen soll und nicht nur eine einfache Charakterisierung der konvexen Gebiete in \mathcal{D} erlaubt, sondern auch ein einfaches hinreichendes Kriterium für die Teilbarkeit einer Region im einfachen Fall der Lozengeteilungen liefert. Bei der Argumentation kann auf algebraische Argumente völlig verzichtet werden.

Zunächst sei $\mathcal{R} \subset \mathcal{D}$ ein einfach zusammenhängendes Gebiet. Man färbe das Dreieckgitter \mathcal{D} derart schwarz und weiß, daß je zwei Dreiecke, die eine Kante gemeinsam haben, unterschiedlich gefärbt sind. Das ergibt eine Art Schachbrettmuster, wie es in Abbildung 7 beispielhaft für ein nichtkonvexes Gebiet dargestellt ist.

Es ist nützlich folgende Eigenschaften zu definieren:

Definition 2.1.1 (Färbungseigenschaft) *Sei $\mathcal{R} \subset \mathcal{D}$ und \mathcal{D} wie oben gefärbt. Dann hat \mathcal{R} die Färbungseigenschaft, wenn die Anzahl weißer und schwarzer Dreiecke in \mathcal{R} gleich groß ist. Es bezeichne $W_{\mathcal{R}}$ bzw. $S_{\mathcal{R}}$ die Anzahl weißer bzw. schwarzer Dreiecke, die in \mathcal{R} enthalten sind.*

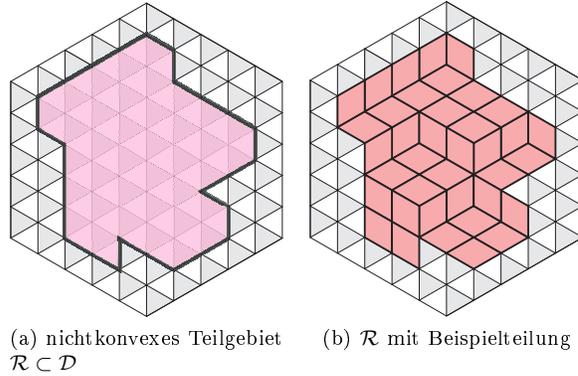


Abbildung 7: Beispiel eines nichtkonvexen Gebietes mit exemplarischer Teilung

Definition 2.1.2 (Teilbarkeit) Sei $\mathcal{R} \subset \mathcal{D}$. Dann heißt \mathcal{R} teilbar, wenn \mathcal{R} vollständig ohne Lücken oder Überlappungen mit Lozenges⁶ ausgelegt werden kann. Eine solche Auslegung sei im Folgenden Teilung genannt. Die Menge aller Teilungen einer Region \mathcal{R} sei mit $\mathcal{L}_{\mathcal{R}}$ bezeichnet.

Es ist nun offensichtlich, daß jede teilbare Region \mathcal{R} auch die Färbungseigenschaft hat, denn ist \mathcal{R} teilbar, so überdeckt jedes Lozenge je ein weißes und ein schwarzes Dreieck, es gilt also $\mathcal{W}_{\mathcal{R}} = \mathcal{S}_{\mathcal{R}}$.

Die Umkehrung gilt jedoch nicht, Abb. 8 liefert als Gegenbeispiel ein nicht teilbares Gebiet mit Färbungseigenschaft.

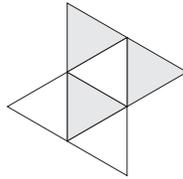


Abbildung 8: nichtteilbare Region mit Färbungseigenschaft

Nach eingehender Betrachtung möglicher Regionen ergab sich das folgende

Lemma 2.1.1 Sei $\mathcal{R} \subset \mathcal{D}$ konvex mit Färbungseigenschaft. Dann ist $\mathcal{R} \in \mathcal{H}$.⁷ Insbesondere ist dann \mathcal{R} teilbar.

Beweis. Zunächst ist zu zeigen, dass jede Region $\mathcal{H}_{a,b,c} \in \mathcal{H}$ teilbar ist. Dazu betrachte Abb. 9. Es ist klar, daß man jede Region $\mathcal{H}_{a,b,c}$ auf diese Art als isometrische Projektion einer Box der Höhe a , der Breite b und Tiefe c interpretieren kann.

⁶Ein Lozenge ist nach Kapitel 1, ein aus zwei Dreiecken mit gemeinsamer Kante zusammengesetzter Rhombus.

⁷Es sei daran erinnert, daß mit \mathcal{H} die Menge aller $\mathcal{R} \subset \mathcal{D}$ bezeichnet war, die Hexagons mit Seitenlängen a,b,c,a,b,c sind, siehe auch Abschnitt 1.1.

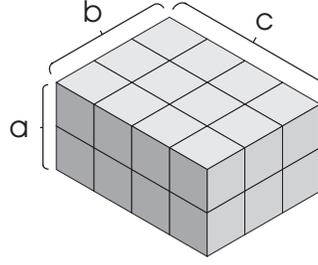


Abbildung 9: $\mathcal{H}_{2,3,4}$. Die spezielle Auslegung mit Lozenges suggeriert eine Box der Größe $a \times b \times c$

Insbesondere für $a = 0$ degeneriert die hexagonale Region $\mathcal{H}_{a,b,c}$ zur isometrischen Projektion einer Box der Höhe null. In dieser ist offenbar die Boxinterpretation auch die einzig mögliche Teilung (analog auch für $b = 0$ bzw. $c = 0$). Weiterhin bemerkt man, dass es nur fünf Typen konvexer Regionen in \mathcal{D} geben kann. Das sind gleichseitige Dreiecke, Parallelelogramme bzw. Trapeze, unregelmäßige Fünfecke und Hexagons, spezieller die Regionen aus \mathcal{H} , sowie unregelmäßige hexagonale Regionen mit Seitenlängen a, b, c, d, e, f und

$$a \neq d, b \neq e, c \neq f.$$

Nun ist zunächst offenbar keine dreieckige Region $\Delta \subset \mathcal{D}$ teilbar, denn mittels vollständiger Induktion lassen sich leicht die Anzahlen der weissen und schwarzen Dreiecke in Δ angeben, wenn man voraussetzt, dass die Dreiecke in den Ecken der Region schwarz gefärbt sind und $n \in \mathbb{N}, n \geq 1$ die Länge einer Dreiecksseite beschreibt.⁸ Dann gilt:

$$\mathcal{S}_\Delta = \frac{n(n+1)}{2} \neq \frac{n(n-1)}{2} = \mathcal{W}_\Delta$$

d.h. Δ ist nicht färbbar, also auch nicht teilbar.

Die parallelogrammartigen Regionen dagegen sind offensichtlich aus \mathcal{H} und damit nach obiger Bemerkung teilbar. Dahingegen sind die trapezförmigen Regionen in ein Dreieck und ein Parallelogramm zerlegbar und damit nicht färbbar, also erst recht nicht teilbar.

Genauso kann man jede fünfeckige Region mit einem (nicht-teilbaren) Dreieck zu einem (teilbaren) Parallelogramm ergänzen. Daraus ergibt sich unmittelbar, daß fünfeckige Regionen nicht färbbar sind, also nicht teilbar.

Es bleiben die unregelmäßigen sechseckigen Regionen, sei eine solche durch $\mathcal{H}_{a,b,c,d,e,f}$ gekennzeichnet, siehe auch Abb. 10a. Da $\mathcal{H}_{a,b,c,d,e,f}$ unregelmäßig, gilt

$$a \neq d, b \neq e, c \neq f,$$

⁸Eine duale Aussage gilt, wenn man die Farben vertauscht.

da sonst die Region zu einer Region aus \mathcal{H} degeneriert. Nun kann diese Region durch zwei Dreiecke Δ_a und Δ_d der Seitenlängen a und d zu einem teilbaren Parallelogramm ergänzt werden (siehe Abb. 10b).

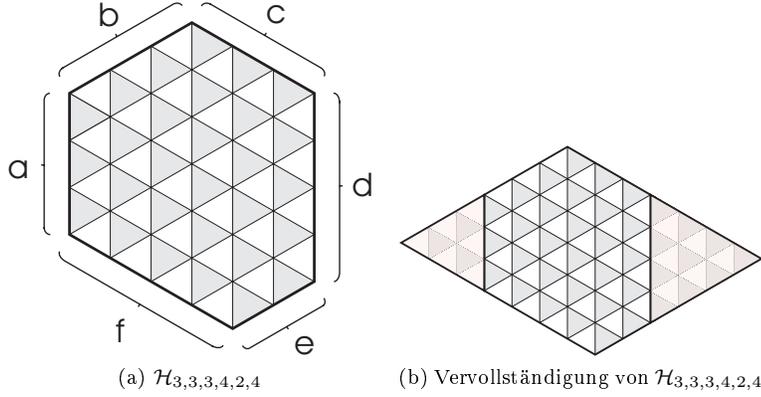


Abbildung 10: irreguläres Hexagon mit Vervollständigung zum Parallelogramm

Es bleibt zu zeigen, dass die Anzahl der weißen resp. schwarzen Dreiecke in Δ_a und Δ_d zusammen ungleich ist. Man bemerkt zuerst, daß die zu betrachtenden Dreiecke dual gefärbt sind, die Ecken beider Dreiecke also unterschiedlich gefärbt sein müssen, siehe Abb. 10b. Daraus ergeben sich die Anzahlen gleichgefärbter Dreiecke:

$$\mathcal{W}_{\Delta_a \cup \Delta_d} = \frac{a(a+1)}{2} + \frac{d(d-1)}{2}$$

$$\mathcal{S}_{\Delta_a \cup \Delta_d} = \frac{a(a-1)}{2} + \frac{d(d+1)}{2}$$

Es sei angenommen es gelte $\mathcal{W}_{\Delta_a \cup \Delta_d} = \mathcal{S}_{\Delta_a \cup \Delta_d}$, dann ergäbe sich:

$$\begin{aligned} \mathcal{W}_{\Delta_a \cup \Delta_d} = \mathcal{S}_{\Delta_a \cup \Delta_d} &\Leftrightarrow a(a+1) + d(d-1) = a(a-1) + d(d+1) \\ &\Leftrightarrow a(a+1-a+1) = d(d+1-d+1) \\ &\Leftrightarrow a = d \end{aligned}$$

im Widerspruch zu $a \neq d$. Eine analoge Schlussweise gilt auch für die anderen beiden möglichen Ergänzungen durch Dreiecke. Damit ist auch $\mathcal{H}_{a,b,c,d,e,f}$ nicht färbbar und somit ebenfalls nicht teilbar. \square

Das Lemma charakterisiert also die Regionen $\mathcal{R} \in \mathcal{H}$ als einzige konvexe Regionen, die überhaupt teilbar sind. Aus dieser Charakterisierung folgt unmittelbar

Lemma 2.1.2 *Sei $\mathcal{R} \in \mathcal{D}$ endlich. Dann ist \mathcal{R} teilbar, wenn \mathcal{R} als Vereinigung von endlich vielen Regionen $\mathcal{R}_i \in \mathcal{H}, i \in I \subset \mathbb{N}$ darstellbar ist, d.h.*

$$\mathcal{R} \text{ teilbar} \Leftrightarrow \exists \mathcal{R}_i \in \mathcal{H}, i \in I : \bigsqcup_{i \in I} \mathcal{R}_i = \mathcal{R}.^9$$

⁹Die Zerlegung ist offenbar nicht eindeutig bestimmt.

Insbesondere kann jede Teilung $T \in \mathcal{L}_{\mathcal{R}}$ als eine solche disjunkte Vereinigung aufgefasst werden, denn jedes Lozenge ist als $\mathcal{H}_{0,1,1}$, $\mathcal{H}_{1,0,1}$ oder $\mathcal{H}_{1,1,0}$ darstellbar, daher existiert zu jeder teilbaren Region mindestens eine solche Zerlegung. T ist in diesem Sinne die *feinste* Zerlegung, denn ein Lozenge ist die kleinste teilbare Region des Dreieckgitters \mathcal{D} . Insgesamt hat man mit diesem Lemma nun für beliebige und damit auch nichtkonvexe resp. nicht zusammenhängende Regionen ein Teilbarkeitskriterium an der Hand.

Der weiter oben erwähnte Algorithmus zur Entscheidung der Teilbarkeit einer Region von Thurston basiert auf einer solchen „Dekomposition“, in jedem Schritt des Algorithmus werden Lozenges entfernt. Lässt sich dieses Verfahren fortführen, bis nichts mehr übrig bleibt, so war die Ausgangsregion teilbar. Bleibt ein nicht teilbarer Rest übrig, so war entsprechend die Ausgangsregion nicht teilbar. Abbildung 11 illustriert eine mögliche Zerlegung einer Region mit Löchern in teilbare Subregionen.

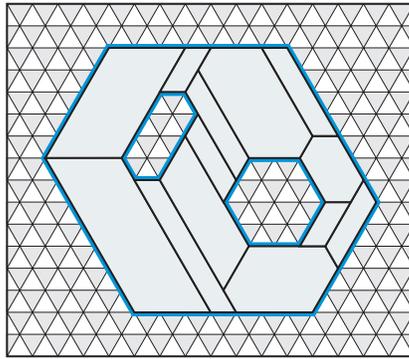


Abbildung 11: Mögliche Zerlegung der dunkelblau umrandeten Region in teilbare Subregionen.

2.2 Darstellung von Lozengeteilungen durch deBruijn-Pfade

Für die Berechnung der Anzahl der möglichen Teilungen einer hexagonalen Region $\mathcal{R} \in \mathcal{H}$ wird sich das Lemma von Gessel-Viennot als hilfreich erweisen. Es stellt auf überraschende Weise einen Zusammenhang zwischen der Anzahl eckendisjunkter Pfade (sogenannte Pfadsysteme) in einem Graphen mit bestimmten Eigenschaften und der Berechnung einer Determinante her. Ziel wird es also zunächst sein, die Lozenge-Teilungen mit einem Pfadsystem zu identifizieren. Dies leisten die von Nicolaas Govert deBruijn in [10] eingeführten und nach ihm benannten deBruijn-Pfade. Es ist möglich diese auch für allgemeinere und speziell auch mehrdimensionale Teilungen zu spezifizieren (siehe z.B. [14]). Für den Fall der Lozengeteilungen betrachte folgende

Definition 2.2.1 (deBruijn-Pfad, deBruijn-Familie) Sei $\mathcal{H}_{a,b,c} \in \mathcal{H}$ und

T eine Teilung aus $\mathcal{L}_{\mathcal{H}_{a,b,c}}$. Weiter sei an das Koordinatensystem aus Abb. 6 erinnert.

Für ein fixiertes Lozenge der Teilung T und eine fixierte Richtung \vec{r} wird die Menge aller zu diesem Lozenge adjazenten Lozenges, mit Kanten parallel zu \vec{r} , deBruijn-Pfad genannt.

Die so entstehende Sequenz von Lozenges ist eindeutig bestimmt und verbindet die zu \vec{r} parallelen Seiten von $\mathcal{H}_{a,b,c}$ miteinander.

Die Menge aller deBruijn-Pfade einer Teilung T zu \vec{r} von $\mathcal{H}_{a,b,c}$ heißt deBruijn-Familie (siehe Abbildung 12), die Menge aller Familien zu \vec{r} Faser.

Bemerkung. Es gibt so viele Familien wie Teilungen, aber zu jeder Richtung \vec{r} nur eine Faser.

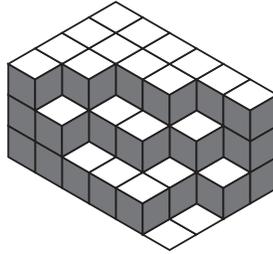


Abbildung 12: deBruijn-Familie in $\mathcal{H}_{3,4,6}$ für die a-Richtung

Eine deBruijn-Faser¹⁰ lässt sich speziell für die Lozengeteilungen einfach mit gerichteten Pfaden im gerichteten euklidischen Gitter identifizieren (die Gitterpunkte seien der Einfachheit halber Elemente von \mathbb{Z}^2).¹¹ Abbildung 13 zeigt diese Bijektion für $\mathcal{H}_{3,4,6}$.

Dabei bedeute *gerichtet*, dass für zwei Punkte (=Ecken des euklidischen Graphen) $x = (x_1, x_2)$ und $y = (y_1, y_2)$ mit $x, y \in \mathbb{Z}^2$ eine Kante existiert, wenn der euklidische Abstand gleich 1 ist.

Eine Kante von x nach y sei in „Nordrichtung“ orientiert, falls $x_1 = y_1$ und $y_2 - x_2 = 1$. Analog sei eine Kante von x nach y in „Ostrichtung“ orientiert, falls $y_1 - x_1 = 1$ und $x_2 = y_2$ (siehe auch Abb. 13 rechts).

Für die konvexen Regionen $\mathcal{R} \in \mathcal{H}_{a,b,c}$ fixiere im gerichteten, euklidischen Gitter die a Knoten A_1, \dots, A_a und B_1, \dots, B_a , wie in Abbildung 13 rechts.

Nach Konstruktion sind die so im gerichteten euklidischen Gitter definierten Pfade von A_i nach B_j eckendisjunkt, d.h. kein Pfad hat eine Ecke mit einem anderen Pfad gemein. Denn angenommen es gäbe zwei deBruijn-Pfade mit gemeinsamer Ecke, so gäbe es in der zugehörigen Lozengeteilung T zwei Lozenges, die sich überlagern. Damit wäre aber T keine zulässige Teilung mehr.

¹⁰Es gibt zu jeder hexagonalen Region genau drei Fasern.

¹¹Es genügt bereits eine der drei möglichen Fasern, um die Menge der Lozengeteilungen $\mathcal{L}_{\mathcal{H}_{a,b,c}}$ vollständig zu charakterisieren, im folgenden sei daher immer die Faser oder Familie in \vec{a} -Richtung gemeint.

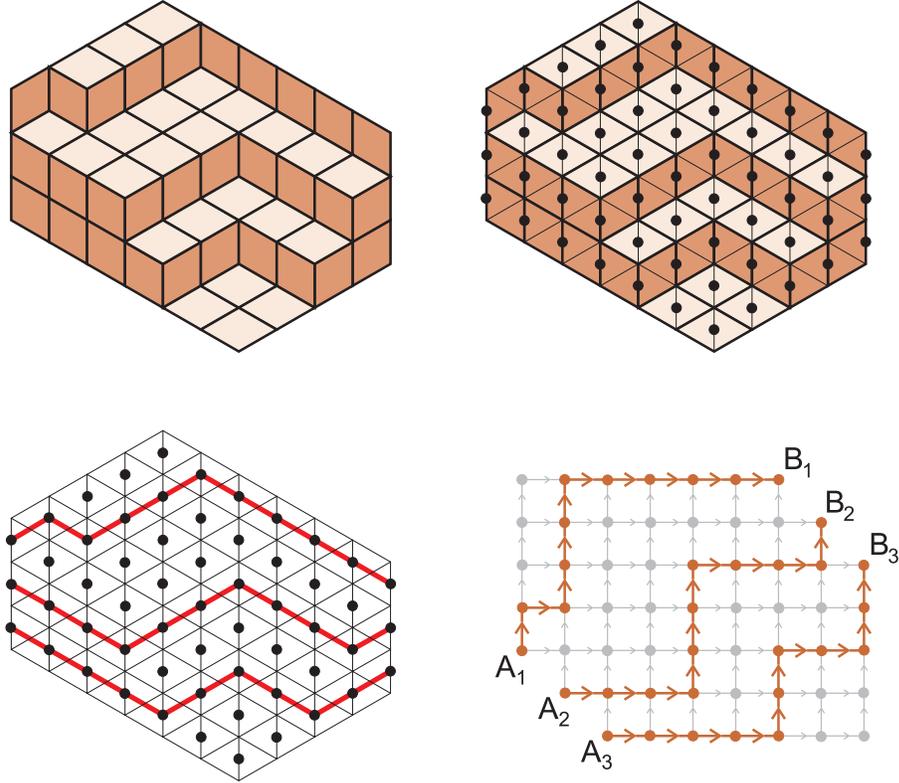


Abbildung 13: Übersetzung von $T \in \mathcal{L}_{\mathcal{H}_{3,4,6}}$ mittels deBruijn-Familie in eine Darstellung im gerichteten euklidischen Gitter.

2.3 Anzahl Teilungen konvexer hexagonaler Regionen

2.3.1 Lemma von Gessel-Viennot

Bevor das Lemma von Gessel-Viennot formuliert werden kann, ist es notwendig einige Begriffe einzuführen. Im wesentlichen wird dabei und bei der Formulierung resp. dem Beweis des Lemmas von Gessel-Viennot auf die Ausführungen in [1] ab Seite 3 von Martin Aigner zurückgegriffen, hier in der Verallgemeinerung für beliebige sechseckige Regionen $\mathcal{H}_{a,b,c}$ mit $a, b, c \in \mathbb{N}_0$.¹² Im folgenden wird ein endlicher, gerichteter, reell-gewichteter und azyklischer Graph kurz *egga-Graph* -Graph genannt. Sei nun $\underline{G} = (E, K)$ ein *egga-Graph* -Graph und $\mathbb{A} = \{A_1, \dots, A_n\}$ bzw. $\mathbb{B} = \{B_1, \dots, B_n\}$ Teilmengen der Eckenmenge E des Graphen, die nicht notwendigerweise disjunkt sein müssen.

Offenbar gibt es in \underline{G} nur endlich viele oder gar keine gerichteten Pfade zwischen einem $A_i \in \mathbb{A}$ und einem $B_j \in \mathbb{B}$, wobei auch der triviale Pfad¹³ zugelassen sei,

¹²In der gegebenen Referenz wird $a = b = c = n$ vorausgesetzt.

