

Wilfried Huss

INTERNAL DIFFUSION-LIMITED AGGREGATION

Diplomarbeit

zur Erlangung des akademischen Grades Diplomingenieur der Studienrichtung
Technische Mathematik eingereicht an der

Technischen Universität Graz

Betreuer: Univ.-Prof. Dipl.-Ing. Dr.rer.nat Wolfgang Woess

Institut für Mathematik C

Technische Universität Graz

November 2005

Inhaltsverzeichnis

1	Einführung	5
1.1	Aufbau der Arbeit	5
2	Das Modell	7
3	Internal Diffusion-Limited Aggregation auf \mathbb{Z}^d	13
3.1	Vorraussetzungen	14
3.2	Die untere Schranke	28
3.3	Die obere Schranke	35
3.4	Weitere Ergebnisse	48
4	Internal Diffusion-Limited Aggregation als Computing Modell	51
4.1	Die Klasse CC	53
4.2	iDLA Vorhersage	54
5	Internal-Diffusion-Limited Aggregation auf Bäumen	59
5.1	Bäume mit endlich vielen Kegelklassen	59
5.2	Das Simulationsprogramm	62

6	Simulationsergebnisse	67
6.1	Homogene Bäume	67
6.2	Der Kamm	69
6.2.1	Einfache Irrfahrt	70
6.2.2	Inhomogene Irrfahrten	74
6.3	Sternförmige Kämmen	77
6.4	Kämmen in höheren Dimensionen	78
7	Jüngere Entwicklungen	87
7.1	Gruppen mit exponentiellem Wachstum	87
7.2	Rotor Router	91
7.3	Ausblick	93

Kapitel 1

Einführung

Diese Arbeit beschäftigt sich mit dem Studium der sogenannten *internal Diffusion-Limited Aggregation*, einem diskreten stochastischen Prozess welcher 1991 von Diaconis und Fulton im Zuge einer Untersuchung bestimmter algebraischer Strukturen eingeführt wurde [5]. Internal Diffusion-limited Aggregation ist ein Wachstumsprozess auf einem unendlichen Graphen, das heißt der Prozess bildet eine aufsteigende Folge von zufälligen Teilmengen des zugrundeliegenden Graphen.

1.1 Aufbau der Arbeit

In Kapitel 2 wird das stochastische Wachstumsmodell *internal Diffusion-Limited Aggregation* (iDLA) eingeführt. Weiters enthält das Kapitel einen Vergleich von iDLA mit dem verwandten, älteren Modell (*external*) *Diffusion-limited aggregation* (DLA) von Witten und Sander.

In Kapitel 3 wird die erste Arbeit zu iDLA von Lawler, Bramson und Griffeath [9] aus dem Jahre 1991 präsentiert. Sie enthält das erste Gestaltstheorem für iDLA bezüglich der einfachen Irrfahrt auf dem d -dimensionalen Gitter. Die hier eingeführten Methoden wurden mittlerweile für iDLA auf anderen Graphen verallgemeinert. Wir zitieren auch kurz noch Ergebnisse zur Größe der Abweichungen der Form der iDLA-Cluster von der theoretischen Gestalt. Diese Arbeiten sind von Lawler [10] sowie Blachère [3].

In Kapitel 4 untersuchen wir iDLA aus der Sicht der theoretischen Informatik. Moore und Machta [16] definieren mittels iDLA ein computing Modell. Dieses ist allerdings schwächer als klassische Turingberechenbarkeit, da iDLA Information zwar verarbeiten, aber nicht beliebig kopieren kann.

Kapitel 5 stellt eine spezielle Klasse rekursiv definierter Bäume vor. Diese sogenannten *Bäume mit endlich vielen Kegelklassen* werden durch einen endlichen Automaten definiert. Weiters stellen wir ein Simulationsprogramm vor, welches in der Lage ist iDLA-Cluster auf diesen Graphen zu simulieren. Die zugrundeliegenden Irrfahrten sind dabei jeweils an die Baumstruktur angepasst.

In Kapitel 6 besprechen wir einige Simulationsergebnisse, die mit dem oben genannten Programm erzielt wurden. Behandelt werden dabei hauptsächlich der zweidimensionale Kamm, das ist der Teilgraph des zweidimensionalen Gitters aus dem alle horizontalen Kanten mit Ausnahme der y -Achse entfernt wurden, sowie höherdimensionale Verallgemeinerungen dieses Graphen. Neben den rein numerischen Ergebnissen werden auch einige Vermutungen für die Gestalt der iDLA-Cluster in diesen Fällen aufgestellt.

Kapitel 7 schließlich stellt jüngere theoretische Ergebnisse vor. Einen, zum Resultat von Lawler, Bramson und Griffeath äquivalenten, Satz von Blachère und Brofferio [4] aus dem Jahr 2005, der die Gestalt der Aggregations-Cluster erklärt, wenn iDLA auf Cayleygraphen von Bäumen mit exponentiellen Wachstum durchgeführt wird. Außerdem stellen wir noch ein zu iDLA ähnliches, deterministisches Modell vor. Anstatt von Irrfahrten wird dabei das von Jim Propp eingeführte Rotor-Routing verwendet. Levine und Peres zeigen in [12] die kugelförmige Gestalt von Rotor-Routing-Clustern im d -dimensionalen Zahlengitter, allerdings ist dieses Resultat etwas schwächer als im Fall von iDLA.