¹³Es sei der triviale Pfad der (gerichtete) Pfad $e \rightarrow e$ bzw. die Kante $(e, e) \in K$ mit $e \in E$

falls $A_i = B_j$. Die Pfade sind selbst von endlicher Länge.¹⁴
 Es bezeichne

$$P : A_i \rightarrow B_j$$

einen gerichteten Pfad von A_i nach B_j . Analog schreibe man für eine beliebige Permutation $\sigma \in S_n$

$$P_{(\sigma,i)} : A_i \rightarrow B_{\sigma(i)}.$$

Nun kann auf kanonische Weise jedem Pfad ein Gewicht zugeordnet werden, dies geschieht mittels

Definition 2.3.1 (Pfadgewicht) *Es sei $P : A_i \rightarrow B_j$ ein beliebiger Pfad von A_i nach B_j .*

Dann ist das Gewicht von P definiert durch

$$\omega(P) = \begin{cases} \prod_{k \in P} \omega(k) & \text{falls } P \text{ nicht trivial} \\ 1 & \text{falls } P \text{ trivial} \end{cases},$$

wobei $\omega(k) \in \mathbb{R}$ das Gewicht der Kante $k \in P \subset K$ sei.

Betrachte nun weiter

Definition 2.3.2 (Pfadsystem) *Es seien n Pfade $P_{(\sigma,i)} : A_i \rightarrow B_{\sigma(i)}$ gegeben, wobei $i \in \{1, \dots, n\}$.*

Dann heißt die Menge

$$\mathcal{P}_\sigma := \{P_{(\sigma,i)} : A_i \rightarrow B_{\sigma(i)}, i \in \{1, \dots, n\}\}$$

ein Pfadsystem zur Permutation σ .

Mit Hilfe der folgenden Definition kann nun auch einem Pfadsystem ein Gewicht zugeordnet werden.

Definition 2.3.3 (Pfadsystemgewicht) *Es sei*

$$\omega(\mathcal{P}_\sigma) := \prod_{i=1}^n \omega(P_{(\sigma,i)} : A_i \rightarrow B_{\sigma(i)})$$

das Gewicht des Pfadsystems \mathcal{P}_σ .

Bemerkung. Zu einer beliebigen Permutation $\sigma \in S_n$ ist das Pfadsystem zu σ nicht eindeutig bestimmt. Es kann durchaus mehrere mögliche Pfadsysteme geben, insbesondere dann, wenn es mehr als einen gerichteten Pfad zwischen mindestens einem A_i und $B_{\sigma(i)}$ gibt.

Mithin ist auch die Existenz eines Pfadsystems zu einer beliebigen Permutation nicht gesichert, insbesondere dann, wenn es ein i gibt, so daß kein Pfad

$$P : A_i \rightarrow B_{\sigma(i)}$$

mit Länge 0 und Gewicht 1

¹⁴Die Pfadlänge ist gleich der Anzahl nichttrivialer Kanten.

existiert.

Man kann diese Lücke schließen, indem man eine neue gerichtete Kante $(A_i, B_{\sigma(i)}) \in K$ mit Gewicht 0 einführt und so einen Pfad $P : A_i \rightarrow B_{\sigma(i)}$ erhält. Das Pfadgewicht und damit auch das Pfadsystemgewicht ist dann Null.

Definition 2.3.4 (Signum eines Pfadsystems) *Es sei $\sigma \in S_n$, $n \in \mathbb{N}$, eine beliebige Permutation und \mathcal{P}_σ ein Pfadsystem zu σ . Dann sei*

$$\text{sign}(\mathcal{P}_\sigma) := \text{sign}(\sigma)$$

Schließlich benötigt man zur Formulierung des Lemmas noch

Definition 2.3.5 (Pfadmatrix) *Man definiere die quadratische $n \times n$ Matrix $M := (m_{ij})_{ij}$ mit $i, j \in \{1, \dots, n\}$ durch*

$$m_{ij} := \sum_{P: A_i \rightarrow B_j} \omega(P)$$

und setze

$$m_{ij} := 0,$$

falls kein gerichteter Pfad $P : A_i \rightarrow B_j$ existiert.

Weiterhin sei ein Pfadsystem *eckendisjunkt*, falls je zwei Pfade des Pfadsystem keine gemeinsamen Ecken besitzen.

Nun kann das Lemma von Gessel-Viennot (siehe [22]) folgendermaßen formuliert werden:

Lemma 2.3.1 (Gessel-Viennot) *Seien $\underline{G} = (E, K)$ ein egga-Graph und $\mathbb{A}, \mathbb{B} \subset E$ n -elementige Teilmengen, $n \in \mathbb{N}$, M die zugehörige Pfadmatrix. Dann gilt*

$$\det M = \sum_{\mathcal{P}_\sigma \text{ eckendisjunkt}} \text{sign}(\mathcal{P}_\sigma) \omega(\mathcal{P}_\sigma).$$

Beweis. Zunächst ist nach Gottfried Leibniz die Determinante von M gegeben durch

$$\det M = \sum_{\sigma \in S_n} \text{sign}(\sigma) m_{1\sigma(1)} m_{2\sigma(2)} \dots m_{n\sigma(n)}$$

Nach Definition der Pfadmatrix ist aber:

$$\begin{aligned} & \text{sign}(\sigma) m_{1\sigma(1)} m_{2\sigma(2)} \dots m_{n\sigma(n)} \\ &= \text{sign}(\sigma) \left(\sum_{P_1: A_1 \rightarrow B_{\sigma(1)}} \omega(P_1) \right) \cdot \left(\sum_{P_2: A_2 \rightarrow B_{\sigma(2)}} \omega(P_2) \right) \cdot \dots \cdot \left(\sum_{P_n: A_n \rightarrow B_{\sigma(n)}} \omega(P_n) \right) \end{aligned}$$

und damit offenbar

$$\det M = \sum_{\mathcal{P}_\sigma} \text{sign}(\mathcal{P}_\sigma) \omega(\mathcal{P}_\sigma).$$

Definiert man nun

$$\mathcal{N} := \{\mathcal{P}_\sigma \mid \mathcal{P}_\sigma \text{ nicht eckendisjunkt}\},$$

so ist nur noch

$$\sum_{\mathcal{P}_\sigma \in \mathcal{N}} \text{sign}(\mathcal{P}_\sigma) \omega(\mathcal{P}_\sigma) = 0 \quad (1)$$

zu zeigen.

Sei nun $\mathcal{P}_\sigma \in \mathcal{N}$. Dann existiert ein Pfad $P_{i_0} : A_{i_0} \rightarrow B_{\sigma(i_0)}$ derart, daß i_0 der kleinste Index ist, so daß P_{i_0} eine Ecke mit einem anderen Pfad aus dem Pfadssystem gemein hat. Sei $X \in P_{i_0}$ die erste Ecke für die ein Pfad P' existiert, so daß $P_{i_0} \cap P' \neq \emptyset$ und $P_{j_0} : A_{j_0} \rightarrow B_{\sigma(j_0)}$ der Pfad mit kleinstem Index $j_0 > i_0$ mit $X \subset P_{i_0} \cap P_{j_0}$.

Dann definiere die Abbildung $\pi : \mathcal{N} \rightarrow \mathcal{N}$ durch folgende Konstruktion für ein Pfadssystem \mathcal{P}_σ :

- wähle

$$P_{i_0} = ((A_{i_0}, e_1), (e_1, e_2), \dots, (e_k, X), (X, e_{k+1}), \dots, (e_m, B_{\sigma(i_0)}))$$

und

$$P_{j_0} = ((A_{j_0}, \hat{e}_1), (\hat{e}_1, \hat{e}_2), \dots, (\hat{e}_k, X), (X, \hat{e}_{k+1}), \dots, (\hat{e}_m, B_{\sigma(j_0)}))$$

wie oben.

- $\pi(\mathcal{P}_\sigma) = \mathcal{P}'_{\sigma'}$ mit

$$P'_{i_0} = ((A_{i_0}, e_1), (e_1, e_2), \dots, (e_k, X), (X, \hat{e}_{k+1}), \dots, (\hat{e}_m, B_{\sigma(j_0)}))$$

und

$$P'_{j_0} = ((A_{j_0}, \hat{e}_1), (\hat{e}_1, \hat{e}_2), \dots, (\hat{e}_k, X), (X, e_{k+1}), \dots, (e_m, B_{\sigma(i_0)})).$$

Dabei sei

$$\sigma' = \sigma \circ (i_0 \ j_0)$$

das Produkt von σ mit der Transposition $(i_0 \ j_0)$

Mit dieser Konstruktion gilt nun offenbar

- $\forall \mathcal{P}_\sigma \in \mathcal{N} : \pi(\pi(\mathcal{P}_\sigma)) = \mathcal{P}_\sigma$
- $\text{sign}(\pi(\mathcal{P}_\sigma)) = -\text{sign}(\mathcal{P}_\sigma)$
- $\omega(\pi(\mathcal{P}_\sigma)) = \omega(\mathcal{P}_\sigma)$

Daher muss (1) verschwinden, denn zu jedem Pfadssystem aus \mathcal{N} wird mittels der Involution π ein Pfadssystem mit entgegengesetzten Signum assoziiert. Diese Vorgehensweise nennt man auch *tailswapping*, Abb. 14 konkretisiert das Vorgehen.

Damit ist die Behauptung gezeigt. \square

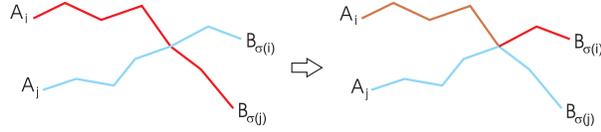


Abbildung 14: Vertauschen der Enden zweier Pfade (tailswap)

2.3.2 Formel von MacMahon

Im Abschnitt über die deBruijn-Pfade wurde bereits ein Pfadsystem definiert, das in Bijektion mit den Lozengeteilungen steht. Mittels des vorgestellten Lemmas von Gessel-Viennot genügt es nun die zugehörige Pfadmatrix zu bestimmen und deren Determinante zu berechnen, um die Anzahl der möglichen Teilungen einer Region $\mathcal{H}_{a,b,c}$ zu erhalten.

Mit der Beschränkung auf die Teilungen von $\mathcal{H}_{a,b,c}$ betrachtet man ein Pfadsystem, wie es bereits bei der Definition der deBruijn-Pfade angegeben wurde (siehe Abb. 13, letztes Teilbild).

Jeder der Pfade kann in dem Rechteck der Größe $b \times c$ variieren, siehe dazu Abb. 15. Jeder Pfad von A_i nach B_i für ein festes $i \in \{1, \dots, a\}$ hat dann die Länge $b + c$.

Sei nun jedes Kantensegment mit Ausrichtung nach Norden durch 1 kodiert, jedes Kantensegment mit Ausrichtung nach Osten durch 0. Dann ist jeder Pfad von A_i nach B_i eine Sequenz von b Nullen und c Einsen. Die Menge aller solcher Pfade hat demnach die Mächtigkeit $\binom{b+c}{b}$ und nach Definition der Pfadmatrix setzt man nun

$$m_{ii} = \binom{b+c}{b}.$$

Es bleiben noch die Anzahlen möglicher Pfade zwischen A_i und B_j mit $i \neq j$. Dazu betrachte die Abb. 15. Es wird deutlich, daß die Pfade nur noch in dem Rechteck der Größe $(b+i-j) \times (c-i+j)$ variieren können; die Pfadlänge bleibt unverändert. Nach der vorgeschlagenen Kodierung der Pfade und der Definition der Pfadmatrix ist nun analog zu vorher

$$m_{ij} = \binom{b+c}{b+i-j}$$

zu setzen.

Insgesamt ergibt sich damit die Pfadmatrix

$$M := (m_{ij})_{ij} = \left(\binom{b+c}{b+i-j} \right)_{ij}.$$

Im Abschnitt über die deBruijn-Pfade wurde schon erwähnt, daß die betrachteten deBruijn-Pfade eckendisjunkt sein müssen, das bedeutet aber nichts anderes als das nun das Lemma von Gessel-Viennot angewendet werden kann und die Anzahl Teilungen gerade die Determinante von M ist.

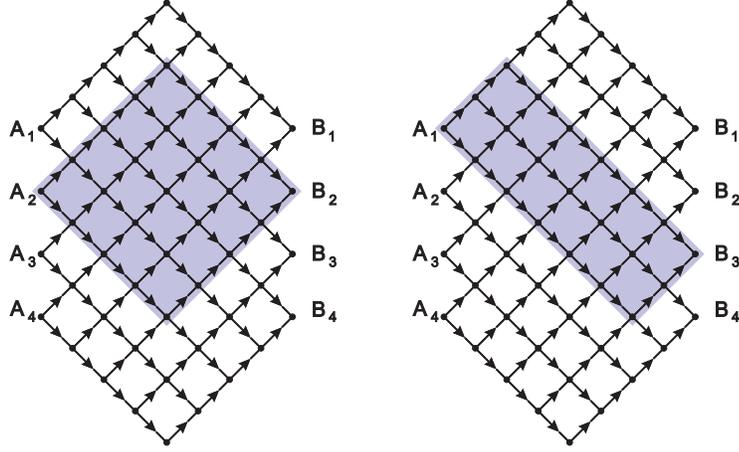


Abbildung 15: Variationsbereiche möglicher Pfade von $A_2 \rightarrow B_2$ resp. $A_1 \rightarrow B_3$ sind blau unterlegt.

Die folgenden Berechnungen stellen eine Detaillierung der Vorschläge für die Berechnung von $\det M_a(b, c)$ aus [29] ab S. 11 dar.

Bevor die Determinante berechnet werden kann, sei ein kleines Lemma erwähnt, das helfen wird einen Beweis mittels vollständiger Induktion für beliebige $a \geq 0$ zu führen. Dazu führe folgende Schreibweise ein:

Definition 2.3.6 *Es sei A eine $n \times n$ Matrix. Dann bezeichne $A_{n_1, \dots, n_m}^{n_1, \dots, n_{m'}}$ die Matrix, die man durch Streichen der Zeilen n_1, \dots, n_m und durch Streichen der Spalten $n_1, \dots, n_{m'}$ erhält.*

Auf das erwähnte Lemma wird in [29], S. 11 verwiesen. Es liefert eine hilfreiche Identität für die Berechnung von Determinanten, die unter anderem auf C. L. Dodgson¹⁵ zurückgeht, siehe dazu auch [16]. Man nennt diese Methode auch Kondensation; diese liefert ein iteratives Verfahren zur Berechnung der Determinante durch die Berechnung gewisser Minoren der Ausgangsmatrix. Dies präzisiert

Lemma 2.3.2 *Sei A eine $n \times n$ Matrix, $n \geq 3$. Dann gilt:*

$$\det A \cdot \det A_{1,n}^{1,n} = \det A_1^1 \cdot \det A_n^n - \det A_1^n \cdot \det A_n^1.$$

Die Frage nach der Anzahl von Teilungen einer beliebigen Region $\mathcal{R} \in \mathcal{H}$ lässt sich also so formulieren:

Lemma 2.3.3 *Sei $\mathcal{H}_{a,b,c} \in \mathcal{H}$, dann ist*

$$\#\mathcal{L}_{\mathcal{H}_{a,b,c}} = \det M_a(b, c) = \prod_{i=1}^a \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2} \quad (2)$$

¹⁵bekannter als Autor von „Alice in Wonderland“ unter dem Pseudonym Lewis Carroll.

für beliebiges $a \in \mathbb{N}_0$ und $b, c \in \mathbb{N}$.

Beweis. (per Induktion) Zunächst bemerkt man, daß im Fall $a = 0$ nur genau eine Teilung der nun rautenförmigen Region möglich ist. Dies wurde schon im Beweis von Lemma 2.1.1 bemerkt. Also

$$\det M_0(b, c) = 1 \quad (3)$$

Für $a = 1$ ergibt sich

$$\begin{aligned} \det M_1(b, c) &= \det_{i,j=1} \left(\binom{b+c}{b} \right) = \binom{b+c}{b} = \prod_{j=1}^b \prod_{k=1}^c \frac{j+k}{j+k-1} \\ &= \prod_{i=1}^1 \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2} \end{aligned} \quad (4)$$

Weiter gilt für beliebiges $a \geq 2$:

$$\begin{aligned} (M_a(b, c))_1^1 &= \det_{2 \leq i, j \leq a} \left(\binom{b+c}{b+i-j} \right) \\ &= \det_{2 \leq i, j \leq a} \left(\binom{b+c}{b+(i-1)-(j-1)} \right) = M_{a-1}(b, c) \end{aligned}$$

Analog gilt auch

$$\begin{aligned} (M_a(b, c))_n^n &= M_{a-1}(b, c) \\ (M_a(b, c))_n^1 &= M_{a-1}(b+1, c-1) \\ (M_a(b, c))_1^n &= M_{a-1}(b-1, c+1) \\ (M_a(b, c))_{1,n}^{1,n} &= M_{a-2}(b, c) \end{aligned}$$

Nach diesen Überlegungen gilt nun nach Lemma 2.3.3 folgende Rekursionsgleichung für die zu berechnende Determinante:

$$\begin{aligned} \det M_a(b, c) \cdot \det M_{a-2}(b, c) \\ = (\det M_{a-1}(b, c))^2 - \det M_{a-1}(b+1, c-1) \cdot \det M_{a-1}(b-1, c+1) \end{aligned}$$

Wenn diese Rekursion auch für die rechte Seite von (2) gezeigt werden kann ist der Induktionsbeweis zusammen mit (3) und (4) abgeschlossen.

Daher zeige zunächst für beliebige $i, b, c \in \mathbb{N}$:

$$\prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2} = \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}}$$

es ergibt sich

$$\begin{aligned}
& \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2} \\
&= \left(\frac{1+i}{i} \cdot \frac{2+i}{1+i} \cdots \frac{c+i}{c+i-1} \right) \cdots \left(\frac{b+i}{b+i-1} \cdot \frac{b+i+1}{b+i} \cdots \frac{b+c+i-1}{b+c+i-2} \right) \\
&= \frac{c+i}{i} \cdots \frac{b+c+i-1}{b+i-1} \\
&= \frac{(i+b+c-1)!}{c!(i+b-1)!} \cdot \frac{c!(i-1)!}{(i+c-1)!} \\
&= \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}}
\end{aligned}$$

woraus nun sofort:

$$\prod_{i=1}^a \prod_{j=1}^b \prod_{k=1}^c \frac{i+j+k-1}{i+j+k-2} = \prod_{i=1}^a \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}}$$

folgt.

Mit dieser handlicheren Formel gilt nun offenbar

$$\begin{aligned}
& \prod_{i=1}^{a-1} \left(\frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \right)^2 - \prod_{i=1}^{a-1} \frac{\binom{i+b+c-1}{c+1} \binom{i+b+c-1}{c-1}}{\binom{i+c}{c+1} \binom{i+c-2}{c-2}} \\
&= \left(1 - \frac{bc}{(a+b-1)(a+c-1)} \right) \prod_{i=1}^{a-1} \left(\frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \right)^2 \\
&= \left(1 - \frac{bc}{(a+b-1)(a+c-1)} \right) \frac{\binom{a+b+c-2}{c} \binom{a+c-1}{c}}{\binom{a+c-2}{c} \binom{a+b+c-1}{c}} \prod_{i=1}^a \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \prod_{i=1}^{a-2} \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \\
&= \frac{(a+b-1)(a+c-1) - bc}{(a+b-1)(a+c-1)} \cdot \frac{(a+c-1)(a+b-1)}{(a-1)(a+b+c-1)} \prod_{i=1}^a \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \prod_{i=1}^{a-2} \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \\
&= \prod_{i=1}^a \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}} \prod_{i=1}^{a-2} \frac{\binom{i+b+c-1}{c}}{\binom{i+c-1}{c}}
\end{aligned}$$

und damit die Behauptung. \square

Im oft betrachteten Fall des regelmäßigen Hexagons $\mathcal{H}_{a,a,a}$, $a \geq 1$ vereinfacht sich die angegebene Formel zu

$$\prod_{i=1}^a \prod_{j=1}^a \prod_{k=1}^a \frac{i+j+k-1}{i+j+k-2} = \prod_{i=1}^a \prod_{j=1}^a \frac{i+j+a-1}{i+j-1} = \prod_{i=1}^a \frac{\binom{i+2a-1}{a}}{\binom{i+a-1}{a}}.$$

In dieser Form ist die Formel beispielsweise in [1] als Lösung des Problems D zu finden, welches die etwas einfachere Frage nach der Anzahl von Teilungen von $\mathcal{H}_{n,n,n}$, $n \in \mathbb{N}$ als eine Anwendung des Lemmas von Gessel-Viennot beantwortet. Im noch einfacheren Fall $a = 1, b = c = n$ wird eine direkte Verbindung mit den Catalanzahlen

$$C_n := \frac{1}{n+1} \binom{2n}{n}$$

deutlich. Es ist dann

$$\#\mathcal{L}_{\mathcal{H}_{1,n,n}} = \prod_{i=1}^n \frac{\binom{2n+i-1}{n}}{\binom{n+i-1}{n}} = \binom{2n}{n} = (n+1)C_n.$$

Tatsächlich korrespondieren die möglichen Pfade der de Bruijn-Darstellung einer $1 \times n \times n$ Lozengeteilung mit dem sogenannten „stair case walk“-Pfaden, siehe z.B. [39] für eine Definition.

2.4 Wachstumsverhalten der Anzahl Teilungen

Es soll noch eine Abschätzung des Wachstums der Anzahl Lozengeteilungen des Hexagons $\mathcal{H}_{n,n,n}$ gegeben werden. Diese basiert direkt auf empirischen Versuchen mit dem Computer-Algebra-System Maple. Es erlaubte die Anzahl Teilungen von für große n in logarithmischer Skala zu betrachten. Auffällig ist die Parabelform der Kurve, deren Aussehen schließlich zu den nachfolgenden Betrachtungen führte.

Mittels der im vorigen Kapitel vorgestellten und bewiesenen Formel erhält man zunächst die in Tabelle 1 aufgelisteten (teilweise gerundeten) Anzahlen von Teilungen von $n = 1$ bis $n = 30$. Es zeigt sich ganz deutlich, daß der Zustandsraum schnell an Größe zunimmt.