Mein Dank gilt vor allem meiner Familie für ihre Unterstützung. Weiters danken möchte ich Herrn Dr. Elmar Teuffl, der mir bei Computer Problemen immer mit Rat und Tat behilflich war, sowie Frau Birgit Vogtenhuber für das aufmerksame Korrekturlesen des Manuskripts. Für alle verbliebenen Fehler trage natürlich ich die volle Verantwortung. Und nicht zuletzt danke ich auch meinem Betreuer Herrn Univ.-Prof. Dr. Wolfgang Woess.

Kapitel 2

Das Modell

Zur Definition des stochastischen Wachstumsmodells *internal Diffusion-Limited Aggregation* benötigen wir Irrfahrten (bzw. Markov-Ketten). Für Details siehe [20] und [23]. Wir folgen weitgehend der Notation aus [23].

Definition 1. Eine Irrfahrt ist gegeben durch ihren Zustandsraum S , sowie einer Übergangsmatrix $P = (p(x, y))_{x, y \in S}$. P ist stochastisch, das heißt für alle $x \in S$ gilt:

$$\sum_{y \in S} p(x, y) = 1$$

Zusätzlich wird noch ein Startpunkt x_0 benötigt.

Die Elemente $p(x, y)$ geben die Einschrittübergangswahrscheinlichkeit von x nach y an. Die Zufallsvariable $X(t)$ bezeichnet die zufällige Position der Irrfahrt zum Zeitpunkt t .

Die Verteilung der Trajektorie $X(t)$ ergibt sich aus den Übergangswahrscheinlichkeiten wie folgt:

$$\mathbb{P}_x[X(0) = x_0, X(1) = x_1, \dots, X(n) = x_n] = \delta_x(x_0)p(x_0, x_1) \cdots p(x_{n-1}, x_n)$$

\mathbb{P}_x bezeichnet dabei die Wahrscheinlichkeit eines Ereignisses bezüglich einer Irrfahrt mit Startpunkt x . Für den Erwartungswert verwenden wir analog \mathbb{E}_x .

Die n -schritt Übergangswahrscheinlichkeiten

$$p^{(n)}(x, y) := \mathbb{P}_x[X(n) = y]$$

sind gegeben durch die n -te Potenz der Übergangsmatrix P . $P^n = (p^{(n)}(x, y))_{x, y \in S}$.

Im Folgenden ist der Zustandsraum der betrachteten Irrfahrten jeweils ein lokalendlicher Graph $G = (V, E)$. Wir nennen einen Graph lokalendlich, wenn der Knotengrad $\deg(x) < \infty$, für alle $x \in V$. Weiters bezeichnen wir mit \sim die Nachbarschaftsrelation auf G . Wenn keine Verwechslungsgefahr besteht, setzen wir, um die Notation zu vereinfachen, den Graphen mit seiner Knotenmenge gleich.

Definition 2. Sei G ein lokalendlicher Graph, dann heißt die durch die Übergangswahrscheinlichkeiten

$$p(x, y) = \begin{cases} \frac{1}{\deg x} & \text{wenn } x \sim y \\ 0 & \text{sonst} \end{cases}$$

definierte Irrfahrt, *einfache Irrfahrt* auf G .

Internal Diffusion-limited Aggregation ist ein stochastischer Clusterwachstumsprozeß. Von einem Aggregationskeim, typischerweise ein einzelner Knoten, ausgehend erhalten wir eine aufsteigende Folge von zufälligen Mengen $\{A(n)\}_{n \geq 0}$. Für jeden Wachstumsschritt wird eine Irrfahrt gestartet. Sobald sie den aktuellen Cluster $A(n)$ verlässt, wird die Irrfahrt gestoppt, und der Aggregationscluster um den Knoten, an dem der Cluster $A(n)$ verlassen wurde, erweitert. Diese Beschreibung konkretisieren wir in der folgenden Definition.

Definition 3 (internal Diffusion-Limited Aggregation - iDLA). Seien G ein lokalendlicher Graph mit Wurzel \mathcal{O} , und $\{X^i(t)\}_{i \geq 0}$ eine Folge von unabhängigen, identisch verteilten Irrfahrten auf G . Weiters gelte $X^i(0) = \mathcal{O}$ für alle i . Die Folge $\{A(n)\}_{n \in \mathbb{N}}$ wird zusammen mit der Stoppzeit σ^i wie folgt definiert:

$$A(1) = \mathcal{O}$$

und für $i > 1$

$$\begin{aligned} \sigma^i &= \min \{t \geq 0 : X^i(t) \notin A(i-1)\} \\ A(i) &= A(i-1) \cup X^i(\sigma^i). \end{aligned}$$

σ^i ist der Zeitpunkt, an dem die i -te Irrfahrt die Menge $A(i - 1)$ verlässt.

Der stochastische Prozess $\{A(n)\}_{n \in \mathbb{N}}$ heißt *internal Diffusion-Limited Aggregation*, kurz iDLA. Die Menge $A(n)$ heißt n -ter *Aggregationscluster*.