Ziel des nächsten Abschnittes ist es, das exponentielle Wachstum zu zeigen.

n	$\#\mathcal{L}_{\mathcal{H}_{n,n,n}}$	n	$\#\mathcal{L}_{\mathcal{H}_{n,n,n}}$	n	$\#\mathcal{L}_{\mathcal{H}_{n,n,n}}$
1	2	11	$1,3 \cdot 10^{41}$	21	$1,5 \cdot 10^{150}$
2	20	12	$9,1 \cdot 10^{48}$	22	$6,8 \cdot 10^{164}$
3	980	13	$3 \cdot 10^{57}$	23	$1,5 \cdot 10^{180}$
4	232848	14	$4,8 \cdot 10^{66}$	24	$1,5 \cdot 10^{196}$
5	$2,6 \cdot 10^8$	15	$3,6 \cdot 10^{76}$	25	$7,7 \cdot 10^{212}$
6	$1,5 \cdot 10^{12}$	16	$1,3 \cdot 10^{87}$	26	$1,9 \cdot 10^{230}$
7	$3,9 \cdot 10^{16}$	17	$2,4 \cdot 10^{98}$	27	$2,2 \cdot 10^{248}$
8	$5,1 \cdot 10^{21}$	18	$2 \cdot 10^{110}$	28	$1,2 \cdot 10^{267}$
9	$3,1 \cdot 10^{27}$	19	$8,1 \cdot 10^{122}$	29	$3,2 \cdot 10^{286}$
10	$9,3 \cdot 10^{33}$	20	$1,6 \cdot 10^{136}$	30	$4,2 \cdot 10^{306}$

Tabelle 1: Liste der Anzahlen Lozengeteilungen für regelmäßige Hexagons

Es ergab sich aus der Betrachtung des logarithmischen Plots und einigen numerischen Näherungen mit Maple

Satz 2.4.1 *Es gilt*

$$e^{\frac{n^2}{2}} \leq \#\mathcal{L}_{H_{n,n,n}} \leq e^{n^2}$$

für alle $n \geq 4$.

Um diese Aussage zu beweisen liefert das folgende Lemma zunächst eine Ungleichung, die im Beweis des Satzes eine essentielle Rolle spielt.

Lemma 2.4.1 *Es gilt für alle $n \geq 4$*

$$\frac{3}{2} \leq \frac{1}{n}\rho(n) \leq 2,$$

wobei

$$\rho(n) := \log \left(\frac{\binom{4n}{2n-1} \binom{4n+1}{2n} \binom{3n}{n}}{\binom{4n+1}{n} \binom{4n}{n} \binom{2n}{n}} \right).$$

Beweis. Es gilt zunächst

$$\begin{aligned} \rho(n) &= \log \left(\frac{\binom{4n}{2n-1} \binom{4n+1}{2n} \binom{3n}{n}}{\binom{4n+1}{n} \binom{4n}{n} \binom{2n}{n}} \right) \\ &= \log \left(\frac{4n!(4n+1)!3n!n!(3n+1)!n!3n!n!}{(2n-1)!(2n+1)!2n!(2n+1)!2n!n!(4n+1)!4n!2n!} \right) \\ &= \log \left(\frac{(3n+1)2n}{(2n+1)^2} \right) + 3 \log \left(\frac{3n!n!}{2n!^2} \right) \end{aligned} \quad (5)$$

Daraus folgt nun weiter

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n}\rho(n) &= \lim_{n \rightarrow \infty} \underbrace{\frac{1}{n} \log \left(\frac{(3n+1)2n}{(2n+1)^2} \right)}_{=0} + \lim_{n \rightarrow \infty} \frac{3}{n} \log \left(\frac{3n!n!}{2n!^2} \right) \\ &= \lim_{n \rightarrow \infty} \frac{3}{n} \log \left(\frac{3n!n!}{2n!^2} \right) \end{aligned}$$

Mit Hilfe der Stirlingformel

$$n! \sim \frac{\sqrt{2\pi n} \cdot n^n}{e^n}$$

kann man weiter berechnen

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{3}{n} \log \left(\frac{3n!n!}{2n!^2} \right) \\ &= \lim_{n \rightarrow \infty} \frac{3}{n} \log \left(\frac{\sqrt{6\pi n} \cdot (3n)^{3n}}{e^{3n}} \cdot \frac{\sqrt{2\pi n} \cdot n^n}{e^n} \cdot \frac{e^{4n}}{(\sqrt{4\pi n} \cdot (2n)^{2n})^2} \right) \\ &= \lim_{n \rightarrow \infty} \frac{3}{n} \log \left(\frac{\sqrt{3}}{2} \left(\frac{3^3}{2^4} \right)^n \right) \\ &= \log \left(\frac{\sqrt{3}}{2} \right) \underbrace{\lim_{n \rightarrow \infty} \frac{3}{n}}_{=0} + \lim_{n \rightarrow \infty} \log \left(\frac{3^3}{2^4} \right)^3 = \log \left(\frac{3^9}{2^{12}} \right), \end{aligned}$$

d.h.

$$\lim_{n \rightarrow \infty} \frac{1}{n} \rho(n) = \log \left(\frac{3^9}{2^{12}} \right) \approx 1,569744431 < 2.$$

Daraus folgt unmittelbar

$$\rho(n) \in \mathcal{O}(n).$$

Weiterhin ist $\rho(n)$ monoton steigend, denn

$$\begin{aligned} & \rho(n+1) - \rho(n) \\ &= \log \left(\frac{(3n+4)2(n+1)(2n+1)^2}{(2n+3)^2(3n+1)2n} \right) + 3 \log \left(\frac{(3n+3)!(n+1)!(2n)!^2}{(2n+2)!^2(3n)!n!} \right) \\ &= \log \left(\frac{(3n+4)(n+1)(2n+1)^2}{(2n+3)^2(3n+1)n} \right) + 3 \log \left(\frac{3(n+1)(3n+2)(3n+1)(n+1)}{4(n+1)^2(2n+1)} \right) \\ &= \log \left(\frac{3^3}{2^6} \cdot \frac{n+1}{n} \cdot \frac{(3n+4)(3n+2)^3(3n+1)^2}{(2n+1)^4(2n+3)^2} \right) \\ &\geq \log \left(\left(\frac{\sqrt{3}}{2} \cdot \frac{3n+1}{2n+3} \right)^6 \underbrace{\frac{n+1}{n}}_{\geq 1} \right) \geq 6 \log \left(\frac{\sqrt{3}}{2} \cdot \frac{3n+1}{2n+3} \right) \end{aligned}$$

Für die Positivität der Differenz $\rho(n+1) - \rho(n)$ betrachte nun

$$\frac{\sqrt{3}}{2} \left(\frac{3n+1}{2n+3} \right) \geq 1 \Leftrightarrow n \geq \frac{6 - \sqrt{3}}{\sqrt{3} - 4} \approx 3,568064660 > 3$$

Daraus folgt die Behauptung des Lemmas. \square

Nun kann der Satz bewiesen werden.

Beweis. Der Satz wird mittels Induktion bewiesen.
Die Behauptung ist nach logarithmieren äquivalent zu

$$\frac{n^2}{2} \leq \log \prod_{i=1}^n \frac{\binom{2n+i-1}{n}}{\binom{n+i-1}{n}} \leq n^2.$$

Für $n = 4$ gilt nun

$$8 = \frac{16}{2} \leq \log \frac{\binom{8}{4} \binom{9}{4} \binom{10}{4} \binom{11}{4}}{\binom{4}{4} \binom{5}{4} \binom{6}{4} \binom{7}{4}} = \log(232848) \approx 12,35814116 < 16.$$

Man nimmt nun an, die Behauptung gelte für $n \in \mathbb{N}$ mit $n > 4$. Wegen

$$\binom{2n+i+1}{n+1} = \frac{(2n+i+1)(2n+i)}{(n+1)(n+i)} \binom{2n+i-1}{n}$$

und

$$\binom{n+i}{n+1} = \frac{n+i}{n+1} \binom{n+i-1}{n}$$

gilt

$$\begin{aligned} \log \prod_{i=1}^{n+1} \frac{\binom{2n+i+1}{n+1}}{\binom{n+i}{n+1}} &= \log \prod_{i=1}^{n+1} \frac{(2n+i+1)(2n+i)}{(n+i)^2} + \log \frac{\binom{3n}{n}}{\binom{2n}{n}} + \log \prod_{i=1}^n \frac{\binom{2n+i-1}{n}}{\binom{n+i-1}{n}} \\ &= \rho(n) + \log \prod_{i=1}^n \frac{\binom{2n+i-1}{n}}{\binom{n+i-1}{n}} \end{aligned}$$

nach der Definition von $\rho(n)$ im vorhergehenden Lemma.
Nach Induktionsvoraussetzung genügt es

$$\frac{(n+1)^2}{2} = \frac{n^2}{2} + n + \frac{1}{2} \leq \rho(n) + \frac{n^2}{2}$$

resp.

$$\rho(n) + n^2 \leq n^2 + 2n + 1 = (n+1)^2$$

zu zeigen.

Zusammenfassen lässt sich dies zu

$$n + \frac{1}{2} \leq \rho(n) \leq 2n + 1.$$

Nach Division der Ungleichung durch n ergibt sich die äquivalente Behauptung

$$1 + \frac{1}{2n} \leq \frac{1}{n} \rho(n) \leq 2 + \frac{1}{n}.$$

Dank Lemma gilt nun für $n \geq 4$

$$1 + \frac{1}{2n} \leq \frac{3}{2} \leq \frac{1}{n} \rho(n) \leq 2 \leq 2 + \frac{1}{n}.$$

Damit ist die Behauptung gezeigt. \square

Bemerkung. Mit Hilfe von Maple kann man berechnen, daß die Abschätzung aus dem Satz schon ab $n = 1$ gilt, denn die Berechnungen ergeben

$e^{\frac{n^2}{2}}$	$\#\mathcal{L}_{\mathcal{H}_{n,n,n}}$	e^{n^2}
1,648721271	2	2,718281828
7,389056099	20	54,59815003
90,01713130	980	8103,083928

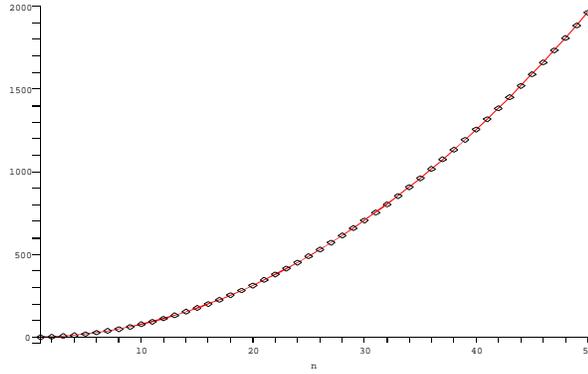


Abbildung 16: Logarithmischer Plot der Anzahl Lozenge-Teilungen von $\mathcal{H}_{n,n,n}$ für $n = 1 \dots 50$ (Punkte) und überlagerter logarithmischer Plot der Funktion $e^{0,779538296 \cdot x^2}$ (rot)

Bemerkung. Die Argumentation des Satzes lässt sich vermutlich analog auch auf die verschärfte Behauptung

$$e^{0,7n^2} \leq \#\mathcal{L}_{\mathcal{H}_{n,n,n}} \leq e^{0,8n^2}$$

anwenden.

Betrachtet man nocheinmal die Ausgangsungleichung so ergibt sich für $n > 6$

$$\begin{aligned} \log \prod_{i=1}^n \frac{\binom{2n+i-1}{n}}{\binom{n+i-1}{n}} &= \sum_{i=1}^n \left(\log \binom{2n+i-1}{n} - \log \binom{n+i-1}{n} \right) \\ &= \sum_{i=1}^n (\log(2n+i-1)! - 2 \log(n+i-1)! + \log(i-1)!) \\ &\geq \int_{\frac{1}{2}}^n \log \left(\frac{\Gamma(2n+x)}{\Gamma(n+x)^2 \Gamma(x)} \right) dx =: f(n) \end{aligned}$$

Ein Plot mittels Maple von $f(n)$ legt die Vermutung nahe, daß f eine quadratische Funktion in n ist.

Tatsächlich scheint

$$k := \lim_{n \rightarrow \infty} \frac{1}{n^2} \int_{\frac{1}{2}}^n \log \left(\frac{\Gamma(2n+x)}{\Gamma(n+x)^2 \Gamma(x)} \right) dx$$

gegen eine Konstante zu konvergieren. Nach einer numerischen Näherung ergab sich $k \approx 0,7795382960$.¹⁶

¹⁶Kurz vor Abschluss dieser Arbeit ist dem Autoren noch der Artikel [6] in die Hände

Betrachtet man nun die Approximation

$$e^{kn^2},$$

so ist dies eine gute untere Schranke für $\#\mathcal{L}_{\mathcal{H}_{n,n,n}}$, siehe auch Abb. 16.

Die vorangegangenen Ausführungen zeigen nocheinmal ganz deutlich, daß der Versuch alle möglichen Lozengeteilungen zu berechnen, um dann mittels der Pseudo-Inversen der Verteilungsfunktion Stichproben zu nehmen, allein wegen der Größe der Menge aller Teilungen zwangsläufig scheitern muss.

Die folgenden Kapitel werden Auswege aus diesem Dilemma durch Konstruktion einer Kopplung von Markovketten aufzeigen.

gefallen, in dem auf Seite 4 Korollar 3 das genaue Wachstumsverhalten mit Hilfe der Summenformel von Euler-MacLaurin berechnet wurde. Für $a = b = c = 1$ und $m = 0$ ergibt sich mit der auf Seite 5 der Referenz angegebenen Formel und der Notation des Korollars obiges k zu $0,5 \cdot (9 \log 3 - 12 \log 2) \approx 0,784872216$, was gut mit der durchgeführten numerischen Approximation übereinstimmt.

3 Coupling From The Past (CFTP)

In den folgenden Abschnitten sollen neben der notwendigen Notation auch einige wichtige Sätze und Definitionen Erwähnung finden. Daneben werden die wichtigsten Begriffe und Methoden entwickelt, um *exakt* nach der Gleichgewichtsverteilung verteilte Stichproben von zufälligen Lozengeteilungen zu erzeugen. Wichtigster Punkt ist dabei neben der Kopplung von Markovketten der Algorithmus von Propp und Wilson, der oft unter dem englischen Schlagwort „Coupling From The Past“ oder, übersetzt, „Kopplung aus der Vergangenheit“, auftaucht. Mit Hilfe eines Monotoniebegriffes und einer partiellen Ordnung auf dem Zustandsraum kann dieser Algorithmus noch wesentlich effizienter gemacht werden.

3.1 Markovketten als M-Folgen und Notation

Die Sichtweise, eine Markovkette nicht als durch die Übergangsmatrix determinierten stochastischen Prozess, sondern als iterierte, zufällige Funktion zu sehen, findet man z.B. in [15] und wird insbesondere bei der Simulation von Markovketten mit Hilfe eines Rechners benutzt. Es zeigt sich, daß beide Definitionen äquivalent sind, in dem Sinne, daß zu jeder homogenen, diskreten Markovkette eine iterierte, zufällige Funktion konstruiert werden kann und umgekehrt, dies wird später präzisiert.

Zuvor werden jedoch noch einige Definitionen und wichtige Sätze gegeben.

Für die folgenden Definitionen ist eine Markovkette $(X_n)_{n \geq 0}$ als Markovkette mit einem endlichen Zustandsraum E und der Anfangsverteilung ν_0 zu verstehen.¹⁷ Es wird weiterhin die Homogenität vorausgesetzt und mit Q sei die zugehörige Übergangs- oder Transitionsmatrix bezeichnet.

Definition 3.1.1 (Kommunikation) *Seien $x, y \in E$, dann schreibe*

$$x \rightarrow y,$$

falls ein $k \in \mathbb{N}$ existiert, so daß $Q^k(x, y) > 0$. Man sagt x und y kommunizieren miteinander, wenn $x \rightarrow y$ und $y \rightarrow x$. Abkürzend schreibt man dafür

$$x \leftrightarrow y.$$

Bemerkung: Falls $x \rightarrow y$ und $y \rightarrow x$ gilt zwar

$$\exists k_1 : Q^{k_1}(x, y) > 0$$

und

$$\exists k_2 : Q^{k_2}(y, x) > 0,$$

aber nicht zwangsläufig

$$k_1 = k_2.$$

Mit Hilfe der Definition von Kommunikationsklassen kann man einfach die Irreduzibilität einer Transitionsmatrix zeigen.

¹⁷Im folgenden sei daher E resp. die Markovkette endlich genannt.

Definition 3.1.2 (Kommunikationsklasse) Der Zustandsraum einer endlichen Markovkette lässt sich eindeutig in disjunkte Teilmengen E_i zerlegen, so daß

$$E = \bigsqcup_i E_i$$

gilt und für alle x, y gilt $x \rightarrow y$, wenn $x, y \in E_i$ und $x \not\rightarrow y$, wenn $x \in E_i$ und $y \in E_j$ für $i \neq j$.

Man nennt die E_i dieser Zerlegung Kommunikationsklassen.¹⁸

Definition 3.1.3 (Irreduzibilität) Die zu einer endlichen Markovkette gehörende Übergangsmatrix Q heißt irreduzibel, wenn für je zwei Zustände $i, j \in E$ ein k existiert, so daß $Q^k(i, j) > 0$.

Damit kann man nun zeigen

Lemma 3.1.1 Q ist irreduzibel genau dann, wenn E nur aus einer Kommunikationsklasse besteht.

Beweis. Der Beweis und weitere Ausführungen sind beispielsweise in [4] zu finden. \square

Definition 3.1.4 (stationäre Verteilungen) Die Linkseigenvektoren π mit

$$\pi Q = \pi$$

mit der Eigenschaft

$$\pi(E) := \sum_{i \in E} \pi(i) = 1$$

werden stationäre Maße oder stationäre Verteilungen genannt, manchmal auch Gleichgewichtsverteilung.

Ist die Übergangsmatrix Q irreduzibel, so hat die Eigenwertgleichung genau eine Lösung mit positiven Einträgen, so daß eine Renormierung zu einem Wahrscheinlichkeitsmaß möglich wird.

Definition 3.1.5 (Konvergenz ins Equilibrium) Für eine diskrete Markovkette $(X_n)_{n \geq 0}$ und assoziierter irreduzibler Übergangsmatrix Q und Startbedingung ν_0 spricht man von der Konvergenz ins Gleichgewicht oder Equilibrium, wenn $(X_n)_{n \geq 0}$ in Verteilung gegen das einzige stationäre Maß konvergiert, wenn $n \rightarrow \infty$.

Bemerkung Oft ist in der Literatur nicht explizit angegeben, ob die Definition einer Markovkette bereits die Startbedingung beinhaltet. Fehlt diese Angabe so wird meistens nur die Übergangsmatrix betrachtet und die Startbedingung

¹⁸Da „ \leftrightarrow “ eine Äquivalenzrelation ist, sind die Kommunikationsklassen gerade die Äquivalenzklassen von E bzgl. der Kommunikationsrelation.

ist beliebig. Dann spricht man von Konvergenz ins Equilibrium, wenn die betrachtete Markovkette für alle Startbedingungen in Verteilung gegen die stationäre Verteilung konvergiert. Besonders im Falle periodischer Übergangsmatrizen kann die Konvergenz von der Startbedingung abhängen (für die Definition der Aperiodizität siehe 3.1.5).

Weiterhin lässt sich leicht zeigen

Lemma 3.1.2 *Ist die Transitionsmatrix Q einer Markovkette symmetrisch und irreduzibel, so ist die stationäre Verteilung π gerade die Gleichverteilung auf E , d.h.*

$$\pi Q = \pi = \frac{1}{\#E} \mathbf{1} = \frac{1}{\#E} (1, 1, 1, \dots, 1).$$

Beweis. Man berechnet zunächst das reversible Maß ν von Q , d.h. es ist das Gleichungssystem

$$\nu(i)Q(i, j) = \nu(j)Q(j, i)$$

zu lösen. Es ergibt sich wegen der vorausgesetzten Symmetrie

$$\nu(i)Q(i, j) = \nu(i)Q(j, i) = \nu(j)Q(j, i) \Leftrightarrow \nu(i) = \nu(j)$$

für alle $i, j \in E$. Wegen der Normierungseigenschaft $\sum_{i \in E} \nu(i) = 1$ ist ν die Gleichverteilung auf E .

Wegen der Endlichkeit sind reversible Maße stationär; wegen der Irreduzibilität von Q ist ν gerade das eindeutig bestimmte stationäre Maß und damit die Behauptung gezeigt. \square

Definition 3.1.6 (aperiodisch) *Es sei $(X_n)_{n \geq 0}$ eine Markovkette mit assoziiertem irreduzibler Übergangsmatrix Q und Startbedingung ν_0 . Man definiert zunächst die Menge der Rückkehrzeiten durch*

$$T_i = \{n \geq 1 : X_n = i | X_0 = i\} = \{n \geq 1 : Q(i, i)^n > 0\}$$

für einen Zustand $i \in E$.

Dann heißt die Markovkette aperiodisch, wenn

$$\text{ggT}(T_i) = 1$$

gilt.

Definition 3.1.7 (ergodisch) *Man nennt eine endliche Markovkette ergodisch, wenn ihre Übergangsmatrix aperiodisch und irreduzibel ist.*

Nun kann folgender wichtiger Zusammenhang gezeigt werden

Satz 3.1.1 *Zu jeder diskreten Markovkette $(X_n)_{n \geq 0}$ existiert eine Funktion f und eine Folge unabhängiger, gleichverteilter Zufallsvariablen $(\Theta_n)_{n \geq 1}$, so daß für beliebige, deterministische Anfangsbedingung $X_0 \in E$*

$$X_{n+1} = f(X_n, \Theta_n), \quad n \geq 1$$

gilt. Man nennt f dabei auch Updatefunktion.

Beweis. Den Beweis findet man in [30], Kapitel 1; er wird hier nocheinmal wiedergegeben.

Es sei Q die zu $(X_n)_{n \geq 0}$ assoziierte Transitionsmatrix und die n Zustände des Zustandsraumes seien mit x_1, \dots, x_n bezeichnet.

Betrachtet man das Einheitsintervall $I := [0, 1]$ und konstruiert $(\Theta_n)_{n \geq 1}$ mit $\Theta_i \sim \mathcal{U}_I$, so setze zunächst

$$F_{j,k} := \sum_{i=1}^k Q(x_j, x_i).$$

Eine Funktion $f : E \rightarrow E$ mit

$$f(x_j, \Theta_i) = x_k, \text{ falls } F_{j,k-1} < \Theta \leq F_{j,k}$$

erfüllt die Eigenschaften einer Updatefunktion.

Dafür genügt zu bemerken

$$\mathbb{P}(f(x_j, \Theta_i) = x_k) = \mathbb{P}(F_{j,k-1} < \Theta_i \leq F_{j,k}) = Q(x_j, x_k).$$

□

Bemerkung. Die Darstellung einer Markovkette durch eine iterierte, zufällige Funktion ist nicht eindeutig, so kann das im Beweis konstruierte $f(x_i, \Theta)$ stets durch $f(x_i, 1 - \Theta)$ ersetzt werden.

Der Satz liefert also nur eine Existenzaussage, aber keine Eindeutigkeit.

3.2 Kopplung von Markovketten mit diskretem Zustandsraum

Bevor die Kopplung von endlichen Markovketten besprochen werden kann, werden noch einige Begriffe und Konzepte benötigt.

Definition 3.2.1 (totale Variationsdistanz) Für zwei Wahrscheinlichkeitsmaße α und β auf E definiert man

$$d_{TV}(\alpha, \beta) := \frac{1}{2} \sum_{i \in E} |\alpha(i) - \beta(i)|.$$

Man bemerkt, daß außerdem d_{TV} eine Metrik auf dem Raum der Wahrscheinlichkeitsmaße auf E definiert. Dies folgt direkt aus den Eigenschaften des Betrages.

Im folgenden wird oft etwas unkorrekt von der totalen Variationsdistanz zweier Zufallsvariablen auf E gesprochen, damit seien stets die zugehörigen Wahrscheinlichkeitsmaße gemeint. Aufgrund dessen setze man abkürzend

$$d_{TV}(X, Y) := d_{TV}(\mathbb{P}_X, \mathbb{P}_Y).$$

Man kann die totale Variationsdistanz noch anders ausdrücken, dies zeigt

Lemma 3.2.1 Seien X und Y Zufallsvariablen auf E . Dann gilt

$$d_{TV}(X, Y) = \sup_{A \in E} |\mathbb{P}_X(A) - \mathbb{P}_Y(A)| = \sup_{A \in E} (\mathbb{P}_X(A) - \mathbb{P}_Y(A))$$

Beweis. Der Beweis findet sich beispielsweise in [5], Seite 126, Lemma 1.1. \square

Es kann jetzt die Konvergenz in Variation definiert werden. Dazu

Definition 3.2.2 (Konvergenz in Variation) Sei $(\alpha_n)_{n \geq 0}$ eine Folge von Wahrscheinlichkeitsmaßen auf E . Man sagt die Folge konvergiert in Variation gegen das Wahrscheinlichkeitsmaß β , wenn

$$\lim_{n \rightarrow \infty} d_{TV}(\alpha_n, \beta) = 0.$$

Analog kann man nun für eine ergodische, homogene Markovkette $(X_n)_{n \geq 0}$ die Konvergenz in Variation gegen π festhalten, indem gefordert wird

$$\lim_{n \rightarrow \infty} \sum_{i \in E} |\mathbb{P}(X_n = i) - \pi(i)| = \lim_{n \rightarrow \infty} d_{TV}(\nu_0 Q^n, \pi) = 0.$$

Nun kann die Kopplung von (zwei) Markovketten definiert werden.¹⁹

Definition 3.2.3 (Kopplung) Es seien $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$ zwei endliche Markovketten mit Zustandsraum E und beliebigen Anfangsverteilungen ν_0 und μ_0 . Man nennt einen Prozess $(Z_n)_{n \geq 0}$ mit Zustandsraum $E \times E$ und Anfangsbedingung (ν_0, μ_0) eine Kopplung von $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$, wenn

$$\pi_1(Z_n) \sim X_n, \quad \pi_2(Z_n) \sim Y_n,$$

wobei π_1, π_2 die kanonischen Projektionen sind, d.h. es gilt $\pi_i(t_1, t_2) = t_i$.

Da die Randverteilungen von $(Z_n)_{n \geq 0}$ nur in Verteilung mit den originalen Prozessen übereinstimmen müssen, hat man bei der eigentlichen Definition von $(Z_n)_{n \geq 0}$ großen Spielraum, denn es muss nicht notwendig gelten, daß

$$\mathbb{P}_{Z_n} = \mathbb{P}_{X_n} \otimes \mathbb{P}_{Y_n}$$

ist. Man kann also die Vektorkomponenten auch so wählen, daß sie abhängig voneinander sind.

Man bezeichnet eine Kopplung mit unabhängigen Marginalprozessen auch als *Doebelin-Kopplung*.

Im folgenden sei die Kopplung zweier Prozesse trotzdem mit

$$(Z_n)_{n \geq 0} = ((X_n, Y_n))_{n \geq 0}$$

bezeichnet, auch wenn die definierte Kopplung nur in den Randverteilungen mit den originalen Prozessen übereinstimmt. Ist der Kontext klar oder die genaue

¹⁹Für eine wesentlich allgemeinere Definition der Kopplung beliebiger Prozesse siehe [32].

Gestalt der Kopplung unwichtig, so wird auf die genaue Angabe von $(Z_n)_{n \geq 0}$ verzichtet und schlicht von einer Kopplung von $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$ gesprochen. Insbesondere wird im Folgenden oft X_n mit den kanonischen Projektionen $\pi_1(Z_n)$ und Y_n mit $\pi_2(Z_n)$ gleichgesetzt.

Für die nachfolgenden Ausführungen ist noch eine weitere Eigenschaft wichtig, die

Definition 3.2.4 (Koaleszenzeigenschaft) *Eine Kopplung zweier endlicher Markovketten $(Z_n)_{n \geq 0} = ((X_n, Y_n))_{n \geq 0}$ hat die Koaleszenzeigenschaft, wenn ein n_0 existiert, so daß für alle $m \in \mathbb{N}$ mit $m \geq n_0$*

$$X_m = Y_m \Rightarrow X_{m+1} = Y_{m+1}$$

gilt.

Die Trajektorien einer Kopplung mit Koaleszenzeigenschaft „verkleben“ also nach dem Zusammentreffen und entfernen sich nicht mehr voneinander.

Mit Hilfe der Kopplung kann die Distanz zum Equilibrium abgeschätzt werden, dazu zunächst

Definition 3.2.5 *Es sei für eine irreduzible, endliche und homogene Markovkette $(X_n)_{n \geq 0}$ mit Übergangsmatrix Q*

$$d(n) := \max_{i \in E} d_{TV}(\delta_i Q^n, \pi) = \max_{i \in E} \sum_{j \in E} |(\delta_i Q^n)(j) - \pi(j)|.$$

Analog definiert man eine totale Variationsdistanz maximiert über alle Anfangsbedingungen durch

Definition 3.2.6 *Es sei für eine irreduzible, endliche und homogene Markovkette $(X_n)_{n \geq 0}$ und Übergangsmatrix Q*

$$\bar{d}(n) := \max_{i, j \in E} d_{TV}(\delta_i Q^n, \delta_j Q^n) = \max_{i, j \in E} \sum_{k \in E} |(\delta_i Q^n)(k) - (\delta_j Q^n)(k)|.$$

Die beiden Metriken $d(n)$ und $\bar{d}(n)$ sind äquivalent²⁰ und es gilt

$$d(n) \leq \bar{d}(n) \leq 2d(n).$$

Außerdem ist $\bar{d}(n)$ submultiplikativ, d.h. es gilt für $n, m \in \mathbb{N}$

$$\bar{d}(n + m) \leq \bar{d}(n)\bar{d}(m).$$

Es zeigt sich außerdem, daß $\bar{d}(n)$ nicht steigend ist. Dies garantiert im Folgenden, daß eine ε -Umgebung nach Erreichen nicht mehr verlassen wird (siehe Mischungszeit weiter unten) und [2], sowie [30].

Man definiert noch

²⁰ E ist endlich, daher sind alle Metriken äquivalent.

Definition 3.2.7 (Koaleszenzzeit) Für eine Kopplung $(Z_n)_{n \geq 0}$ endlicher Markovketten $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$ mit den Startverteilungen $\nu_0 = \delta_i$ und $\mu_0 = \delta_j$ mit $i, j \in E$ und Koaleszenzeigenschaft, nennt man

$$\begin{aligned}\tau^{i,j} &= \inf\{n \geq 0 : X_n = Y_n | X_0 = i, Y_0 = j\} \\ &= \inf\{n \geq 0 : d_{TV}(\delta_i Q^n, \delta_j Q^n) = 0\}\end{aligned}$$

*Koaleszenzzeit.*²¹

Mit Hilfe der Koaleszenzzeit kann nun die Distanz vom Equilibrium in totaler Variation abgeschätzt werden, es gilt

Lemma 3.2.2 (Kopplungsungleichung) Für eine Kopplung endlicher Markovketten zweier identischer Kopien $(X_n)_{n \geq 0}$ und $(Y_n)_{n \geq 0}$ mit den Startverteilungen $\nu_0 = \delta_i$ und $\mu_0 = \delta_j$ mit $i, j \in E$ und Koaleszenzeigenschaft gilt

$$d_{TV}(X_n, Y_n) = d_{TV}(\delta_i Q^n, \delta_j Q^n) \leq \mathbb{P}(\tau^{i,j} > n).$$

Beweis. Der Beweis folgt sofort aus der Darstellung der Variationsdistanz aus 3.2.1, denn für beliebiges $A \subset E$

$$\begin{aligned}\mathbb{P}_{X_n}(A) - \mathbb{P}_{Y_n}(A) &= \mathbb{P}_{X_n}(A, \tau^{i,j} > n) + \mathbb{P}_{X_n}(A, \tau^{i,j} \leq n) \\ &\quad - \mathbb{P}_{Y_n}(A, \tau^{i,j} > n) - \mathbb{P}_{Y_n}(A, \tau^{i,j} \leq n) \\ &= \mathbb{P}_{X_n}(A, \tau^{i,j} > n) - \mathbb{P}_{Y_n}(A, \tau^{i,j} > n) \\ &\leq \mathbb{P}_{X_n}(A, \tau^{i,j} > n).\end{aligned}$$

Nimmt man auf beiden Seiten das Maximum, so erhält man die Aussage des Lemmas. \square

Aus dem Lemma erhält man als Folgerung

$$d(n) \leq \max_{i,j \in E} \mathbb{P}(\tau^{i,j} > n)$$

Eine naheliegende Fragestellung ist nun, wie lange eine Markovkette laufen muss, damit sie resp. ihre Verteilung in einer ε -Umgebung (bzgl. der Variationsdistanz) der Gleichgewichtsverteilung ist. Dies lässt sich mit der Definition der sogenannten Mischungszeit beantworten.

Definition 3.2.8 (Mischungszeit) Es sei $(X_n)_{n \geq 0}$ eine endliche, ergodische Markovkette mit stationärer Verteilung π . Man nennt

$$\rho(\varepsilon) := \inf\{n \geq 0 : d(n) \leq \varepsilon\}$$

Mischungszeit.

²¹ $\tau^{i,j}$ ist natürlich eine Stopzeit bezüglich der von $(Z_n)_{n \geq 0}$ erzeugten Filtration und wird in der Literatur oft auch als Kopplungszeit beschrieben. Um Bedeutungs-dopplungen zu vermeiden, wurde hier eine andere Bezeichnung gewählt.

Insbesondere ist nun sicher, daß nach Erreichen des ε -Balls um die stationäre Verteilung π dieser nicht mehr verlassen wird, da $d(n)$ nicht steigend ist. Die Kenntnis der Mischungszeit kann unter anderem dazu verwendet werden, eine Simulation nach einer gewissen (festen!) Anzahl von Schritten abzubrechen und trotzdem mit hoher Wahrscheinlichkeit Stichproben zu bekommen, die höchstens um ε in totaler Variation von der stationären Verteilung abweichen.

Definition 3.2.9 (erwartete Schrittzahl bis zur Koaleszenz) *Die erwartete maximale Schrittzahl für die „schlechteste Wahl“ von Startbedingungen bis zur Koaleszenz lässt sich formalisieren durch die Definition*

$$T := \max_{i,j \in E} \mathbb{E}(\tau^{i,j}).$$

Da die Mischungszeit oft nur sehr schwierig explizit zu berechnen ist, ist der folgende Satz sehr hilfreich und stellt eine Verbindung zwischen Koaleszenz- und Mischungszeit her.

Satz 3.2.1 *Es gilt*

$$\rho(\varepsilon) \leq eT \ln(\varepsilon^{-1}).$$

Ein Beweis und weitere Ausführungen zum Konzept des schnellen Mischens finden sich in [2].

3.3 Der Coupling-From-The-Past-Algorithmus

Der „Coupling From The Past“-Algorithmus²² geht auf die Arbeiten von *James Propp* und *David B. Wilson* zurück, siehe etwa [37].

Für eine gegebene Markovkette $(X_n)_{n \geq 0}$ gibt es eine messbare Funktion f und eine Folge von identisch verteilten, unabhängigen Zufallsvariablen $(\Theta_n)_{n \geq 0}$, so daß für deterministisch gegebenes X_0 die X_{n+1} durch die Rekursion

$$X_{n+1} = f(X_n, \Theta_n) =: f_{\Theta_n}(X_n)$$

für $n \geq 1$ dargestellt werden können.²³

Betrachtet man nun eine endliche, ergodische Markovkette, so möchte man mit Hilfe dieser Darstellung zufällige Stichproben gewinnen, die *genau* nach der (hier einzigen) stationären Verteilung π verteilt sind. Die Ergodizität der Kette legt nahe, daß es dafür genügt die Dynamik „lange genug“ laufen zu lassen, man bildet also

$$X_M = (f_{\Theta_M} \circ f_{\Theta_{M-1}} \circ \dots \circ f_{\Theta_1} \circ f_{\Theta_0})(X_0).$$

Die Frage ist dann, wie groß M gewählt werden muss, damit X_M tatsächlich nach π verteilt ist. M ist mithin selbst eine Zufallsvariable.

²²Die Übersetzung ins Deutsche ist „Kopplung aus der Vergangenheit“; im folgenden wird meist die englisch motivierte Abkürzung „CFTP“ verwendet

²³siehe auch Abschnitt 3.1.

Zwar lässt sich mit Hilfe der Mischungszeit $\rho(\varepsilon)$ diese Frage bis auf einen Fehler beantworten, dieser ist jedoch in vielen Anwendungen störend und verfälscht die Ergebnisse. Eine hinreichende Mischung ist nach Propp und Wilson gegeben, wenn der Fehler $\varepsilon < 10^{-6}$ ist (siehe [36] Seite 2 oben).

Die Grundlegende Idee der Kopplung aus der Vergangenheit ist es nun mehrere Kopien der Dynamik zu betrachten und schrittweise in die Vergangenheit bis zur Zeit $-T$ zurückzugehen und die Dynamik *vorwärts* bis zur Zeit 0 laufen zu lassen. Man inkrementiert T und startet die Dynamik solange mit den Anfangsbedingungen neu, bis die Koaleszenz zur Zeit 0 gegeben ist.

Mit Hilfe der oben erwähnten Darstellung einer Markovkette als zufällige, iterierte Funktion lässt sich zunächst ein *stochastischer Fluß* definieren, der Auskunft darüber geben kann, wann die Dynamik abgebrochen werden kann und eine *exakt* gleichverteilte Stichprobe vorliegt.

Definition 3.3.1 (stochastischer Fluß) *Es sei eine Markovkette $(X_n)_{n \geq 0}$ mit Darstellung durch die Funktion f und die unabhängig, identisch verteilte Folge $(\Theta_n)_{n \geq 0}$ auf einem endlichen Zustandsraum E gegeben. Für fixes $T \in \mathbb{N}$ bezeichnet man den durch die folgende Rekursion gegebenen $|E|$ -dimensionalen zufälligen Vektor F_{-T} mit*

$$F_{-T}(E) = (f_{\Theta_1} \circ f_{\Theta_2} \circ \dots \circ f_{\Theta_{T-1}} \circ f_{\Theta_T})(E)$$

als stochastischen Fluß zu T . F_{-T} wird konstant genannt, wenn alle Komponenten des Vektors identisch sind.

Diesen stochastischen Fluß kann man einerseits als eine Art Aufzeichnung der Trajektorien einer Markovkette (sozusagen der Flußlinien) begreifen, die für *alle* möglichen Anfangsbedingungen die Entwicklung der Kette protokolliert; andererseits als Kopplung von $|E|$ Kopien der Originalkette, die gemäß der Updatefunktion f und $(\Theta_k)_{k \geq 0}$ gleichzeitig aktualisiert werden.

Der Grundansatz ist also die Dynamik zur Zeit „ $-\infty$ “ zu starten, dann ist wegen der Ergodizität der Kette die Verteilung der Dynamik zur Zeit 0 gerade die stationäre Verteilung. Es wird nun ein T gesucht, bei dem die Pfade bei Start in $-T$ zur Zeit 0 bereits zusammengetroffen sind, dann muß der zur Zeit 0 erreichte Zustand nach der Equilibriumsverteilung verteilt sein. Durch das schrittweise Zurückgehen in die Vergangenheit (vergrößern von T) kann dies erreicht werden.

Diesen heuristischen Ansatz präzisiert

Lemma 3.3.1 *Wenn der zu einer endlichen, ergodischen Markovkette $(X_n)_{n \geq 0}$ gehörende stochastische Fluß F_{-T} für ein $T \in \mathbb{N}$ konstant ist, so sind die Einträge von F_{-T} genau nach der Equilibriumsverteilung π verteilt.*

Beweis. siehe [37] Seite 5. \square

Das Lemma liefert das ausschlaggebende Argument für die Funktionalität des folgenden Algorithmus.

Für eine Markovkette, die durch eine Updatefunktion f und eine Folge von identisch verteilten unabhängigen Zufallsvariablen Θ_i beschrieben wird, sei der CFTP-Algorithmus nach [37], Seite 5 oben

$T = 0$;

repeat

$T = T - 1$;

$f_{-T} = \text{RandomMap}()$;

$F_{-T} = F_{-T-1} \circ f_{-T}$

until $F_{-T}(E)$ konstant

return $\pi_1(F_{-T}(E))$

wobei $\text{RandomMap}()$ eine Methode sei, die entsprechend der definierten Markovkette die Einträge des stochastischen Flusses in Abhängigkeit der Folge $(\Theta_n)_{n \geq 1}$ aktualisiert.

Man beachte dabei, daß die Θ_i in jedem Durchgang wieder benutzt werden müssen. Insbesondere müssen diese während der Laufzeit des Programms gespeichert werden.

Eine scheinbare Verbesserung des Algorithmus wäre es einfach die Kopplung in Zeitpunkt 0 zu starten und solange zu warten, bis die Pfade koaleszieren. Zwar hat die Koaleszenzzeit für dieses „Vorwärtskoppeln“ die selbe Verteilung, wie im Fall der Kopplung in die Vergangenheit, jedoch erhält man so Stichproben, die *nicht* nach der Gleichverteilung verteilt sind, siehe Fill siehe [18], Abschnitt 5.1. Dort wird auch ein Algorithmus vorgestellt, der während der Laufzeit abgebrochen werden kann und trotzdem nach der gewünschten Verteilung verteilte Stichproben liefern kann.

Eine mögliche Verbesserung des Algorithmus, der das Speichern der Θ_i vermeidet, ist unter dem Namen „Read-Once-Algorithmus“ bekannt und beispielsweise in [23] Kapitel 11 zu finden.

3.4 Monotonie und Coupling From The Past

Der im vorigen Abschnitt definierte CFTP-Algorithmus ist sehr kostenintensiv (im Sinne von Speicher und Rechenzeit), wenn der Zustandsraum der betrachteten Markovkette sehr groß ist.²⁴

Da also für jeden Zustand des Zustandsraumes eine separate Kopie der Markovkette für den CFTP-Algorithmus gestartet werden müsste, ist die Frage nach einer möglichen Vereinfachung für die Realisierung vieler Simulationen von großer Bedeutung.

²⁴Im Fallbeispiel der in dieser Arbeit behandelten Lozengeteilungen ist der Zustandsraum schon für kleine Regionen geradezu gigantisch (siehe Tabelle 1, im Kapitel 2). Der Zustandsraum „wächst exponentiell“.

Wenn es gelingt eine partielle Ordnung „ \preceq “ der Zustände auf dem Zustandsraum E zu finden und eindeutig bestimmte minimale und maximale Elemente existieren, d.h. Elemente \perp und \top , so daß für alle $x \in E$ gilt

$$\perp \preceq x \preceq \top,$$

so ist es möglich die Anzahl der notwendigen Rechenoperationen wesentlich zu reduzieren.²⁵

Wird für die Dynamik nun eine Kopplung von $|E|$ Kopien definiert und gilt für je zwei Kopien $(X_n^{(i)})_{n \geq 0}, (X_n^{(j)})_{n \geq 0}$ mit $i \neq j$

$$X_n^{(i)} \preceq X_n^{(j)} \Rightarrow X_{n+1}^{(i)} \preceq X_{n+1}^{(j)} \text{ f.s. für alle } n,$$

so sagt man, die Updatefunktion f erhalte die Ordnungsrelation \preceq .

Nun genügt es offenbar, den CFTP-Algorithmus lediglich in den extremen Startbedingungen \perp und \top zu starten, da wegen der Monotonie alle anderen Trajektorien zwischen den Trajektorien der in den extremalen Punkten gestarteten Dynamik liegen müssen. Die Koaleszenz und damit die Konstanz des stochastischen Flusses tritt nun ein, wenn die beiden extremalen Trajektorien koaleszieren, siehe zur Erläuterung auch Abbildung 17.