Bemerkung 1. iDLA wurde zum ersten mal in [5] von Diaconis und Fulton beschrieben.

Bemerkung 2. Zunächst betrachten wir iDLA auf dem d -dimensionalen Gitter \mathbb{Z}^d mit $d \geq 2$, bezüglich der einfachen Irrfahrt.

iDLA ist eine Abwandlung des bekannten Wachstumsmodells *Diffusion-Limited Aggregation* (DLA). Dieses Modell wurde 1981 von Witten und Sander [22] eingeführt, um Aggregate kondensierender Metallpartikel zu studieren. DLA wurde seitdem zur Modellierung diverser physikalischer und biologischer Wachstumsprozesse verwendet [16]. Die Anwendungsgebiete reichen vom Schneeflockenwachstum über die Bildung von Wasserscheiden bis hin zur Untersuchung von Kapillarnetzwerken in Organen und dem Wachstum von Neuronen.

Bei DLA starten die Irrfahrten außerhalb des Aggregationsclusters und stoppen sobald sie einen Gitterknoten erreichen, der an den Wachstumscluster angrenzt. Dieser Knoten wird zum Cluster hinzugefügt und der Prozess startet von Neuem.

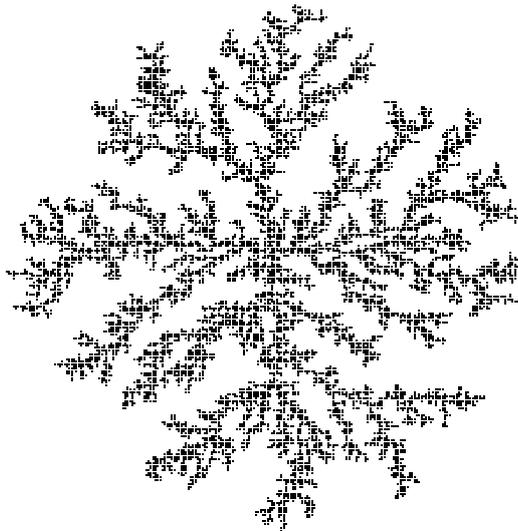


Abbildung 2.1: DLA Simulation mit 10000 Teilchen in \mathbb{Z}^2

Definition 4. Sei $A \subset \mathbb{Z}^d$, dann heißt

$$\partial A = \{x \in \mathbb{Z}^d \setminus A : \exists z \in A, |z - x| = 1\}$$

äußerer Rand von A . Weiters definieren wir den Radius der Menge A durch

$$\text{rad}(A) = \max\{\|x\| : x \in A\}.$$

Definition 5. Sei $\{A(n)\}_{n \in \mathbb{N}}$ mit $A(n) \subset \mathbb{Z}^d$ ein stochastischer Prozess von Teilmengen des d -dimensionalen Gitters mit den folgenden Eigenschaften:

- $A(1) = \{0\}$
- Und für $n > 1$ und $x \in \partial A(n)$ gelte

$$\mathbb{P}[A(n+1) = A(n) \cup \{x\} | A(n)] = H_{\partial A(n)}(x).$$

Dann heißt $\{A(n)\}_{n \in \mathbb{N}}$ *Diffusion-Limited Aggregation* auf \mathbb{Z}^d .

$H_{\partial A(n)}(x)$ ist dabei die Wahrscheinlichkeit, dass eine Irrfahrt „aus dem Unendlichen“ kommend den Rand der Menge $A(n)$ im Punkt x trifft.

„Aus dem Unendlichen“ bedeutet hier konkret, für jeden Zeitpunkt d definieren wir zwei Kreise mit Mittelpunkt 0 und den Radien $r_n^t \gg r_n^g$. Weiters soll gelten $r_n^g > \text{rad}(A(n))$, also der innere der beiden Kreise soll den Aggregationscluster zu jedem Zeitpunkt voll umschließen.

Die n -te Irrfahrt startet jeweils in einem zufälligen Punkt aus dem „Geburts-Kreis“ mit Radius r_n^b . Verlässt die i Irrfahrt den „Todes-Kreis“ mit Radius r_n^t , beginnt die Irrfahrt erneut in einem zufälligen Punkt im „Geburts-Kreis“. Dieser Vorgang wird solange wiederholt bis sich das Teilchen an den Aggregationscluster anlagert. Durch diese Prozedur lässt sich ein endlicher Erwartungswert für Anzahl der Schritte, die für die Anlagerung eines Partikels benötigt werden, garantieren.

Es gibt bisher nur wenige rigorose Resultate über die Eigenschaften von DLA-Clustern. Durch Computersimulationen erzeugte Cluster zeigen eine relativ „ausgedünnte“, dendritische Struktur. Die Cluster haben eine nicht ganzzahlige fraktale Dimension $\frac{1}{\alpha}$, mit α definiert als

$$\alpha = \limsup_{n \rightarrow \infty} \frac{\ln \mathbb{E}[\text{rad}(A_n)]}{\ln n}.$$

Computersimulationen ergaben zum Beispiel für DLA in 2 Dimensionen eine fraktale Dimension von $1.715 \pm .004$. Eine Beschreibung von effizienten Methoden zur Simulation von Diffusion-Limited Aggregation sowie eine Reihe von Simulationsergebnissen für DLA in 2 bis 8 Dimensionen finden sich bei Tolman und Meakin [21].

Bisher ist keine untere Schranke für die Dimension bekannt. In Abschnitt 2.6 von [7] liefert Lawler neben einer kurzen Einführung in die Thematik auch eine obere Schranke, die allerdings von den numerischen Ergebnissen noch recht weit entfernt ist.

Kapitel 3

Internal Diffusion-Limited Aggregation auf \mathbb{Z}^d

In diesem Kapitel beschäftigen wir uns mit dem Studium des Grenzverhaltens der *internal Diffusion-Limited Aggregation* auf dem diskreten d -dimensionalen Gitter.

Wir suchen also eine geometrisch möglichst einfach zu beschreibende Familie von Teilmengen $\mathcal{F}(n) \subset \mathbb{Z}^d$ ($n \in \mathbb{N}$), so dass sich die Folge der Aggregationscluster $A(i)$ für $i \rightarrow \infty$ den Mengen $\mathcal{F}(i)$ mit Wahrscheinlichkeit 1 beliebig annähert.

Bereits eine einfache numerische Simulation für $d = 2$ (Abbildung 3.1) zeigt, dass die Aggregationscluster $A(i)$ sehr schnell Kugelgestalt annehmen.

Diese Vermutung wird im folgenden Satz von Lawler, Bramson und Griffeath [9] konkretisiert. Zunächst benötigen wir allerdings noch eine Definition der diskreten, euklidischen Kugeln.

Definition 6. Die Menge $\mathcal{B}_r = \{x \in \mathbb{Z}^d : \|x\| < r\}$ heißt d -dimensionale, diskrete, euklidische Kugel mit Radius r .

Weiters bezeichne $\omega_d = |\{x \in \mathbb{R}^d : \|x\| \leq 1\}|$ das Volumen der euklidischen Einheitskugel.

Bemerkung 3. Im Folgenden werden reellwertige Indices jeweils abgerundet. Das heißt für $t \in \mathbb{R}$ wird $\mathcal{B}_t := \mathcal{B}_{\lfloor t \rfloor}$ und $A(t) := A(\lfloor t \rfloor)$ definiert.

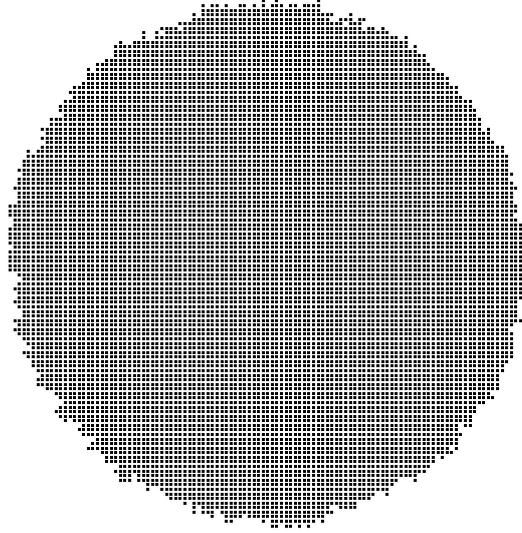


Abbildung 3.1: iDLA Simulation von 10000 Partikel in \mathbb{Z}^2

Satz 1. Für jedes $\varepsilon > 0$ existiert eine Konstante n_0 , sodass gilt

$$\mathbb{P} [\mathcal{B}_{n(1-\varepsilon)} \subseteq A(\omega_d n^d) \subseteq \mathcal{B}_{n(1+\varepsilon)}, \forall n \geq n_0] = 1. \quad (3.1)$$

Bemerkung 4. Der Beweis von Satz 1 erfolgt in zwei Schritten. Zuerst zeigen wir in Abschnitt 3.2 die innere Schranke

$$\mathbb{P} [\mathcal{B}_{n(1-\varepsilon)} \subseteq A(\omega_d n^d), n \geq n_0] = 1 \quad (3.2)$$

In Abschnitt 3.3 wird schließlich die äußere Schranke

$$\mathbb{P} [A(\omega_d n^d) \subseteq \mathcal{B}_{n(1+\varepsilon)}, n \geq n_0] = 1 \quad (3.3)$$

auf (3.2) zurück geführt.

3.1 Voraussetzungen

Der Beweis von Satz 1 stützt sich in wesentlichen Punkten auf asymptotische Abschätzungen von *Greenschen Funktionen*. Dieser Abschnitt stellt die benötigten technischen Lemmata bereit, auf die der Beweis im weiteren zurückgreift.

Definition 7. Für $d \geq 3$ und $y \in \mathbb{Z}^d$ ist die *Greensche Funktion* G definiert als

$$G(y, z) = \mathbb{E}_y \left[\sum_{t=0}^{\infty} \mathbb{1}_{\{X(t)=z\}} \right] = \sum_{t=0}^{\infty} \mathbb{P}_y[X(t) = z] \quad \text{für } z \in \mathbb{Z}^d.$$

Für die Stoppzeit $\tau_n = \min\{t : X(t) \notin \mathcal{B}_n\}$ definieren wir die *gestoppte Greensche Funktion* G_n als

$$G_n(y, z) = \mathbb{E}_y \left[\sum_{t=0}^{\tau_n-1} \mathbb{1}_{\{X(t)=z\}} \right] \quad \text{für } z \in \mathcal{B}_n.$$

Bemerkung 5.