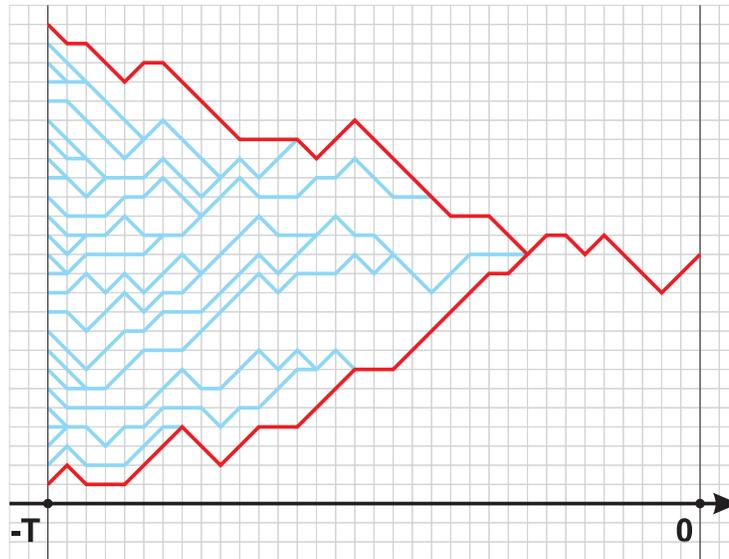


Abbildung 17: Hier dargestellt der stochastische Fluß einer Kopplung von 25 Kopien einer Markovkette mit Ordnungserhaltung. Die roten Pfade indizieren die Entwicklung der extremen Trajektorien.

²⁵Im Fall der Lozengeteilungen ist $\underline{E} = (E, \preceq)$ sogar ein Verband.

Die so erhaltene Stichprobe ist dann ebenso *exakt* nach der Equilibriumsverteilung verteilt. Dieses Verfahren nennt man *monotone Kopplung aus der Vergangenheit* und sei im Folgenden mit *mCFTP* abgekürzt.

Bemerkung. Die Möglichkeit der Benutzung des mCFTP hängt stark davon ab, ob die auf dem Zustandsraum definierte Ordnungsrelation durch die Dynamik resp. durch die Updatefunktion einer Darstellung der Kette als zufällige, iterierte Funktion erhalten wird und ist keinesfalls kanonisch gegeben.

Weitere und genauere Ausführungen hierzu, dem CFTP-Algorithmus und verwandten Fragen sind in [36], [37] und [21] zu finden.

Dieser Abschnitt basiert auf den Argumentationen in [36].

4 Simulation von uniform verteilten Lozengeteilungen

Die Idee zur Lösung des Problems des Samplings uniformverteilter Teilungen ist, eine ergodische Markovkette zu konstruieren, die als stationäre Verteilung die Gleichverteilung besitzt. Gelingt es weiter eine (partielle) Ordnung zu definieren und erhält die zu konstruierende Dynamik diese, so kann man den mCFTP-Algorithmus aus dem vorigen Kapitel verwenden, um so *exakt* gleichverteilte Teilungen zu erzeugen.

Zunächst soll die Dynamik in einer abstrakteren Form dargestellt werden, danach wird mittels der Charakterisierung von Markovketten als zufällige, iterierte Funktion eine explizite Darstellung der Dynamik gegeben, um damit schließlich eine Kopplung definieren zu können.

Um das ganze mathematisch gut fassen zu können, wird zunächst noch eine weitere Darstellungsmöglichkeit der Lozengeteilungen besprochen, die *plane partitions*.

4.1 Definition der Dynamik als zufällige, iterierte Funktion und Kopplung

Es sei zunächst definiert:

Definition 4.1.1 (plane partition) *Es sei $P = (p^{(i,j)})_{ij} \in \mathbb{R}^{b \times c}$ eine Matrix mit Einträgen aus \mathbb{N}_0 . Gilt*

- P, P^T haben monoton fallende Spalteneinträge,
- $\forall (i, j) \in \{1, \dots, b\} \times \{1, \dots, c\} : 0 \leq p^{(i,j)} \leq a,$

so heiße P eine $b \times c$ plane partition zur Magnitude a .

Die Menge aller solcher plane partitions sei mit $\mathcal{P}_{b,c}^a$ bezeichnet.

Zu jeder plane partition kann ein sogenanntes *YOUNG-Diagramm* assoziiert werden. Daher

Definition 4.1.2 (YOUNG-Diagramm) *Sei $P = (p^{(i,j)})_{ij} \in \mathcal{P}_{b,c}^a$, dann ordnet das YOUNG-Diagramm jedem $(i, j) \in \{1, \dots, b\} \times \{1, \dots, c\}$ einen Stapel von $p^{(i,j)}$ Einheitskuben zu.*

Die Abb. 18 zeigt das YOUNG-Diagramm für die plane partition $A \in \mathcal{P}_{3,3}^3$ mit

$$A = \begin{array}{|c|c|c|} \hline 3 & 3 & 2 \\ \hline 2 & 2 & 1 \\ \hline 1 & 0 & 0 \\ \hline \end{array}$$

Die Abbildung legt nahe, daß die isometrische Projektion entlang des Vektors $(1, 1, 1)$ gerade wieder eine Lozengeteilung ergibt; diese Abbildung ist bijektiv. Zeichnet man Höhengniveaulinien ein, so erhält man wieder die deBruijn-Pfade aus den vorherigen Kapiteln (in der Abbildung rot eingezeichnet).

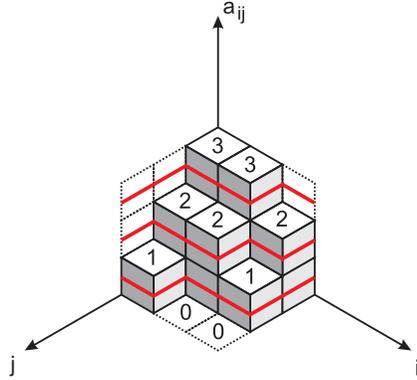


Abbildung 18: isometrische Projektion des YOUNG-Diagramms aus dem Beispiel mit eingezeichneten de-Bruijn-Pfaden (rot).

Im weiteren benötigt man einen Abstands begriff, daher definiere zuerst

$$\|P\|_L := \sum_{i=1}^b \sum_{j=1}^c |p^{(i,j)}|$$

für ein $P \in \mathcal{P}_{b,c}^a$. Man nenne diese Abbildung $\mathcal{P}_{b,c}^a \rightarrow \{0, \dots, B\}$ mit $B := abc$ im folgenden *Lozenge-Norm*.²⁶

$\|\cdot\|_L$ kann als Abzählung der Anzahl Einheitskuben im YOUNG-Diagramm von P interpretiert werden.

Mit der Lozenge-Norm kann man einen Abstands begriff auf der Menge der plane partitions definieren, der damit zum metrischen Raum wird. Es sei

$$d_L(P_1, P_2) := \|A - B\|_L = \sum_{i=1}^b \sum_{j=1}^c |p_1^{(i,j)} - p_2^{(i,j)}|,$$

für $P_1 = (p^{(i,j)})_{ij}, P_2 = (p_2^{(i,j)})_{ij} \in \mathcal{P}_{b,c}^a$.

Mit $d_L(\cdot)$ mißt man also, wieviele Einheitskuben zum YOUNG-Diagramm von P_1 hinzugefügt oder weggenommen werden müssen, um P_2 zu erhalten.

Dieses „Hinzufügen“ und „Wegnehmen“ wird die Grundlage der Definition der Markovkette \mathcal{M}_{loz} sein.

Definition 4.1.3 (Basismatrizen) Die Basiselemente $E^{(i,j)}$ von $\mathbb{R}^{b \times c}$ sind gerade die Matrizen $E^{(i,j)} = (e^{(k,l)})_{kl}$ mit $e^{(i,j)} = 1$ und $e^{(k,l)} = 0 \forall k \neq i, l \neq j$.

²⁶Der Begriff Norm überträgt sich hier aus der Normeigenschaft dieser Abbildung auf dem Ring der reellen $b \times c$ Matrizen; $\mathcal{P}_{b,c}^a$ ist selbst kein Vektorraum. Allerdings ist $\underline{P} = (P_{b,c}^a, \min, \max)$ ein distributiver Verband, wobei \min und \max komponentenweise zu verstehen sind. Im folgenden soll wegen der analogen Eigenschaften diese Abbildung trotzdem als Norm bezeichnet werden.

Es sei bemerkt, daß jede *plane partition* aus $\mathcal{P}_{b,c}^a$ durch diese Basismatrizen und normale Matrixaddition auf kanonische Weise dargestellt werden kann, denn für $P = (p^{(i,j)})_{ij} \in \mathcal{P}_{b,c}^a$ gilt

$$P = \sum_{k=1}^b \sum_{l=1}^c p^{(k,l)} E^{(k,l)}.$$

Definition 4.1.4 (κ -Nachbarschaft) *Es sei $P_1 \in \mathcal{P}_{b,c}^a$. Die κ -Nachbarschaft von P_1 sei dann*

$$V_{P_1, \kappa} := \{P_2 \in \mathcal{P}_{b,c}^a : \exists(i, j) : P_1 = P_2 \pm \kappa E^{(i,j)}\},$$

für $\kappa \geq 1$.

Aus der Forderung $P_1 = P_2 \pm \kappa E_{i,j}$ folgt im übrigen $d_L(P_1, P_2) = \kappa$. Die Umkehrung ist jedoch nicht richtig.

Nun kann \mathcal{M}_{loz} definiert werden:

Definition 4.1.5 (\mathcal{M}_{loz}) *Sei $(Z_n)_{n \geq 1}$ eine Markovkette mit $(\Omega := \text{Pot}(\mathcal{P}_{b,c}^a), \mathcal{F}, \mathbb{P})$ mit Transitionsmatrix Q , so daß*

$$Q(P_1, P_2) = \begin{cases} \frac{1}{2abc\kappa} & \text{falls } P_2 \in V_{P_1, \kappa} \\ 1 - \sum_{\kappa} \sum_{P_2' \in V_{P_1, \kappa}} Q(P_1, P_2') & \text{falls } P_1 = P_2 \\ 0 & \text{sonst} \end{cases},$$

dann heie $(Z_n)_{n \geq 0} \mathcal{M}_{loz}$.²⁷

Diese Definition scheint zunchst willkrlich. Die Definition der bergangswahrscheinlichkeiten wird aber einsichtig, wenn man die Dynamik als *iterierte, zufllige Funktion* beschreibt.

Satz 3.1.1 aus Kapitel 3 liefert die Existenz einer Darstellung von \mathcal{M}_{loz} als iterierte, zufllige Funktion; eine mgliche liefert

Definition 4.1.6 (\mathcal{M}_{loz} als iterierte, zufllige Funktion mit Werten in $\mathcal{P}_{b,c}^a$) *Es sei $(\theta_n)_{n \geq 0}$ eine Folge unabhngiger, identisch verteilter Zufallsvariable mit*

$$\theta_i := (X_i, Y_i, L_i, R_i) \sim \mathcal{U}_{\{1, \dots, b\}} \otimes \mathcal{U}_{\{1, \dots, c\}} \otimes \mathcal{U}_{\{1, \dots, a\}} \otimes \mathcal{U}_{\{0, 1\}}.$$

Definiere zur Abkrzung

$$K_n^{(1)} := \min(Z_n^{(X_{n-1}, Y_n)} - Z_n^{(X_n, Y_n)}, Z_n^{(X_n, Y_{n-1})} - Z_n^{(X_n, Y_n)})$$

$$K_n^{(2)} := \min(Z_n^{(X_n, Y_n)} - Z_n^{(X_{n+1}, Y_n)}, Z_n^{(X_n, Y_n)} - Z_n^{(X_n, Y_{n+1})})$$

²⁷Die Bezeichnung $\text{Pot}(A)$ bedeute die Potenzmenge von A .

und

$$H_n^{(1)} := \max(L_n - Z_n^{(X_n, Y_n)}, 0)$$

$$H_n^{(2)} := \max(Z_n^{(X_n, Y_n)} - L_n + 1, 0)$$

Wegen der Definition der plane partitions (fallende Spalten- und Zeileneinträge) sind $K_n^{(i)}$ und $H_n^{(i)}$ nicht negativ, $i = 1, 2$.

Dann sei $(Z_n)_{n \geq 0}$ definiert durch

$$Z_0 := P \in \mathcal{P}_{b,c}^a$$

$$Z_{n+1} := f(Z_n, \theta_n)$$

$$:= Z_n + H_n^{(1)} E^{(X_n, Y_n)} \mathbb{1}_{H_n^{(1)} > K_n^{(1)}} \mathbb{1}_{R_i \leq \frac{1}{2H^{(1)}}} - H_n^{(2)} E^{(X_n, Y_n)} \mathbb{1}_{H_n^{(2)} > K_n^{(2)}} \mathbb{1}_{R_i \geq 1 - \frac{1}{2H^{(2)}}}$$

Um die Darstellung konsistent zu machen, definiere

$$Z_n^{(0,i)}, Z_n^{(i,0)} := a$$

und

$$Z_n^{(a+1,i)}, Z_n^{(i,a+1)} := 0$$

quasi als Randbedingungen.²⁸

Zur Veranschaulichung kann man sich vorstellen, daß zu jedem Zeitpunkt zufällig ein Eintrag der plane partition Darstellung von $\pi_1(Z_n)$ und $\pi_2(Z_n)$ ausgewählt wird und zudem ein Höhenniveau. Nun wird eine Münze geworfen, ob an der gewählten Stelle mit Kuben aufgefüllt werden soll oder ob Kuben weggenommen werden (dies entspricht dem Ereignis $R_i \leq 0.5$ resp. $R_i > 0.5$). Die Anzahl der hinzuzufügenden oder wegzunehmenden Kuben wird durch die $K^{(i)}$ und $H_i^{(i)}$ determiniert resp. es wird versucht in jedem Schritt so viele Kuben hinzuzufügen oder wegzunehmen, wie es der Betrag der Differenz zwischen dem Höhenniveau und dem ausgewählten Eintrag zulässt ohne die Menge der plane partitions zu verlassen. Die $K_n^{(i)}$ determinieren dabei stets die maximale Anzahl der hinzufügbaren resp. wegnehmbaren Kuben an der ausgewählten Stelle. Um die gewünschte Übergangswahrscheinlichkeit zu erreichen, wird noch mit der Anzahl zu ändernder Kuben gewichtet.

Diese Dynamik präferiert offenbar die Transition zwischen plane partitions deren Distanz 1 ist.

Eine offensichtliche Vereinfachung ist die Beschränkung auf Sprünge zwischen den 1-Nachbarschaften, eine solche Modifikation kann erreicht werden, indem

²⁸Es ist klar, daß diese Definition nur dann Sinn macht, wenn man die oben betrachteten plane partitions als $(b+2) \times (c+2)$ -Matrizen auffasst, die an den Rändern die Randbedingungen enthalten. Diese Einträge werden nicht geändert und werden als Grundlage zur Implementierung der Dynamik dienen, siehe Kapitel 5.

geprüft wird, ob das Hinzufügen oder Wegnehmen an der zufällig ausgewählten Position (X_n, Y_n) noch eine plane partition ergibt, man benötigt dazu die R_i nicht mehr.

Beschränkt man sich auf die Ein-Kuben-Dynamik, so kann die Eigenschaft des schnellen Mischens (rapid mixing) ebenfalls gezeigt werden mit sehr viel schärferen, sogar optimalen Grenzen, dies verlangt aber nach feineren Methoden (siehe [45]).

Eine Untersuchung der Eigenschaft des schnellen Mischens wird für die hier vorgestellte Markovkette \mathcal{M}_{loz} im nächsten Abschnitt 4.4 durchgeführt.

Es kann nun eine Kopplung zweier Kopien von \mathcal{M}_{loz} hergestellt werden.

Sei dazu wieder $(\Theta_n)_n$ wie in der Definition von \mathcal{M}_{loz} eine Folge identischer, unabhängig verteilter Zufallsvariablen.

Dann sei der Prozess $(Z_n)_{n \geq 0}$ definiert durch die Kopplung zweier Kopien $(Z_n^{(1)})_{n \geq 0}, (Z_n^{(2)})_{n \geq 0}$ von \mathcal{M}_{loz} mit

$$Z_0 = (P_1, P_2), \quad P_1, P_2 \in \mathcal{P}_{b,c}^a$$

$$Z_{n+1} = (Z_{n+1}^{(1)}, Z_{n+1}^{(2)}) = (f(Z_n^{(1)}, \Theta_n), f(Z_n^{(2)}, \Theta_n))$$

Es bezeichne im Folgenden \mathcal{M}_{loz}^* diesen gekoppelten Prozess.

Dies stellt im Sinne der Definition von \mathcal{M}_{loz}^* eine gültige Kopplung mit Koaleszenzeigenschaft dar.²⁹

Insbesondere ist in diesem Fall die Kopplung keine Doeblin-Kopplung, d.h. die Vektorkomponenten sind voneinander *abhängig*, wegen der Benutzung der selben Θ_i für alle Vektorkomponenten.

Es sei an dieser Stelle bemerkt, daß die Koaleszenzeigenschaft in dieser Dynamik keine lokale Eigenschaft ist. Die Gleichheit von Einträgen in $Z_n^{(1)}$ und $Z_n^{(2)}$ bedeutet nicht zwangsläufig, daß dies in der weiteren Entwicklung so bleibt, sondern hängt sehr stark von der Nachbarschaft der Einträge ab. Erst wenn *alle* Einträge der beiden plane partitions übereinstimmen kommt es zur Koaleszenz. Die Dynamik wurde aus der (unvollständigen) Definition in [33] sinngemäß auf plane partitions übertragen.

4.2 Monotonie von \mathcal{M}_{loz}

Für den mCFTP-Algorithmus (siehe Abschnitt 3.4) ist die Definition einer partiellen Ordnungsrelation auf dem Zustandsraum wesentlich. Es sei daher für $P_1 = (p_1^{(i,j)})_{i,j}, P_2 = (p_2^{(i,j)})_{i,j} \in \mathcal{P}_{b,c}^a$

$$P_1 \preceq P_2 :\Leftrightarrow \forall (i, j) \in \{1, \dots, b\} \times \{1, \dots, c\} : p_1^{(i,j)} \leq p_2^{(i,j)}$$

Diese Relation ist lediglich eine partielle Ordnung, denn es sind zum Beispiel

²⁹Die Marginalprozesse sind jeweils wieder 1:1 Kopien von \mathcal{M}_{loz} .

$$P_1 = \begin{array}{|c|c|} \hline 2 & 1 \\ \hline 2 & 0 \\ \hline \end{array}$$

und

$$P_2 = \begin{array}{|c|c|} \hline 2 & 2 \\ \hline 1 & 0 \\ \hline \end{array}$$

nicht vergleichbar im Sinne der obigen Definition.

Es zeigt sich vielmehr, daß zwei *plane partitions* P_1, P_2 unvergleichbar sind, wenn

$$\|P_1\|_L = \|P_2\|_L$$

gilt. Jedoch folgt aus der Unvergleichbarkeit nicht, daß die Lozengenorm gleich ist.

Diese partielle Ordnung „ \preceq “ kann auf die deBruijn-Pfade 1:1 übertragen werden, Abbildung 19 zeigt die vollständige Anordnung der Elemente aus $\mathcal{P}_{2,2}^2$ in einem Hasse-Diagramm, zusammen mit den entsprechenden Lozengeteilungen und assoziierten deBruijn-Pfaden.

Die Existenz und Eindeutigkeit des maximalen und minimalen Elements liefert

Lemma 4.2.1 *Die partielle Ordnung „ \preceq “ auf $\mathcal{P}_{b,c}^a$ besitzt genau ein maximales und ein minimales Element, d.h. mit*

$$\perp_{b,c}^a := \underline{0} = \begin{array}{|c|c|c|c|} \hline 0 & 0 & \dots & 0 \\ \hline 0 & 0 & \dots & 0 \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline 0 & 0 & \dots & 0 \\ \hline \end{array} \quad \text{und} \quad \top_{b,c}^a := a \cdot \underline{1} = \begin{array}{|c|c|c|c|} \hline a & a & \dots & a \\ \hline a & a & \dots & a \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline a & a & \dots & a \\ \hline \end{array}$$

gilt

$$\forall P \in \mathcal{P}_{b,c}^a : \perp_{b,c}^a \preceq P \preceq \top_{b,c}^a$$

Beweis. Die Eigenschaften folgen direkt aus der Definition des minimalen und maximalen Elementes und den definierenden Eigenschaften einer *plane partition*. \square

Mit dieser partiellen Ordnung lässt sich nun zeigen

Satz 4.2.1 $\mathcal{M}_{loz}^* = (Z_n)_{n \geq 0}$ erhält „ \preceq “, d.h.

$$\pi_1(Z_N) \preceq \pi_2(Z_N) \Rightarrow \forall m \geq N : \pi_1(Z_{N+m}) \preceq \pi_2(Z_{N+m})$$

Beweis. Es genügt den Fall $m = 1$ zu betrachten. Dann folgt aber die Eigenschaft direkt aus der Definition der Updatefunktion der gekoppelten Dynamik. Insbesondere werden immer nur Kuben hinzugefügt oder nur Kuben weggenommen, dies wird durch die Konstruktion der Dynamik sichergestellt. \square

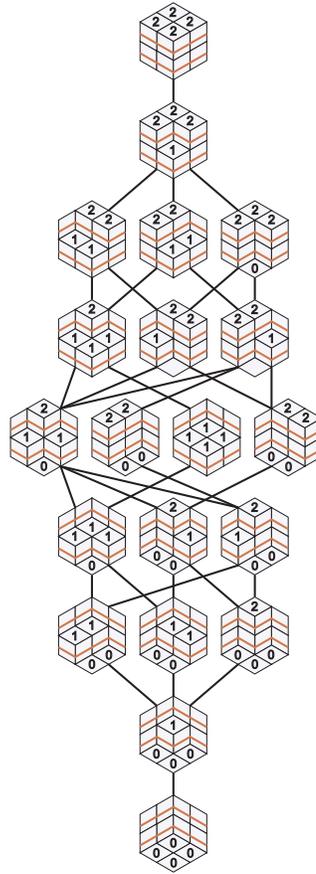


Abbildung 19: Vollständige Anordnung der Elemente aus $\mathcal{P}_{2,2}^2$ bezüglich „ \leq “

4.3 Konnektivität des Zustandsraums

Es sei Q die zu \mathcal{M}_{loz} assoziierte Übergangsmatrix. Um sicherzustellen, daß die Dynamik ins Gleichgewicht konvergiert, genügt die Kenntnis über die Ergodizität der Übergangsmatrix.

Diese Aussage bedeutet, daß Q sowohl aperiodisch als auch irreduzibel ist. Die Aperiodizität folgt aber aus der Definition, denn für jeden Zustand ist die Wahrscheinlichkeit im selben Zustand zu bleiben ungleich Null (die Diagonale von Q ist nirgends Null).

Es genügt also die Irreduzibilität von Q nachzuweisen. Diese lässt sich durch die Kommunikationsrelation zwischen Zuständen charakterisieren, d.h. alle Zustände der Markovkette sind durch einen Pfad mit positiver Wahrscheinlichkeit verbunden resp. es existiert genau eine Kommunikationsklasse.

Die Behauptung lässt sich mittels der Kodierung der Lozengeteilungen als plane

partitions zeigen und in folgendem Satz zusammenfassen:

Satz 4.3.1 \mathcal{M}_{loz} erreicht mit positiver Wahrscheinlichkeit jedes $P \in \mathcal{P}_{b,c}^a$.

Beweis. Zunächst werden die Elemente der plane partition umnummeriert, so dass $i \in \{1, \dots, bc\}$ nacheinander die Diagonalen durchläuft (siehe Abb. 20).

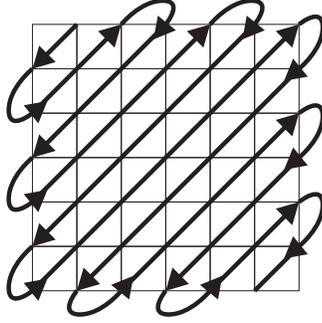


Abbildung 20: Durchlaufreihenfolge der Einträge einer plane partition aus $\mathcal{P}_{6,6}^a$, für a beliebig.

Mit dieser Umnummerierung kann nun jedes $P = (p^{(i,j)})_{ij} \in \mathcal{P}_{b,c}^a$ als Summe dargestellt werden:

$$P = \sum_{i=1}^{bc} p^{(i)} E^{(i)}.$$

Bezeichnet man

$$P^{(l)} = \sum_{i=1}^l p^{(i)} E^{(i)}$$

für $l \in \{1, \dots, bc\}$ und setzt

$$P^{(0)} := \perp_{b,c}^a,$$

so ist jedes $P^{(k)} \in \mathcal{P}_{b,c}^a$ und es gilt nach der Definition der partiellen Ordnung auf $\mathcal{P}_{b,c}^a$:

$$P^{(0)} \preceq P^{(1)} \preceq P^{(2)} \preceq \dots \preceq P^{(bc-1)} \preceq P^{(bc)} = P$$

und für beliebige $l \in \{0, \dots, bc\}$ ist

$$P^{(l)} \in V_{P^{(l+1)}, P^{(l+1)}},$$

sofern $P^{(l)}$ und $P^{(l+1)}$ nicht ohnehin identisch sind.

Der kürzeste Pfad von $\perp_{b,c}^a$ nach P kann als Abschätzung verwendet werden, es gilt

$$\mathbb{P}(Z_{nm} = P | Z_0 = \perp_{b,c}^a) \geq \prod_{i=1}^{bc} \mathbb{P}(Z_i = P^{(i)} | Z_{i-1} = P^{(i-1)}) > 0.$$

Der gewählte Pfad kann auch in umgekehrter Richtung durchlaufen werden, daher gilt auch

$$\mathbb{P}(Z_{nm} = \perp_{b,c}^a | Z_0 = P) > 0.$$

Das heißt zunächst, daß

$$\forall P \in \mathcal{P}_{n,m}^k : P \leftrightarrow \perp_{b,c}^a. \quad (6)$$

Seien nun $P_1, P_2 \in \mathcal{P}_{b,c}^a$, dann folgt aus (6)

$$P_1 \leftrightarrow \perp_{b,c}^a, \quad P_2 \leftrightarrow \perp_{b,c}^a.$$

Wegen der Transitivität und Symmetrie von „ \leftrightarrow “ folgt

$$P_1 \leftrightarrow P_2$$

und damit die Behauptung. \square

4.4 \mathcal{M}_{loz} ist schnell mischend

Es wurde bereits gezeigt, dass die Mischungszeit $\rho(\varepsilon)$ durch die erwartete Koaleszenzzeit T abgeschätzt werden kann (siehe Satz 3.2.1 in Kapitel 3.). Dies wird nun für die Abschätzung der Mischungszeit von \mathcal{M}_{loz} verwendet.

Die bereits definierte Kopplung \mathcal{M}_{loz}^* wird eine solche liefern. Dabei orientieren sich die Argumente an den Ausführungen in [33], wurden aber jeweils an die allgemeine Situation der Betrachtung von $\mathcal{H}_{a,b,c}$ und die Verwendung von plane partitions angepaßt.

Dort wird ein Submartingal $\zeta(n)$ konstruiert und eine verallgemeinerte Waldidentität benutzt, die jedoch die fast sichere Endlichkeit der Stopzeit (hier die Koaleszenzzeit) und beschränkte Zuwächse voraussetzt, um zu zeigen, daß die Dynamik im Mittel die Distanz zwischen den beiden Kopien verringert.

Es ist nicht leicht zu beweisen, daß der Erwartungswert der Koaleszenzzeit endlich ist. Um daher diese technischen Schwierigkeiten zu umgehen, wurde im Rahmen dieser Arbeit bemerkt, daß $-\zeta(n)$ ein nichtnegatives Supermartingal ist. Man kann daher den mit weniger starken Voraussetzungen behafteten Satz 7.6 aus [17] S. 274 benutzen, um die gleiche Aussage zu erhalten.

Außerdem hat sich ein kleiner Fehler in der Argumentation von Lemma 11 in [33] eingeschlichen. Das Problem tritt hier durch die in dieser Arbeit entwickelte Adaption der Argumente an plane partitions deutlicher hervor und konnte so umschifft werden.

Satz 4.4.1 *Es gilt für \mathcal{M}_{loz}*

$$\rho(\varepsilon) \leq 2eB^3 \ln \varepsilon^{-1},$$

wobei $B = abc$.

Der Beweis wird in mehreren Etappen geführt, die wichtigsten Zwischenresultate werden zunächst in den folgenden Lemmas gegeben.

Zuvor sei an

$$\mathcal{M}_{loz}^* = (Z_n)_{n \geq 0} = ((X_n, Y_n))_{n \geq 0}$$

erinnert mit den deterministischen Anfangsbedingungen $X_0 = x$ und $Y_0 = y$ mit $x \preceq y$.

Definiere noch

$$\Phi(n) := d_L(X_n, Y_n)$$

und

$$\Delta\Phi(n) = \Phi(n) - \Phi(n-1).$$

Lemma 4.4.1 *Es gilt*

$$\mathbb{E}(\Delta\Phi(n+1)|Z_n) \leq 0,$$

falls $\Phi(n) > 0$.

Beweis. Für $\Phi(n) = 0$ ist nichts zu zeigen, denn dann gilt

$$\forall m \in \mathbb{N}_0 : X_{n+m} = Y_{n+m} \Leftrightarrow \forall m \in \mathbb{N}_0 : d_L(X_{n+m}, Y_{n+m}) = 0$$

aufgrund der Koaleszenzeigenschaft von \mathcal{M}_{loz}^* .

Sei $\Phi(n) > 0$. Man kann nun $\Phi(n)$ als Volumen zwischen den partitions X_n und Y_n zur Zeit n auffassen und zeigen, daß die Volumenänderung im Mittel negativ ist und damit schrumpft.

Dazu benutzt man eine Art Prinzip von Cavalieri. Man zerlegt dazu X_n und Y_n in plane partitions auf $\mathcal{P}_{b,c}^1$ und betrachtet zunächst die „Scheiben“ ${}^k X_n$ und ${}^k Y_n$, d.h.

$$\forall A \in \mathcal{P}_{b,c}^1 : A = \sum_{k=1}^a {}^k A,$$

wobei

$${}^k A := (\mathbb{1}_{\{a^{i,j} \geq k\}})_{i,j}.$$

Aus der Definition folgt sofort

$${}^a A \preceq {}^{a-1} A \preceq \dots \preceq {}^1 A.$$

Betrachte nun im Folgenden für ein beliebiges, aber festes $k \in \{1, \dots, a\}$ die beiden Scheiben ${}^k X_n$ und ${}^k Y_n$.

Für jeden der Einträge in den Scheiben bezeichne man die Einträge mit „B“, falls das Hinzufügen oder Wegnehmen eines Kubus an dieser Stelle $d_L({}^k X_n, {}^k Y_n)$ vergrößert, mit „G“, falls die Distanz verringert wird.

Da nach Konstruktion von \mathcal{M}_{loz} jeder der Einträge in der plane partition mit gleicher Wahrscheinlichkeit $\frac{1}{bc}$ ausgewählt und mit Wahrscheinlichkeit $\frac{1}{2}$ verändert wird,³⁰ ergibt sich

$$\mathbb{E}(\Delta\Phi(n+1)|Z_n) = \frac{1}{2abc}(\#B - \#G).$$

³⁰Die einzigen erlaubten Transitionen sind nun zwischen Elementen der 1-Nachbarschaften, d.h. $a = \kappa = 1$, ebenso ist es noch erlaubt den aktuellen Zustand unverändert zu lassen.

Es bleibt daher nur noch zu zeigen, daß $\#G \geq \#B$ gilt.

Dazu betrachtet man die Differenz der Scheiben im Sinne der Matrixsubtraktion, so entstehen Polyominos³¹ aus 1-Einträgen. Trägt man nun in diese Differenzmatrix die mit B und G bezeichneten Punkte ein, so wird deutlich, daß es für jedes Polyomino genauso viele G -Einträge wie B -Einträge gibt, wobei die G -Einträge stets im Inneren der Polyominos liegen.

Ist das Polyomino ein Quadrat der Seitenlänge 1, so ist der innere Punkt mit $2G$ zu werten, da dann für beide möglichen Änderungen eine Differenzabnahme folgt.

Ein Beispiel für die Argumentation liefert Abbildung 21.

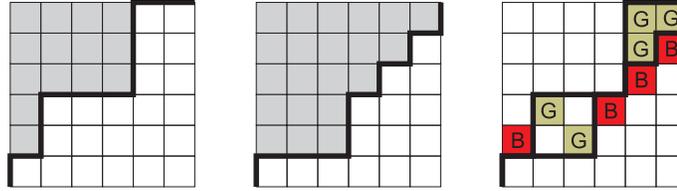


Abbildung 21: Beispiel für die Bezeichnung durch G - und B -Einträge für zwei vergleichbare plane partitions (links, mitte) aus $\mathcal{P}_{1,6,6}$. Es wird deutlich, daß die Anzahl G -Einträge nur deswegen größer ist, da eines der beiden Polyominos den Rand berührt (rechts).

Liegt eines der Polyominos am Rand so gibt es weniger B Einträge (mögliche B -Einträge lägen dann sozusagen außerhalb der plane partition). Daraus folgt insgesamt

$$\#G \geq \#B.$$

Das heißt also $\Delta\Phi(n)$ ist im Mittel nicht positiv auf Scheiben.

Setzt man die Scheiben wieder zur originalen plane partition zusammen, so erhält man das gleiche Ergebnis

$$\mathbb{E}(\Delta\Phi(n+1)|Z_n) = \frac{1}{2abc}(\#B - \#G) \leq 0,$$

da die Dynamik so definiert ist, daß für eine ausgewählte Position *alle* Scheiben geändert werden, in denen eine Transition möglich ist. So wird eine gegenseitige Behinderung der Scheiben vermieden und es folgt daher die Behauptung. \square

Bemerkung. Die Argumentation in [33], daß es für sich berührende Polyominos einen Nettoüberschuss von G -Einträgen pro Polyomino gäbe, ist nicht richtig. Es ist wie gesehen vielmehr so, daß der Überschuss an günstigen Punkten lediglich an den Rändern auftreten kann. $\Delta\Phi(n)$ ist letztlich allein wegen der Randbedingungen nicht im Mittel Null, sondern im Mittel negativ.

³¹In Analogie zu den Dominosteinen nennt man Polygone, die aus Quadraten zusammengesetzt sind Polyominos.

Lemma 4.4.2 *Es gilt*

$$\mathbb{E}((\Delta\Phi(n+1))^2|Z_n) \geq \frac{1}{2B} > 0$$

Beweis. Für jede Transition der Dynamik auf den im vorigen Lemma definierten Scheiben ist die Transitionswahrscheinlichkeit nach Definition der Dynamik gerade $\frac{1}{2bc}$. Daher ist

$$\mathbb{E}((\Delta\Phi(n+1))^2|Z_n) \geq \frac{1}{2bc} \geq \frac{1}{2B} > 0.$$

Dies liefert wegen der Definition der Dynamik in analoger Weise wie im vorigen Lemma die Behauptung. \square

Satz 4.4.2 *Für die Koaleszenzzeit von \mathcal{M}_{loz}^* gilt nun*

$$\mathbb{E}(\tau^{x,y}) \leq \frac{\Phi(0)(2B - \Phi(0))}{V},$$

bei beliebigen deterministischen Startbedingungen $X_0 = x$ und $Y_0 = y$ mit $x \preceq y$ und $B := abc$, wobei $V := \frac{1}{2B} \geq 0$.

Beweis. Definiere zunächst den stochastischen Prozess

$$\zeta(n) := \Phi(n)^2 - 2B\Phi(n) - Vn.$$

Es wird zunächst gezeigt, daß $(\zeta(n))_{n \geq 1}$ ein $(\sigma(Z_n))_{n \geq 1}$ adaptiertes Submartingal ist.

Es gilt:

$$\zeta(n) = \Phi(n)^2 - 2B\Phi(n) - Vn = \underbrace{(\Phi(n) - B)^2}_{\geq 0} - B^2 - Vn$$

Daraus folgt nun weiter wegen $B^2 + Vn \geq 0$

$$-(B^2 + Vn) \leq \zeta(n) \leq \underbrace{\Phi(n)^2}_{\in \{0, \dots, B^2\}} \leq B^2 \leq B^2 + Vn$$

$$\Rightarrow |\zeta(n)| \leq B^2 + Vn$$

$$\Rightarrow \mathbb{E}(|\zeta(n)|) \leq B^2 + Vn$$

$\zeta(n)$ ist also für alle $n \in \mathbb{N}$ integrierbar; die Adaptiertheit zu $(Z_n)_{n \geq 1}$ ist klar.

Weiter gilt:

$$\begin{aligned} \zeta(n+1) - \zeta(n) &= \Phi^2(n+1) - 2B\Phi(n+1) - V(n+1) - \Phi^2(n) + 2B\Phi(n) + Vn \\ &= \Phi^2(n+1) - \Phi^2(n) - 2B\Delta\Phi(n+1) - V \\ &= (\Delta\Phi(n+1))^2 - 2\Phi^2(n) + 2\Phi(n)\Phi(n+1) - 2B\Delta\Phi(n+1) - V \\ &= (\Delta\Phi(n+1))^2 - V + 2\Delta\Phi(n+1)(\Phi(n) - B) \end{aligned}$$

Nimmt man auf beiden Seiten den auf Z_n bedingten Erwartungswert, so erhält man

$$\begin{aligned} & \mathbb{E}(\zeta(n+1) - \zeta(n) | Z_n) \\ &= \underbrace{\mathbb{E}((\Delta\Phi(n+1))^2 | Z_n) - V}_{\geq 0, \text{ nach Lemma 4.4.2}} + 2 \underbrace{(\Phi(n) - B)}_{\leq 0, \text{ da } \Phi(n) \in \{0, \dots, B\}} \underbrace{\mathbb{E}(\Delta\Phi(n+1) | Z_n)}_{\leq 0, \text{ nach Lemma 4.4.1}} \geq 0, \end{aligned}$$

und damit, daß

$$\mathbb{E}(\zeta(n+1) | Z_n) \geq \zeta(n).$$

Insgesamt ist daher $\zeta(n)$ ein Submartingal bezüglich der von $(Z_n)_{n \geq 1}$ erzeugten σ -Algebra.

Eine nähere Analyse von $\zeta(n)$ ergibt noch wegen $\Phi(n) \in \{0, \dots, B\}$

$$\zeta(n) \leq 0$$

für alle $n \in \mathbb{N}$.

Das heißt nichts anderes, als daß $-\zeta(n)$ ein nichtnegatives Supermartingal ist. Da die Koaleszenzzeit $\tau^{x,y}$ eine Stopzeit ist, gilt nach [17] S. 274 Theorem 7.6, Kapitel 4

$$\mathbb{E}(-\zeta(0)) \geq \mathbb{E}(-\zeta(\tau^{x,y})) \Leftrightarrow \mathbb{E}(\zeta(0)) \leq \mathbb{E}(\zeta(\tau^{x,y})).$$

Daraus folgt wegen

$$\mathbb{E}(\zeta(0)) = \Phi(0)^2 - 2B\Phi(0) - V \cdot 0 = \Phi(0)(\Phi(0) - 2B)$$

und

$$\begin{aligned} \mathbb{E}(\zeta(\tau^{x,y})) &= \mathbb{E}(\underbrace{\Phi(\tau^{x,y})^2}_{=0} - 2B \underbrace{\Phi(\tau^{x,y})}_{=0} - V\tau^{x,y}) = \mathbb{E}(0 - 0 - V\tau^{x,y}) \\ &= -V\mathbb{E}(\tau^{x,y}) \end{aligned}$$

die Behauptung. \square

Wegen der partiellen Ordnung auf E ist

$$T = \max_{x,y} \mathbb{E}(\tau^{x,y}) = \mathbb{E}(\tau_{b,c}^{\perp, \top})$$

und daher für $X_0 = \perp$ und $Y_0 = \top$

$$\Phi(0) = d_L(\perp_{b,c}^{\perp}, \top_{b,c}^{\perp}) = abc = B.$$

Lemma 4.4.2 liefert daher zusammen mit den vorangegangenen Lemmas

$$T \leq \frac{B(2B - B)}{V} = \frac{B^2}{V} = 2B^3$$

Damit liefert die Abschätzung der Mischungszeit in Satz 3.2.1 die Behauptung des Satzes.

4.5 Verteilung der Koaleszenzzeit

Für die Abschätzung der Mischungszeit wurde im vorigen Abschnitt die Koaleszenzzeit von \mathcal{M}_{loz}^* herangezogen, eine natürliche Frage ist die nach ihrer Verteilung. Im Folgenden sollen kurz einige Ergebnisse der Simulation bzgl. der Koaleszenzzeiten gezeigt werden.

Mit Hilfe eines kleinen C++-Programms, das für den freien MatLab-Klon *Octave* geschrieben wurde (siehe Kapitel 5), wurden jeweils 1000 Stichproben der Koaleszenzzeit für Teilungen der Grösse $n = 2, \dots, 5$ erzeugt.

Die mit dem Programm erzeugten Daten wurden zunächst zu einem Histogramm verarbeitet und skaliert, so dass die Summe über die Einträge auf 1 normiert ist. Dies liefert eine empirische Dichte. Derartige Plots (siehe Abbildung 22) legen hier nun die Gumbelverteilung für die Koaleszenzzeiten nahe.

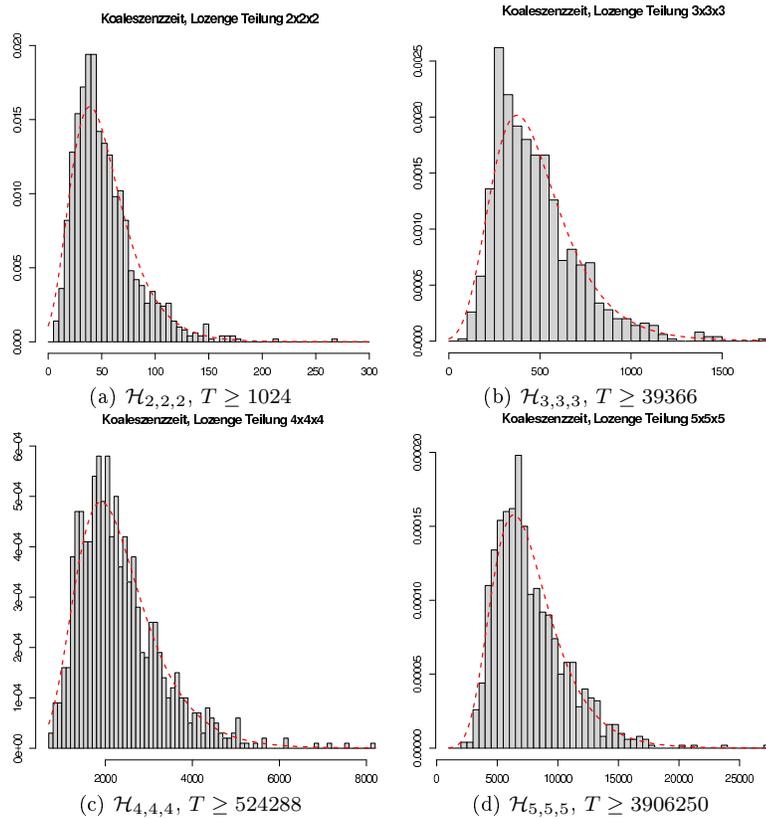


Abbildung 22: Plot der empirischen Dichten der Koaleszenzzeiten und Fit an die Dichte einer Gumbelverteilung, sowie die Schätzung von T aus Abschnitt 4.4.

Die Abbildungen zeigen auch noch einmal ganz deutlich, daß die in Abschnitt 4.4 gegebene Abschätzung der Koaleszenzzeit nicht optimal ist.

5 Programmlisting und Diskussion der Implementierung

Die in Kapitel 4 eingeführte Dynamik eignet sich in der plane partition Darstellung zur Simulation durch einen Computer. Um diese zu realisieren, wird folgender Weg eingeschlagen.