- $G(y, z)$ ist die erwartete Anzahl der Besuche im Punkt z einer im Punkt y startenden Irrfahrt.
- Analog bezeichnet $G_n(y, z)$ die erwartete Anzahl der Besuche im Punkt z , bevor die Irrfahrt die Kugel \mathcal{B}_n zum ersten mal verlässt.
- In weiterer Folge benützen wir die Notationen $G(z) = G(0, z)$ sowie $G_n(z) = G_n(0, z)$.

Bemerkung 6. Satz 1 gilt für beliebige Dimensionen $d \geq 2$. Im folgenden werden wir den Beweis allerdings nur für $d \geq 3$ ausführen. Der Beweis verwendet in wesentlichen Punkten diverse Abschätzungen der Greenschen Funktion $G(y, z)$. Nun ist nach dem Satz von Polya (siehe dazu [23]) die einfache Irrfahrt auf \mathbb{Z}^2 rekurrent, und somit divergiert $G(y, z)$. Es lässt sich in diesem Fall die Greensche Funktion allerdings durch den sogenannten *Potential Kern*

$$a(y, z) = \lim_{n \rightarrow \infty} \mathbb{E}_y \left[\sum_{t=0}^n (\mathbb{1}_{\{X(t)=y\}} - \mathbb{1}_{\{X(t)=z\}}) \right]$$

ersetzen. Für alle benötigten Abschätzungen der Greenschen Funktion können analoge Aussagen für den Potential Kern verwendet werden. Für Details zum Beweis von Satz 1 für $d = 2$ siehe [9].

Die Folgenden Resultate geben die für den Beweis von Satz 1 benötigten Abschätzungen der Greenschen Funktion und der gestoppten Greenschen Funktion. Für die Beweise siehe [7, 9].

3.1. VORRAUSSETZUNGEN

Lemma 1. Sei $d \geq 3$, dann gilt für $a_d = \frac{2}{(d-2)\omega_d}$ und $\|z\| \rightarrow \infty$,

$$G(x) \sim a_d \|z\|^{2-d}.$$

Weiters gilt für alle $\alpha < d$

$$\lim_{\|z\| \rightarrow \infty} \|z\|^\alpha (G(x) - a_d \|z\|^{2-d}) = 0. \quad (3.4)$$

Bemerkung 7. Aus (3.4) folgt für $\alpha = d - 1$,

$$G(z) = a_d \|z\|^{2-d} + o(\|z\|^{1-d}).$$

Lemma 2. Sei $z \in \mathcal{B}_z$ und $d \geq 3$, dann existiert eine Konstante $\mathcal{C} < \infty$ sodass gilt

$$(a) \quad G_n(z) \leq G(z) + \mathcal{C} \cdot n^{2-d},$$

$$(b) \quad G_n(z) \leq a_d (\|z\|^{2-d} - n^{2-d}) + \mathcal{C} \cdot \|z\|^{1-d} \quad \text{für } \|z\| \geq 1.$$

Proposition 1. Sei $\varepsilon > 0$ und $z \in \mathcal{B}_{n(1-\varepsilon)}$, dann gelten die folgenden Ungleichungen

$$G_{\varepsilon n}(0, 0) \leq G_n(z, z) \leq G_{2n}(0, 0).$$

Proposition 2. Sei $A \subset \mathbb{Z}^d$ endlich und $\tau = \inf\{j \geq 1 : X(j) \notin A\}$, dann existieren Konstanten $\mathcal{C} < \infty$ und $\rho = \rho(A) < 1$, so dass $\forall z \in A$ gilt

$$\mathbb{P}_z[\tau \geq n] \leq \mathcal{C} \rho^n. \quad (3.5)$$

Beweis. Sei $R := \sup_{z \in A} \|z\|$, dann gilt, $\forall z \in A$ existiert ein Pfad der Länge $R + 1$ der in z beginnt und außerhalb der Menge A endet. Das heißt, folgt die Irrfahrt diesem speziellen Pfad, so verlässt sie die Menge A nach höchstens $R + 1$ Schritten. Also ist

$$\mathbb{P}_z[\tau \leq R + 1] \geq \left(\frac{1}{2d}\right)^{R+1}.$$

Letzteres ist die Wahrscheinlichkeit eines fixen Pfades der Länge $R + 1$.