Zunächst werden zwei $(b+2) \times (c+2)$ -Matrizen erzeugt und mit den entsprechenden Randbedingungen initialisiert, siehe Abbildung 23.

a	a	a	a	a	a	a	a	a	a
a	0	0	0	0	a	a	a	a	0
a	0	0	0	0	a	a	a	a	0
a	0	0	0	0	a	a	a	a	0
a	0	0	0	0	a	0	0	0	0

Abbildung 23: Initiale Matrizen für die Erzeugung von Stichproben aus $\mathcal{P}_{3,3}^a$. Die Randbedingungen sind blau unterlegt und fix während der ganzen Simulation. Links Initialisierung für den Start der Simulation in der minimalen Teilung \perp , rechts in der maximalen \top .

Nun kann die Simulation mit Hilfe des mCFTP gestartet werden.

Zur Simulation eines Schrittes der Dynamik $\mathcal{M}_{loz}^* = (X_n, Y_n)$ wird die Definition der Dynamik in Abschnitt 4 benutzt.

Zur Feststellung der Koaleszenz wäre eine naive Herangehensweise der Vergleich aller Matrixeinträge pro Zeiteinheit. Dafür müssten in jedem Schritt $b \cdot c$ Einträge verglichen werden.

Aufgrund der Definition von \mathcal{M}_{loz}^* und dessen Monotonieeigenschaften genügt die Protokollierung der Lozengenormen der beiden Teilungen X_n und Y_n im Zeitverlauf.

Die Dynamik ist koalesziert, wenn

$$\|X_n\|_L = \|Y_n\|_L,$$

es genügt also in jedem Schritt die Lozengenormen zu vergleichen (diese werden im Verlauf der Dynamik laufend aktualisiert, entsprechend der Stapelhöhe der wegzunehmenden oder hinzuzufügenden Kuben).

Das Programm kann nach Start von Octave durch

```
> [T,lnorm,Teilung]=lozengesim(a,b,c);
```

gestartet werden.³² T beinhaltet dann die Kopplungszeit, $lnorm$ liefert die Lozengenorm zum Zeitpunkt der Koaleszenz und $Teilung$ liefert eine gleichverteilte Stichprobe aus $\mathcal{P}_{b,c}^a$.

Für die Erfassung der Verteilung der Kopplungszeiten kann

```
> AnzahlStichproben=n;
> for i=1:AnzahlStichproben
    [T]=lozengesim(a,b,c);
    Zeiten(i)=T;
end
```

benutzt werden.

Zeitgleich mit der Programmierung der C++-Routine ist noch eine Visualisierung der Dynamik vorgenommen worden. Dazu wurde das Open-Source-Projekt *Processing* (siehe www.processing.org) verwendet, das es erlaubt mit einfachen Befehlen Visualisierungen unter Verwendung von einfachem JAVA-Code zu verwirklichen.

Da die Implementierung allein zur Visualisierung dient, wurde hier (nicht jedoch beim C++-Programm) auf Implementation der monotonen Kopplung aus der Vergangenheit verzichtet, die Dynamik läuft also vorwärts, startend in der minimalen Verteilung. Ebenso wurde der Algorithmus etwas vereinfacht, es wird in jedem Schritt nur maximal ein Einheitskubus hinzugefügt oder weggenommen. Es ist klar, daß ein Schritt in der in dieser Arbeit vorgestellten Dynamik durch entsprechend viele Schritte auch mit Ein-Kuben-Schritten verwirklicht werden kann.

Dieses Vorgehen erlaubt es, die Entwicklung der Dynamik direkt zu verfolgen. Zudem lässt sich der aktuelle Entwicklungsstand einfach als Bilddatei oder Vektorgraphik exportieren.

Das Programm kann nach Start verschiedene Darstellungsmöglichkeiten aufzeigen, hier eine kurze Liste der Tastenkombinationen zur Steuerung der Darstellung:

```
'a'  arctic circle als gelbe Kurve anzeigen
't'  Einfärben der Teilung zur Visualisierung des YOUNG-Diagramms
'i'  Neustart
'b'  neue zufällige Teilung mit Lozengenorm nahe  $\frac{abc}{2}$  erzeugen
'd'  Umschalten in die Darstellung als System von deBruijn-Pfaden
'v'  Einblenden einer farbkodierten Version der plane partition
'o'  Ausgabe des aktuellen Frames als PDF
'p'  Ausgabe des aktuellen Frames als PNG
```

Sowohl Quellcode als auch das Programm in verschiedenen Versionen für die gängigsten Betriebssysteme liegen der Arbeit auf CD bei. Der Quellcode kann

³²Das Programm muss vor der Benutzung mit

mkocfile lozengesim.cc

kompiliert werden, evtl. sind entsprechende Bibliotheken nachzuladen.

nach Belieben verändert werden, um Teilungen verschiedener Größe zu erzeugen. Alle Abbildungen, im Speziellen die Abbildungen des nächsten Kapitels, wurden mit diesem Programm erzeugt.

Nachfolgend das Listing des Codes der C++-Routine zur dynamischen Verlinkung in Octave

```
#include <time.h>
#include <oct.h>

#include <octave/config.h>
#include <iostream>
#include <octave/defun-dld.h>
#include <octave/error.h>
#include <octave/oct-rand.h>

#include <octave/pager.h>
#include <octave/symtab.h>
#include <octave/variables.h>
#include <octave/Array2.h>

#include <vector>

int min(int a, int b) {
    return (a<b)?a:b;
}

int max(int a, int b) {
    return (a>b)?a:b;
}

// Methode zur Erzeugung der minimalen Teilung mit Randbedingungen
NDArray getMinimalTiling(int a, int b, int c) {
    dim_vector dv(b+2,c+2);

    NDArray out = NDArray(dv);

    for(int i=0; i<b+2; i++) {
        for(int j=0; j<c+2; j++) {
            if((i==0)||j==0){
                out(i,j)=a;
            }
            else {
                out(i,j)=0.0;
            }
        }
    }
}
```

```

    }

    return out;
}

// Methode zur Erzeugung der minimalen Teilung mit Randbedingungen
NDArray getMaximalTiling(int a, int b, int c) {
    dim_vector dv(b+2,c+2);
    NDArray out=NDArray(dv);

    for(int i=0; i<b+2; i++) {
        for(int j=0; j<c+2; j++) {
            if((i==b+1)||(j==c+1)){
                out(i,j)=0;
            }
            else {
                out(i,j)=a;
            }
        }
    }

    return out;
}

// Hauptprogramm
// Definiert eine dynamisch linkbare Funktion, die in
// Octave benutzt werden kann
//
// Rueckgabewerte sind die Kopplungszeit und die erzeugte
// gleichverteilte Stichprobe
//
DEFUN_DLD (
    lozengesim ,
    args , ,
    "simuliert das Kopplungsverhalten von zwei zufällig orientierten
    (hexagonalen) Kristallen"
) {

    // Rueckgabewerte fuer Octave definieren
    octave_value_list l;

    // Pruefen, ob genug Parameter angegeben
    if(args.length()!=3) {
        print_usage("lozengesim");
        return l;
    }
}

```

```

int a = args(0).int_value();
int b = args(1).int_value();
int c = args(2).int_value();

// Bereitstellen der benoetigten Matrizen, xn und yn sind die
// zu koppelnden Kopien von M_loz
//
// xn startet in der minimalen Teilung, yn in der maximalen
dim_vector dv(b+2,c+2);

NDArray xn = NDArray(dv);
NDArray yn = NDArray(dv);

yn=getMaximalTiling(a,b,c);
xn=getMinimalTiling(a,b,c);

// Es gilt hier  $|X_0|_L=0$ ,  $|Y_0|=abc$ 
int lnormyn = a*b*c;
int lnormxn = 0;

// Vorbereiten eines Vektors zufaelliger Zahlen
dim_vector dv1(4,1);

NDArray Theta = NDArray(dv);
Theta = octave_rand::matrix(4,1);

// Speicherplaetze fuer die Komponenten von Theta_n
int thetax = 0;
int thetay = 0;
int level = 0;
float coin = 0;

// Variable fuer die Stapelhoehen
int h=0;
int h1=0;

// erster Kopplungsversuch bei  $T=-1$ 
int T=1;

// Initialisieren der Theta, vector-Container als dynamisches
// Speichermedium
std::vector<int> posx;
std::vector<int> posy;
std::vector<int> heights;
std::vector<float> r;

```

```

posx.push_back((int)(Theta(0,0)*b+1));
posy.push_back((int)(Theta(1,0)*c+1));
heights.push_back((int)(Theta(2,0)*a+1));
r.push_back(Theta(3,0));

// solange T erhoehen bis xn=yn resp. |xn|_L=|yn|_L
while(lnormyn!=lnormxn) {

// start an -T
for(int i=0; i<T; i++) {

// auslesen der schon vorhandenen Zufallszahlen
coin=r[T-i-1];
level=heights[T-i-1];
thetax=posx[T-i-1];
thetay=posy[T-i-1];

// Implementation der Dynamik
if(coin <= 0.5) {
h1=max(level-xn(thetax , thetay), 0);
h=min(
xn(thetax-1, thetay)-xn(thetax , thetay),
xn(thetax , thetay-1)-xn(thetax , thetay)
);

if(h-h1>=0) {
if(coin <= 1.0/(2.0*h1)) {
xn(thetax , thetay)+=h1;
lnormxn+=h1;
}
}

h1=max(level-yn(thetax , thetay), 0);
h=min(
yn(thetax-1, thetay)-yn(thetax , thetay),
yn(thetax , thetay-1)-yn(thetax , thetay)
);

if(h-h1>=0) {
if(coin <= 1.0/(2.0*h1)) {
yn(thetax , thetay)+=h1;
lnormyn+=h1;
}
}
}

```

```

    }
}

else {

    h1=max(xn(thetax , thetay)-level +1 ,0);
    h=min(
        xn(thetax , thetay)-xn(thetax +1, thetay) ,
        xn(thetax , thetay)-xn(thetax , thetay +1)
    );

    if(h-h1>=0) {
        if(coin >=1.0-1.0/(2.0*h1)) {
            xn(thetax , thetay)-=h1;
            lnormxn-=h1;
        }
    }

    h1=max(yn(thetax , thetay)-level +1 ,0);
    h=min(
        yn(thetax , thetay)-yn(thetax +1, thetay) ,
        yn(thetax , thetay)-yn(thetax , thetay +1)
    );

    if(h-h1>=0) {
        if(coin >=1.0-1.0/(2.0*h1)) {
            yn(thetax , thetay)-=h1;
            lnormyn-=h1;
        }
    }

}
}

```

```

// einen Schritt weiter zurueck in die Vergangenheit
T=T+1;

```

```

// neues Theta an die bestehende Liste anhaengen
Theta = octave_rand::matrix(4,1);

```

```

posx.push_back((int)(Theta(0,0)*b+1));
posy.push_back((int)(Theta(1,0)*c+1));
heights.push_back((int)(Theta(2,0)*a+1));
r.push_back(Theta(3,0));

```

```

// nur Reinitialisierern , falls noch nicht koaleszent

```

```

if(lnormxn!=lnormyn) {
  lnormxn=0;
  lnormyn=a*b*c;
  xn=getMinimalTiling(a,b,c);
  yn=getMaximalTiling(a,b,c);

}
}

// Koaleszenzzeit ausgeben
octave_value t=T-1;
l.append(t);

// Lozengenorm des Samples ausgeben
t=lnormxn;
l.append(t);

// Sample ausgeben
l.append(xn);

return l;
}

```

6 Ausblick

6.1 Arctic-Circle-Phänomen

Betrachtet man die Dynamik mit dem beigelegten Processing-Programm in seiner Entwicklung, so stellt man nach einer gewissen Wartezeit fest, daß sich im Zentrum eine fast kreisförmige Region entwickelt, in der die Lozenges relativ wahllos angeordnet sind, während sich am Rand größere Regionen bilden, die nur aus einer Art von Teilen besteht resp. in denen die Teile alle gleiche Orientierung haben. Zeichnet man eine Kurve entlang dieser „eingefrorenen“ Regionen, so ist die eingezeichnete Kurve einem Kreis sehr ähnlich.³³ Die so entstandene Kurve nennt man *arctic circle*, siehe auch Abbildung 24.

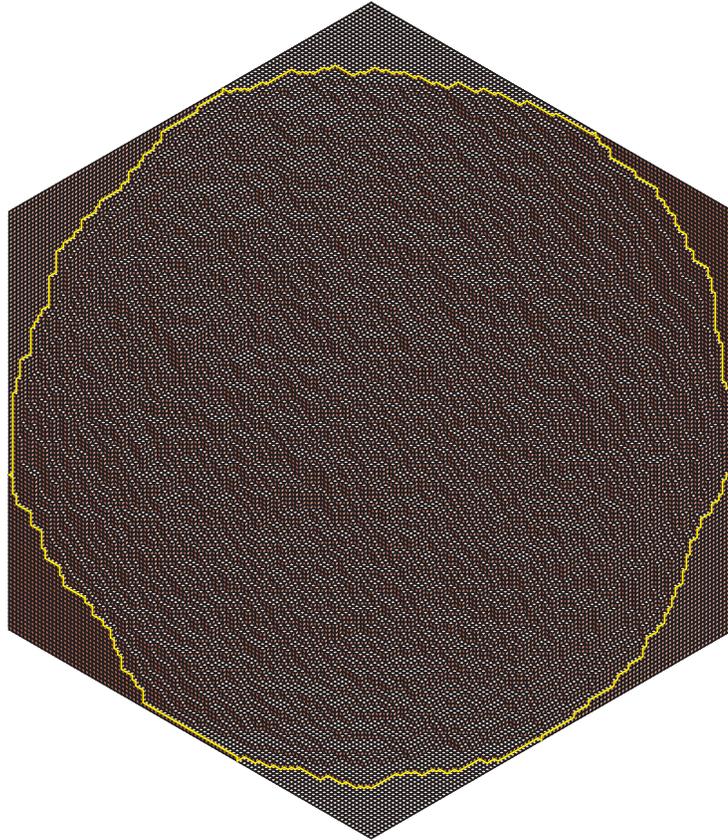


Abbildung 24: Teilung von $\mathcal{H}_{100,100,100}$ mit eingezeichnetem arktischen Kreis (gelbe Linie)

³³Eine solche Kurve kann mittels des beiliegenden Programmes durch das Drücken der Taste 'a' ein- und ausgeblendet werden.

Dieses Phänomen ist zuerst bei Aztech-Domino-Teilungen³⁴ beobachtet worden. James Propp und David B. Wilson haben zahlreiche Untersuchungen dazu angestellt und die Behauptung aufgestellt, daß der *arctic circle* tatsächlich gegen einen die Seitenränder der Region berührenden Kreis konvergieren müsse, wenn man die Anzahl der Teile erhöht und gegen unendlich streben lässt, dabei aber den äußeren Rand der Region fixiert. Man nennt diese Art der Grenzbildung im Englischen häufig *scaling limit* oder im Deutschen Skalierungslimit. Man findet weitere Ausführungen dazu und einen Beweis in [8].

Indem die Lozenzteilungen in Analogie zu den Treppenfunktionen als Treppenförmigkeiten angesehen werden, kann eine Lozenzeteilung als diskrete Näherung an eine glatte Fläche, die durch die Randbedingungen bestimmt wird, interpretiert werden. Dies geschieht natürlich unter der Vorstellung, daß jeder Lozenzeteilung ein dreidimensionales YOUNG-Diagramm zugeordnet werden kann, siehe Abschnitt 4.1, Definition 4.1.2.

Diese sind als eine mögliche Erweiterung der Irrfahrt als zweidimensionale Fläche zu sehen, ähnlich der Erweiterung der Brownschen Bewegung auf Brownsche Blätter.

Betrachtet man nun eine Folge C_n von Randkurven (Jordan C^1 Kurven), die eine teilbare Region umspannen und konvergiert $\frac{C_n}{n}$ gegen eine Kurve $C \subset \mathbb{R}^3$, dann spannt die Grenzrandkurve C eine eindeutig bestimmte Lipschitzfläche auf.

Nimmt man die zu C_n gehörenden Gleichverteilungen μ_n auf der Menge der Teilungen von C_n und betrachtet

$$\lim_{n \rightarrow \infty} \frac{\mu_n}{n},$$

im Sinne der schwachen Konvergenz, so kann eine Art Gesetz der Grossen Zahlen formuliert werden (siehe etwa [7]). So erhält man die Aussage, daß der Grenzwert nur von der Randkurve abhängt und ein Dirac-Maß auf der einzigen Lipschitzfläche ist, die von C aufspannt wird. Diese Grenzfläche wird *limit shape* oder Grenzform genannt und kann als makroskopische Form einer Grenzfläche, wie beispielsweise einer Kristallfläche, angesehen werden. Die vorangegangenen Ausführungen basieren auf [28] Abschnitt 2, Einleitung.

Die Frage nach der quantitativen und qualitativen Beschreibung der „arktischen Kurven“ wird in der selben Referenz beantwortet.

Es sind auch andere Randbedingungen denkbar, siehe etwa Abbildung 25.

Insbesondere kann man auch Teilungen mit Löchern untersuchen oder analog mit fixierten Einträgen, siehe als Beispiel Abbildung 26. Die dabei entstehenden Grenzkurven in der Interpretation wie oben nennen Kenyon und Okunov *cloud curves* oder übersetzt *Wölkchenkurven* (siehe wieder [28]).

Weiter bemerkt man, daß die hexagonale Begrenzung der Teilung starken Einfluß auf die Gestalt nimmt. Man kann sich nun fragen, ob es auch Randbedingungen gibt, die keinen Einfluß auf die innere Struktur nehmen. Diese Frage

³⁴Es handelt sich dabei um Dominoteilungen einer Region, die wie ein auf die Spitze gestelltes Quadrat aussieht. Die Dominos sind in Analogie zu den Lozenges aus zwei Quadraten zusammengesetzt und parkettieren das euklidische Gitter resp. die Ebene.

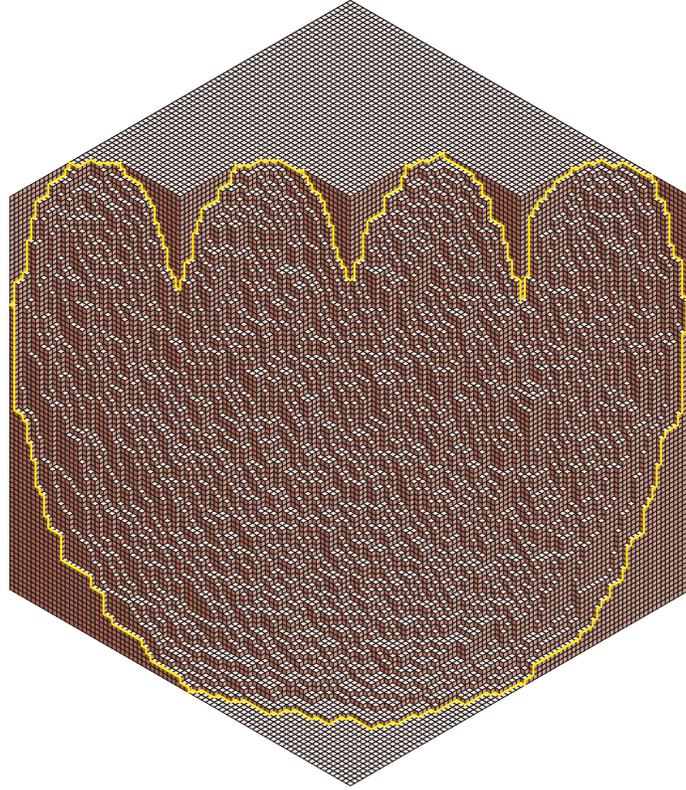


Abbildung 25: Eine Art arktische Tulpenkurve entsteht, wenn eine Dreiecksblockmatrix fixiert wird. Die Teilung ist aus $\mathcal{H}_{80,80,80}$.

kann bejaht werden. Die in Abbildung 27 gezeigten Randbedingungen haben keinen Einfluß auf die innere Strukturierung der Teilung.

Diese spezielle Eigenschaft kann man zum Beispiel nutzen, um die statistische Eigenschaften einer die ganze Ebene ausfüllenden zufälligen Teilung abzuschätzen.

6.2 Oktagonale Teilungen

Die in dieser Arbeit vorgestellten Lozengeteilungen gehören zu einer weitaus größeren Klasse von Teilungen, den Rhombusteilungen.

Man verallgemeinert zunächst die Situation, indem man m paarweise linear unabhängige Vektoren v_1, \dots, v_m aus \mathbb{R}^2 auswählt.³⁵

³⁵Im allgemeineren Fall können auch Vektoren aus \mathbb{R}^n ausgewählt werden, dies führt zu den Rhomboederteilungen.

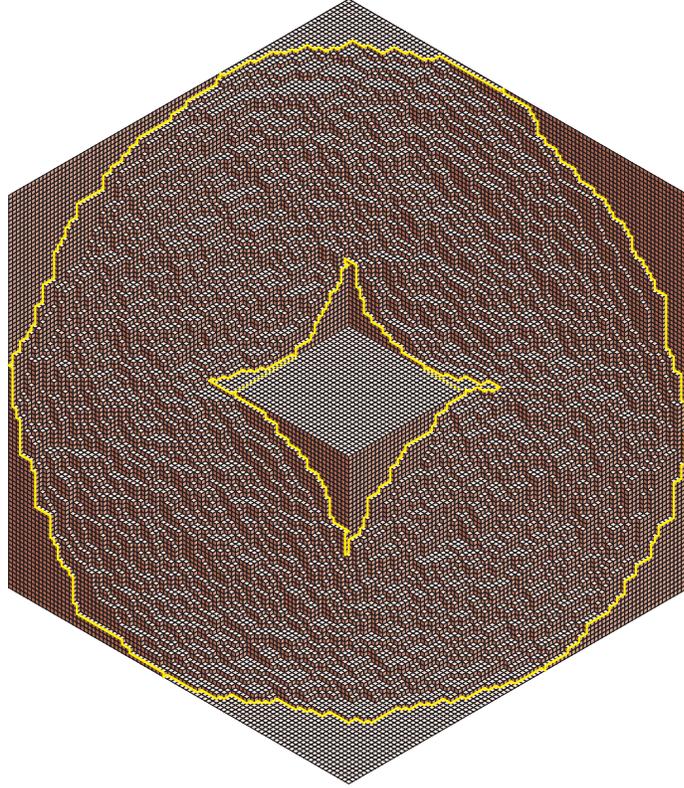


Abbildung 26: Die Kurve die bei fixierten mittleren Einträgen entsteht ist nicht mehr zusammenhängend. Simuliert wurde eine Teilung von $\mathcal{H}_{90,90,90}$ mit auf den Wert 45 fixierten Einträgen in der Mitte.