Wir beweisen (3.5) zunächst nur für spezielle Werte von n mit der Form $n = k(R + 1)$ und $k \in \mathbb{N}$. Dazu betrachten wir die Wahrscheinlichkeit:

$$\mathbb{P}_z[\tau > k(R + 1)] = \mathbb{P}_z[\tau > k(R + 1) \cap \tau > (k - 1)(R + 1)] =$$

Anwenden der Definition der bedingten Wahrscheinlichkeit ergibt

$$= \mathbb{P}_z[\tau > (k - 1)(R + 1)] \cdot \mathbb{P}_z[\tau > k(R + 1) | \tau > (k - 1)(R + 1)].$$

Nun ist, wenn $\tau > (k - 1)(R + 1)$ gilt, das Element $z' = X((k - 1)(R + 1)) \in A$. Da aufgrund der Markoveigenschaft die Wahrscheinlichkeit eines Ereignisses für eine Irrfahrt nicht von ihrer Geschichte abhängt, gibt es ein $z' \in A$ sodass

$$\mathbb{P}_z[\tau > k(R + 1) | \tau > (k - 1)(R + 1)] = \mathbb{P}_{z'}[\tau > (R + 1)].$$

Daraus folgt nun

$$\begin{aligned} \mathbb{P}_z[\tau > k(R + 1)] &= \mathbb{P}_z[\tau > (k - 1)(R + 1)] \cdot \mathbb{P}_{z'}[\tau > R + 1] = \\ &= \mathbb{P}_z[\tau > (k - 1)(R + 1)] \cdot (1 - \mathbb{P}_{z'}[\tau \leq R + 1]) \leq \\ &\leq \mathbb{P}_z[\tau > (k - 1)(R + 1)] \cdot \left(1 - \left(\frac{1}{2d}\right)^{R+1}\right). \end{aligned}$$

Mit Hilfe von vollständiger Induktion über k erhält man daraus schließlich

$$\mathbb{P}_z[\tau > k(R + 1)] \leq \left(1 - \left(\frac{1}{2d}\right)^{R+1}\right)^k.$$

Setzt man $\rho = \left(1 - \left(\frac{1}{2d}\right)^{R+1}\right)^{\frac{1}{R+1}}$ dann erhält man die Abschätzung

$$\mathbb{P}_z[\tau > k(R + 1)] \leq \rho^{k(R+1)}.$$

Es bleibt noch dieses Ergebnis auf alle natürlichen Zahlen auszuweiten. Dazu schreiben

3.1. VORRAUSSETZUNGEN

wir $n = k(R + 1) + j$ mit $0 \leq j \leq R$.

$$\begin{aligned} \mathbb{P}_z[\tau \geq n] &\leq \mathbb{P}_z[\tau \geq k(R + 1)] \leq \\ &\leq \rho^{k(R+1)} = \rho^{-j} \cdot \rho^n \leq \\ &\leq \rho^{-R} \cdot \rho^n = C\rho^n \end{aligned}$$

liefert das gewünschte Ergebnis. □

Satz 2 (Optional Stopping). *Sei $\{M(n)\}_{n \geq 0}$ ein Martingal und T eine Stoppzeit bezüglich der Filtration $\{\mathcal{F}(n)\}_{n \geq 0}$. Sei weiters $\mathbb{P}[T < \infty] = 1$, $\mathbb{E}[|M(T)|] < \infty$ und*

$$\lim_{m \rightarrow \infty} \int_{T > m} |M(m)| d\mathbb{P} = 0,$$

dann gilt

$$\mathbb{E}[M(T)] = \mathbb{E}[M(0)].$$

Lemma 3. *Sei $z \in \mathcal{B}_n$ und $\tau_n = \min\{t : X(t) \notin \mathcal{B}_n\}$, dann gilt*

$$n^2 - \|z\|^2 \leq \mathbb{E}_z[\tau_n] \leq (n + 1)^2 - \|z\|^2.$$

Beweis. Wir betrachten den Prozess $M(t) = \|X(t)\|^2 - t$ und die kanonische Filtration bezüglich $X(t)$

$$\mathcal{F}_t = \sigma(\{X(1), \dots, X(t)\}).$$

$$\begin{aligned} \mathbb{E}[M(t + 1)|\mathcal{F}_t] &= \mathbb{E}[\|X(t + 1)\|^2 - (t + 1)|\mathcal{F}_t] \\ &= \mathbb{E}[\|X(t) + Y\|^2|\mathcal{F}_t] - (t + 1) \\ &= \mathbb{E}\left[\sum_{i=1}^d (X(t)_i + Y_i)^2|\mathcal{F}_t\right] - (t + 1) \\ &= \sum_{i=1}^d \mathbb{E}[X_i^2(t)|\mathcal{F}_t] + 2 \sum_{i=1}^d \mathbb{E}[X_i(t)Y_i|\mathcal{F}_t] + \sum_{i=1}^d \mathbb{E}[Y_i^2|\mathcal{F}_t] - (t + 1) \\ &= \sum_{i=1}^d X_i^2(t) + 2 \sum_{i=1}^d X_i(t)\mathbb{E}[Y_i] + \sum_{i=1}^d \mathbb{E}[Y_i^2] - (t + 1) \end{aligned}$$

Die letzte Gleichung gilt, da $X_i(t)$ \mathcal{F}_t -messbar und Y_i unabhängig von \mathcal{F}_t ist. Nun ist aber $\mathbb{E}[Y_i] = 0$ und $\mathbb{E}[Y_i^2] = \frac{1}{d}$ und somit

$$\mathbb{E}[M(t+1)|\mathcal{F}_t] = \|X(t)\|^2 - t = M(t).$$

Der Prozess $M(t)$ ist also ein Martingal. Das Optional-Stopping Theorem angewendet auf M und die Stoppzeit τ_n ergibt $\mathbb{E}_z[M(\tau_n)] = \mathbb{E}_z[M(0)]$, und weiters

$$\mathbb{E}_z [\|X(\tau_n)\|^2 - \tau_n] = \mathbb{E}_z [\|X(\tau_n)\|^2] - \mathbb{E}_z[\tau_n] = \mathbb{E}_z[\|X(0)\|^2] = \|z\|^2.$$

Mit $n^2 \leq \|X(\tau_n)\|^2 \leq (n+1)^2$ ergibt sich das Lemma.