Es sei nun für $n \in \mathbb{N}$ die Menge

$$Z_m^n := \left\{ \sum_{i=1}^m \lambda_i v_i : 0 \leq \lambda_i \leq n \right\}.$$

Die konvexe Hülle von Z_m^n liefert die zu teilende Region. Weiterhin sei ein Pflasterstein der aus der Kombination zweier verschiedener Vektoren entstehende Rhombus. Es gibt offenbar insgesamt $\binom{m}{2}$ verschiedene Pflastersteine. Wählt man speziell $m = 3$ und die Vektoren

$$v_k = e^{t \frac{k\pi}{6}}, k = 0, 1, 2$$

wird die konvexe Hülle von Z_3^n zum bekannten Hexagon der Seitenlängen n und als Pflastersteine ergeben sich die drei verschieden orientierten Lozenges.³⁶

³⁶Es ist Zufall, daß hier die Pflastersteine durch Rotation auseinander hervorgehen. Im Allgemeinen sind die entstehenden Rhomben nicht kongruent.

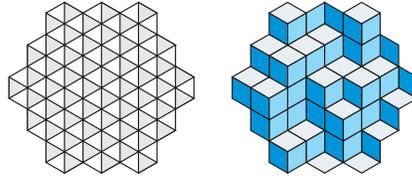


Abbildung 27: Hexagonale Region mit „spannungsfreien Rändern“ und Beispieltteilung

Für den Fall $m = 4$ und die Wahl $v_i = e^{\frac{ik\pi}{4}}$, $k = 0, 1, 2, 3$ ist Z ein regelmäßiges Oktagon mit Seitenlängen n und es gibt insgesamt 6 Pflastersteine (wovon nur zwei nicht kongruent sind). Diese werden zusammen mit einer exemplarischen Teilung in Abbildung 28 dargestellt. Abbildung 29 zeigt eine Prozession von exemplarischen Teilungen für verschiedene n für die die Vektoren so gewählt wurden, daß regelmäßige Polygone entstehen.

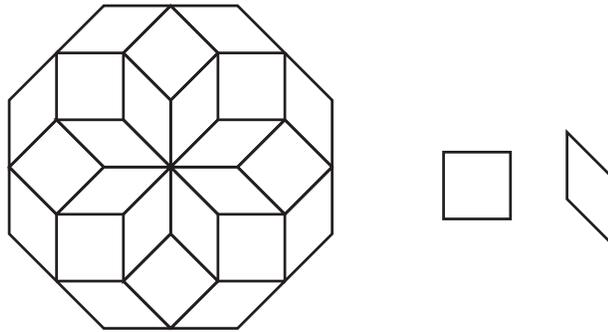


Abbildung 28: Oktagonteilung und die beiden einzigen nicht-kongruenten Pflastersteine der Teilung.

Im weiteren werden nur die sogenannten oktagonalen Teilungen betrachtet, schon hier wird deutlich werden, mit welchen Schwierigkeiten man beim Übergang zu allgemeineren Teilungen konfrontiert ist.

Es wurde schon in Abschnitt 2 bemerkt, daß für die Lozengeteilungen eine deBruijn-Faser ausreicht um die Menge der Teilungen eineindeutig zu beschreiben. Daher konnte man sich auf die Simulation der deBruijn-Pfade bzw. die Simulation von n sich gegenseitig meidenden Pfaden im euklidischen Gitter beschränken.

Im allgemeinen ist dies nicht mehr der Fall, für die oktagonalen Teilungen benötigt man nunmehr mindestens *zwei* nichtidentische Fasern um die Menge der Teilungen eineindeutig zu beschreiben. Von diesen gibt es nun insgesamt vier verschiedene (siehe Abbildung 30).

Im Fall der Lozengeteilungen waren die deBruijn-Pfade mit Pfaden im euklidi-

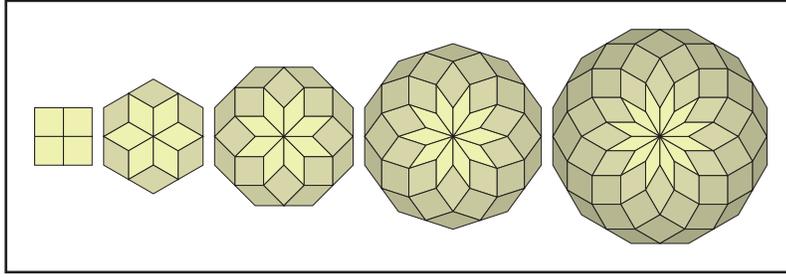


Abbildung 29: Verschiedene Rhombusteilungen für $2n$ -Ecke, $n = 2..5$.

schen Gitter zu identifizieren; die hier auftretenden Pfade sind unter dem Namen *Motzkin-Pfade* bekannt. Sie bestehen hier aus jeweils gleich vielen Schritten in $(1, 1)$, $(1, -1)$ und $(1, 0)$ -Richtung.

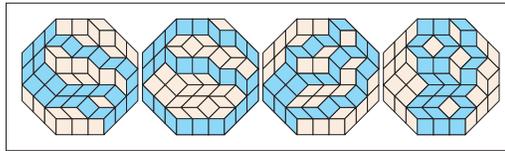


Abbildung 30: Die vier de Bruijn-Familien für eine oktagonale Rhombusteilung.

Ähnlich wie bei den Lozengeteilungen möchte man nun eine Dynamik definieren, die nur lokale Änderungen durchführt. Die Idee dabei ist, drei zufällig ausgewählte Pflastersteine, die zusammen in einem (irregulären) Hexagon liegen um 180 Grad zu drehen; dieses Drehen nennt man auch *phason-flip*. Im Fall der Lozengeteilungen wird dies (nach isometrischer Projektion) durch das Hinzufügen oder Wegnehmen eines Einheitskubus in der plane-partition-Interpretation realisiert.

Diese Vorgehensweise funktioniert im oktagonalen Fall nicht mehr, da man zwar die oktagonalen Teilungen als Projektion einer dreidimensionalen partition, einer sogenannten hyperplane partition, auffassen könnte, es aber durch die Projektion vom \mathbb{R}^4 auf den \mathbb{R}^2 zu unerwünschten Überlappungen kommt oder genauer die Identifikation der oktagonalen Teilungen mit der isometrischen Projektion von hyper plane partitions nicht mehr eindeutig ist.

Aus diesem Grunde muss ein anderer Weg eingeschlagen werden. Destainville schlägt in [13] ein einfaches Verfahren vor. Es wird zunächst eine hexagonale Teilung erzeugt. Auf dieser können dann n Random Walks definiert werden, die sich zwar berühren, jedoch nicht überkreuzen dürfen und die jeweils gemeinsam in einer Ecke der zu teilenden Region beginnen und enden und die Kanten der Lozenges als Pfade wählen können.

Entlang dieser erzeugten Random Walks, resp. ihrer Trajektorien, kann man

die Hexagonteilung sozusagen aufbrechen und eine weitere deBruijn-Familie in Abhängigkeit der erzeugten Trajektorien einfügen. Nach einer passenden Skalierung erhält man dann die gewünschten Pflastersteine und die entsprechende oktagonale Teilung (zur Verdeutlichung siehe Abbildung 31).

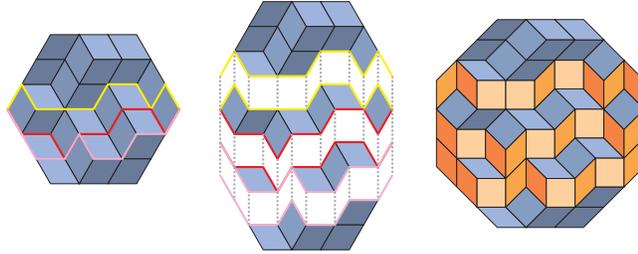


Abbildung 31: Das Verfahren zur Erzeugung einer oktagonalen Teilung aus einer Lozengeteilung, verdeutlicht für den Fall $n = 3$. Die verschiedenfarbigen Pfade sind die erwähnten Random Walks auf einer zuvor erzeugten Teilung von $\mathcal{H}_{3,3,3}$. Orange hier die eingefügte vierte deBruijn-Familie (links).

Dieses Verfahren kann man auf beliebige Teilungen von $2n$ -gonen erweitern, indem es iterativ angewandt wird, wobei in jedem Schritt eine neue deBruijn-Familie wie oben beschrieben eingefügt wird. Es entsteht in Abhängigkeit von der gewählten Ausgangsteilung eines n -Ecks ein $(n + 2)$ -Eck.

Ebenso wie im einfachsten Fall der einer Lozengeteilung sind die oktagonale Teilungen mit der vorgeschlagenen Dynamik schnell mischend. Jedoch benötigt man zum Beweis dieser Eigenschaft feinere Methoden, Destainville benutzt z.B. in [13] ein Dekompositionsverfahren. Dieses erlaubt es die Originaldynamik in Teildynamiken zu zerlegen, deren Zustandsräume sich auf gewisse Weise überlappen. Man kann dann von den Teildynamiken die Eigenschaft des schnellen Mischens zeigen und daraus Rückschlüsse auf das Verhalten des Originals ziehen.

Für die Erweiterung auf drei- und mehrdimensionale Teilungen durch analog definierte Rhomboeder gibt es bis dato sehr wenige Ergebnisse und Vermutungen. Insbesondere für die dreidimensionalen Dominoteilungen, siehe z.B. [38], gibt es ähnliche Vermutungen bzgl. arktischer Grenzkörper wie im zweidimensionalen Fall.

Bei den Rhomboederteilungen scheinen darüberhinaus Schleifeneffekte aufzutreten, die im mehrdimensionalen die Eigenschaft der schnellen Mischung vereiteln können, Gedanken dazu findet man in [11].

A Anhang

Abbildungsverzeichnis

1	Die originalen Calissons aus Aix (Frankreich).	1
2	Beispielteilung im Hexagon $\mathcal{H}_{3,4,6}$	2
3	Dreieckgitter, Bienenwabengraph und kubischer Graph, jeweils mit perfect matching (die Dimere werden durch die blauen Balken kodiert).	3
4	Die Abbildung zeigt die den perfect matchings der vorigen Abbildung zugeordneten Teilungen.	4
5	Interpretation des Isingmodells (links) als perfect matching auf dem Fisher-Graphen (rechts). Rechts sind die Dimere rot markiert, die Spins des Isingmodells werden durch die beiden Farben grau und weiß dargestellt, die gestrichelten Kreise geben ein Beispiel für die Interpretation von Kanten zwischen gegensätzlich geladenen Spins und Teilen des Fisher-Graphen mit perfect matching.	4
6	Man kann auf dem Dreieckgitter eine Art Koordinatensystem durch die drei Richtungen a , b und c definieren und damit beliebige Randpolygone beschreiben.	8
7	Beispiel eines nichtkonvexen Gebietes mit exemplarischer Teilung	9
8	nichtteilbare Region mit Färbungseigenschaft	9
9	$\mathcal{H}_{2,3,4}$. Die spezielle Auslegung mit Lozenges suggeriert eine Box der Größe $a \times b \times c$	10
10	irreguläres Hexagon mit Vervollständigung zum Parallelogramm .	11
11	Mögliche Zerlegung der dunkelblau umrandeten Region in teilbare Subregionen.	12
12	deBruijn-Familie in $\mathcal{H}_{3,4,6}$ für die a -Richtung	13
13	Übersetzung von $T \in \mathcal{L}_{\mathcal{H}_{3,4,6}}$ mittels deBruijn-Familie in eine Darstellung im gerichteten euklidischen Gitter.	14
14	Vertauschen der Enden zweier Pfade (tailswap)	18
15	Variationsbereiche möglicher Pfade von $A_2 \rightarrow B_2$ resp. $A_1 \rightarrow B_3$ sind blau unterlegt.	19
16	Logarithmischer Plot der Anzahl Lozenge-Teilungen von $\mathcal{H}_{n,n,n}$ für $n = 1 \dots 50$ (Punkte) und überlagerter logarithmischer Plot der Funktion $e^{0,779538296 \cdot x^2}$ (rot).	26
17	Hier dargestellt der stochastische Fluß einer Kopplung von 25 Kopien einer Markovkette mit Ordnungserhaltung. Die roten Pfade indizieren die Entwicklung der extremen Trajektorien.	38
18	isometrische Projektion des YOUNG-Diagramms aus dem Beispiel mit eingezeichneten de-Bruijn-Pfaden (rot).	41
19	Vollständige Anordnung der Elemente aus $\mathcal{P}_{2,2}^2$ bezüglich „ \preceq “ . .	46
20	Durchlaufrichtung der Einträge einer plane partition aus $\mathcal{P}_{6,6}^a$, für a beliebig.	47

21	Beispiel für die Bezeichnung durch G- und B-Einträge für zwei vergleichbare plane partitions (links, mitte) aus $\mathcal{P}_{1,6,6}$. Es wird deutlich, daß die Anzahl G-Einträge nur deswegen größer ist, da eines der beiden Polyminos den Rand berührt (rechts).	50
22	Plot der empirischen Dichten der Koaleszenzzeiten und Fit an die Dichte einer Gumbelverteilung, sowie die Schätzung von T aus Abschnitt 4.4.	53
23	Initiale Matrizen für die Erzeugung von Stichproben aus $\mathcal{P}_{3,3}^a$. Die Randbedingungen sind blau unterlegt und fix während der ganzen Simulation. Links Initialisierung für den Start der Simulation in der minimalen Teilung \perp , rechts in der maximalen \top	54
24	Teilung von $\mathcal{H}_{100,100,100}$ mit eingezeichnetem arktischen Kreis (gelbe Linie)	62
25	Eine Art arktische Tulpenkurve entsteht, wenn eine Dreiecksblockmatrix fixiert wird. Die Teilung ist aus $\mathcal{H}_{80,80,80}$	64
26	Die Kurve die bei fixierten mittleren Einträgen entsteht ist nicht mehr zusammenhängend. Simuliert wurde eine Teilung von $\mathcal{H}_{90,90,90}$ mit auf den Wert 45 fixierten Einträgen in der Mitte.	65
27	Hexagonale Region mit „spannungsfreien Rändern,“ und Beispieltteilung	66
28	Oktagonteilung und die beiden einzigen nicht-kongruenten Pflastersteine der Teilung.	66
29	Verschiedene Rhombusteilungen für $2n$ -Ecke, $n = 2..5$	67
30	Die vier de Bruijn-Familien für eine oktagonale Rhombusteilung.	67
31	Das Verfahren zur Erzeugung einer oktagonalen Teilung aus einer Lozengeteilung, verdeutlicht für den Fall $n = 3$. Die verschiedenfarbigen Pfade sind die erwähnten Random Walks auf einer zuvor erzeugten Teilung von $\mathcal{H}_{3,3,3}$. Orange hier die eingefügte vierte deBruijn-Familie (links).	68

Tabellenverzeichnis

1	Liste der Anzahlen Lozengeteilungen für regelmäßige Hexagons	22
---	--	----

Literatur

- [1] M. Aigner. Lattice paths and determinants. *Lecture Notes in Computer Science, Buch: Computational Discrete Mathematics*, 2122/2001:1–12, 2001.
- [2] D. J. Aldous. Random walks on finite groups and rapidly mixing markov chains. *Séminaire de probabilités de Strasbourg*, 17:243–297, 1983.
- [3] M. Baake. A guide to mathematical quasicrystals. *arXiv:math-ph/9901014v1*, 1999.
- [4] E. Behrends. *Introduction to Markov Chains*. Vieweg Verlag, 2000.
- [5] P. Brémaud. *Markov chains, Gibbs fields, Monte Carlo simulation, and queues*. Springer, New York, 1999.
- [6] M. Ciucu, T. Eisenkölbl, C. Krattenthaler, and D. Zare. Enumeration of lozenge tilings with central triangular hole. *Journal of Combinatorial Theory, Series A*, 95:251–334, 2001.
- [7] H. Cohn, R. Kenyon, J. Propp, D. Pieter, and W. Kasteleyn. A variational principle for domino tilings. *Journal of the American Mathematical Society*, 14:297–346, 1998.
- [8] H. Cohn, M. Larsen, and J. Propp. The shape of a typical boxed plane partition. *New York J. Math*, 4:137–166, 1998.
- [9] G. a. T. David. The problem of the calissons. *American Mathematical Monthly*, 96(5):429–431, 1989.
- [10] N. G. deBruijn. Dualization of multigrids. *Proceedings of the International Workshop Aperiodic Crystals, Journal de Physique, Colloque C3, supplement to nr. 7*, 47:85–94, 1960.
- [11] V. Desoutter and N. Destainville. Flip dynamics in three-dimensional random tilings. *Journal of Physics A: Mathematical and General*, pages 17–45, 2005.
- [12] S. Desreux. An algorithm to generate exactly once every tiling with lozenges of a domain. *Theoretical Computer Science archive*, 303:375 – 408, 2003.
- [13] N. Destainville. Mixing times of plane random rhombus tilings. *Discrete Mathematics and Theoretical Computer Science, Proceedings AA*, pages 1–22, 2001.
- [14] N. Destainville, M. Widom, R. Mosseri, and F. Bailly. Random tilings of high symmetry: I. mean-field theory. *Journal of Statistical Physics*, 120:799, 2005.

- [15] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Review*, 41:45–76, 1999.
- [16] C. L. Dodgson. Condensation of determinants, being a new and brief method for computing their arithmetical values. *Proceedings of the Royal Society of London*, pages 150–155, 1866.
- [17] R. Durrett. *Probability : theory and examples*. Thomson Brooks/Cole, Duxbury advanced series, 2005.
- [18] J. A. Fill. An interruptible algorithm for perfect sampling via markov chains. *Annals of Applied Probability*, 8:131–162, 1998.
- [19] M. Fischer and H. Temperley. Association problem in statistical mechanics - critique of the treatment. *Reviews of Modern Physics*, 32(4):1029–1031, 1960.
- [20] M. Fisher. Statistical mechanics of dimers on a plane lattice. *Physical Review*, 124(6):1664–&, 1961.
- [21] S. G. Foss and R. L. Tweedie. Perfect simulation and backward coupling. *Stoch. Models*, 14:187–203, 1998.
- [22] I. Gessel and G. Viennot. Binomial determinants, paths, and hook length formulae. *Advances in Mathematics* 58, pages 300–321, 1985.
- [23] O. Häggström. *Finite Markov chains and algorithmic applications, 5th reprint*. Cambridge University Press, 2008.
- [24] P. Kasteleyn. Statistics of dimers on a lattice, number of dimer arrangements on a quadratic lattice. *Physica*, 27(12):1209–76*, 1961.
- [25] P. Kasteleyn. Dimer statistics and phase transitions. *Journal Of Mathematical Physics*, 4(2):287–&, 1963.
- [26] R. Kenyon. The planar dimer model with boundary: a survey. In *AMS-CRM Monogr*, pages 307–328. press, 2000.
- [27] R. Kenyon. An introduction to the dimer model - lecture notes. *arXiv:math/0310326v1 [math.CO]*, 2008.
- [28] R. Kenyon and A. Okounkov. Limit shapes and the complex burgers equation, 2007.
- [29] C. Krattenthaler. Advanced determinant calculus, paper b42q. *Sém. Lotharingien Combin.* 42 ("The Andrews Festschrift"), 1999.
- [30] D. A. Levin, Y. Peres, and und Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, 2008.
- [31] B. Lindstroem. On the vector representation of induced matroids. *Bulletin London Mathematical Society*, 5:85–90, 1972.

- [32] T. Lindvall. *Lectures on the coupling method*. Courier Dover Publications, 2002.
- [33] M. Luby, D. Randall, and A. Sinclair. Markov chain algorithms for planar lattice structures. *SIAM J. Comput.*, 31(1):167–192 (electronic), 2001.
- [34] P. A. MacMahon. Memoir on the theory of the partitions of numbers. part v: Partitions in two-dimensional space. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 211:75–110, 1912.
- [35] P. A. MacMahon. *Combinatory Analysis*, volume 2. Cambridge University Press, 1960.
- [36] J. Propp and D. Wilson. Coupling from the past: a user’s guide. *Microsurveys in Discrete Probability, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 41:181–192, 1998.
- [37] J. Propp and D. B. Wilson. Exact sampling with coupled markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9:223–252, 1996.
- [38] D. Randall and G. Yngve. Random three-dimensional tilings of aztec octahedra and tetrahedra: an extension of domino tilings. *Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms*, pages 636–645, 2000.
- [39] R. P. Stanley. *Enumerative Combinatorics*, volume II. Cambridge University Press, 2001.
- [40] R. P. Stanley. *Enumerative Combinatorics*, volume I. Cambridge University Press, 2002.
- [41] W. P. Thurston. Conway’s tiling groups. *Special Geometry Issue*, 97:757–773, 1990.
- [42] W. P. Thurston. Conway’s tiling groups. *American Mathematical Monthly*, 97:757–773, 1990.
- [43] M. Widom, N. Destainville, R. Mosseri, and F. Bailly. Random tilings of high symmetry: I. mean-field theory. *Journal Of Statistical Physics*, 120(5-6):799–835, 2005.
- [44] M. Widom, N. Destainville, R. Mosseri, and F. Bailly. Random tilings of high symmetry: II. boundary conditions and numerical studies. *Journal Of Statistical Physics*, 120(5-6):837–873, 2005.
- [45] D. B. Wilson. Mixing times of lozenge tiling and card shuffling markov chains. *Annals of Applied Probability*, 14:274–325, 2004.

B Eidesstattliche Erklärung

Ich versichere hiermit an Eides Statt, daß ich die von mir eingereichte Diplomarbeit selbständig verfasst und ausschließlich die angegebenen Hilfsmittel benutzt habe.

Potsdam, den 12. Februar 2009