Zu zeigen bleibt noch, dass die Voraussetzungen für Optional-Stopping erfüllt sind:

- (a) $\mathbb{P}_z[\tau_n < \infty] = 1$
- (b) $\mathbb{E}_z[|M(\tau_n)|] < \infty$
- (c) $\lim_{m \rightarrow \infty} \int_{\tau_n > m} |M(m)| d\mathbb{P}_z = 0$

Die Voraussetzungen (a) und (b) sind klarerweise erfüllt.

Zum Beweis von Voraussetzung (c) benötigen wir Proposition 2:

$$\begin{aligned} \int_{\tau_n > m} |M(m)| d\mathbb{P}_z &\leq \sup_{\omega \in \Omega} |M(m, \omega)| \cdot \mathbb{P}_z[\tau_n > m] \\ &\leq \sup_{\omega \in \Omega} |\|X(m, \omega)\|^2 - m| \cdot C\rho^m \\ &\leq |(n+m)^2 - m| \cdot C\rho^m \xrightarrow{m \rightarrow \infty} 0. \end{aligned}$$

Die letzte Ungleichung gilt Aufgrund der Tatsache, dass $X(m)$ höchstens m Schritte von z entfernt sein kann. Somit ist $\|X(m)\| \leq n+m$. \square

Lemma 4. Für $0 < \varepsilon < 1$, $d \geq 3$, $\zeta \geq \frac{1}{1-\varepsilon}$ und $\delta = \delta_\varepsilon \geq 0$ genügend klein, gilt die folgende Ungleichung:

$$\frac{d(1+\delta)}{d-2} \zeta^2 \leq 1 + \frac{2(1+\delta)}{d-2} \zeta^d$$

3.1. VORRAUSSETZUNGEN

Beweis. Es genügt zu zeigen, dass die Funktion

$$\begin{aligned} f(\zeta) &= 1 + \frac{2(1+\delta)}{d-2}\zeta^d - \frac{d(1+\delta)}{d-2}\zeta^2 \\ &= 1 + \frac{1+\delta}{d-2}(2\zeta^d - d\zeta^2) \end{aligned}$$

unter den Voraussetzungen des Lemmas positiv ist. f ist differenzierbar und

$$f'(\zeta) = \frac{1+\delta}{d-2}2d\zeta(\zeta^{d-2} - 1) \geq 0, \text{ da } \zeta > 1.$$

Somit ist f monoton wachsend und es genügt für $0 < \varepsilon < 1$ die Existenz eines δ_ε zu zeigen, sodass für diese Wahl $f(\frac{1}{1-\varepsilon}) \geq 0$ ist.

Dazu setzen wir $\delta = k\varepsilon$ mit $k = k_\varepsilon$ passend gewählt. Dann ist

$$g(\varepsilon) = f\left(\frac{1}{1-\varepsilon}\right) = 1 + \frac{1+k\varepsilon}{d-2} \cdot \frac{2(1-\varepsilon)^{2-d} - d}{(1-\varepsilon)^2},$$

und somit $g(0) = 0$. Für die Ableitung von g gilt

$$g'(\varepsilon) = \frac{(1-\varepsilon)^{-2}}{d-2} \left[k(2(1-\varepsilon)^{2-d} - d) + \frac{2d(1+k\varepsilon)}{1-\varepsilon} ((1-\varepsilon)^{2-d} - 1) \right]. \quad (3.6)$$

Hier ist nun der zweite Summand positiv und zwar mit einer von k unabhängigen unteren Schranke. Der erste Summand ist für $\varepsilon < 1 - (2/d)^{\frac{1}{2-d}}$ negativ, aber nach unten beschränkt durch $k(2-d)$.

Setzen wir nun

$$C_\varepsilon := \frac{(1-\varepsilon)^{2-d} - 1}{1-\varepsilon},$$

dann ist der Klammerausdruck in (3.6) größer als

$$k(2-d) + 2d(1+k\varepsilon) \cdot C_\varepsilon = 2dC_\varepsilon + k(2-d + 2d\varepsilon C_\varepsilon).$$

Da $C_\varepsilon > 0$ für $0 < \varepsilon < 1$ existiert somit für jedes ε ein k , für welches die Ableitung $g'(\varepsilon)$ positiv ist.

Das heißt, g und damit auch f sind positiv unter den Voraussetzungen des Lemmas. \square

Proposition 3. *Sei $z \in \mathcal{B}_n$ und die Stoppzeit $\tau_n = \min\{t : X(t) \notin \mathcal{B}_n\}$, dann gilt für die gestoppten Greenfunktionen G_n*

$$\sum_{y \in \mathcal{B}_n} G_n(z, y) = \mathbb{E}_z[\tau_n].$$

3.1. VORRAUSSETZUNGEN

Beweis. Der Beweis erfolgt durch einfaches Umformen der Definition der gestoppten Greenfunktion. Es gilt

$$\begin{aligned}
\sum_{y \in \mathcal{B}_n} G_n(z, y) &= \sum_{y \in \mathcal{B}_n} \mathbb{E}_z \left[\sum_{t=0}^{\tau_n-1} \mathbb{1}_{\{X(t)=y\}} \right] \\
&= \sum_{y \in \mathcal{B}_n} \sum_{i=0}^{\infty} \mathbb{P}_z[\tau_n = i] \sum_{t=0}^{i-1} \mathbb{P}_z[X(t) = y \mid \tau_n = i] \\
&= \sum_{i=0}^{\infty} \mathbb{P}_z[\tau_n = i] \sum_{t=0}^{i-1} \sum_{y \in \mathcal{B}_n} \mathbb{P}_z[X(t) = y \mid \tau_n = i] \\
&= \sum_{i=0}^{\infty} \mathbb{P}_z[\tau_n = i] \sum_{t=0}^{i-1} 1 = \sum_{i=0}^{\infty} i \cdot \mathbb{P}_z[\tau_n = i] \\
&= \mathbb{E}_z[\tau_n].
\end{aligned}$$

□

Lemma 5. Sei $\varepsilon > 0$, $n > n_\varepsilon$ und $z \in \mathcal{B}_{n(1-\varepsilon)}$, dann gilt

$$\sum_{y \in \mathcal{B}_n} G_n(y, z) \leq \omega_d n^d G_n(0, z). \tag{3.7}$$

Beweis. Aufgrund der Symmetrie der Greenfunktion und Proposition 3 ist die linke Seite der Ungleichung (3.7) gleich $\mathbb{E}_z[\tau_n]$. Weiters gilt nach Lemma 3

$$\mathbb{E}_z[\tau_n] \leq (n+1)^2 - \|z\|^2.$$

Aus Lemma 2(b) folgt für $d \geq 3$ und $\|z\| \geq 1$ andererseits

$$\omega_d n^d G_n(0, z) \geq \frac{2n^d}{d-2} (\|z\|^{2-d} - n^{2-d}) + \mathcal{C} \cdot n^d \|z\|^{1-d}.$$

Es genügt also, zum Beweis des Lemmas die folgende Ungleichung

$$(n+1)^2 - \|z\|^2 \leq \frac{2n^d}{d-2} (\|z\|^{2-d} - n^{2-d}) + \mathcal{C} \cdot n^d \|z\|^{1-d} \tag{3.8}$$

zu zeigen. Division von (3.8) durch $\|z\|^2$ ergibt

$$\left(\frac{n+1}{\|z\|}\right)^2 \leq 1 + \frac{2}{d-2} \left[\left(\frac{n}{\|z\|}\right)^d - \left(\frac{n}{\|z\|}\right)^2 \right] + \frac{\mathcal{C}}{\|z\|} \left(\frac{n}{\|z\|}\right)^d,$$

und durch Herausheben von Termen gleicher Ordnung erhält man daraus zunächst

$$\frac{2n+1}{\|z\|^2} + \frac{d}{d-2} \left(\frac{n}{\|z\|}\right)^2 \leq 1 + \left[\frac{2}{d-2} + \frac{\mathcal{C}}{\|z\|} \right] \left(\frac{n}{\|z\|}\right)^d,$$

und schließlich

$$\frac{d}{d-2} \left[1 + \frac{d-2}{d} \frac{2n+1}{n^2} \right] \left(\frac{n}{\|z\|}\right)^2 \leq 1 + \frac{2}{d-2} \left[1 + \frac{d-2}{2} \frac{\mathcal{C}}{\|z\|} \right] \left(\frac{n}{\|z\|}\right)^d.$$

Mit geeignet gewählten Konstanten \mathcal{J} und $\tilde{\mathcal{C}}$ impliziert somit

$$\frac{d}{d-2} \left[1 + \frac{\mathcal{J}}{n} \right] \left(\frac{n}{\|z\|}\right)^2 \leq 1 + \frac{2}{d-2} \left[1 + \frac{\tilde{\mathcal{C}}}{\|z\|} \right] \left(\frac{n}{\|z\|}\right)^d \quad (3.9)$$

die ursprüngliche Ungleichung (3.8).

Nach Voraussetzung ist $z \in \mathcal{B}_{n(1-\varepsilon)}$. Somit gilt $\frac{n}{\|z\|} \geq \frac{1}{1-\varepsilon}$. (3.9) folgt aus Lemma 4, wenn $n \geq \|z\| \geq r_\varepsilon$ und r_ε so gewählt ist, dass gleichzeitig $\frac{\mathcal{J}}{n} < \delta$ und $\frac{\tilde{\mathcal{C}}}{\|z\|} < \delta$ erfüllt sind.

Es bleibt noch den Fall $\|z\| < r_\varepsilon$ abzuhandeln. Hier folgt aus Lemma 2(a) die hinreichende Bedingung

$$(n+1)^2 - \|z\|^2 \leq \omega_d n^d (G(z) + \mathcal{C}n^{2-d}).$$

Diese Ungleichung ist für genügend große n erfüllt. □

Lemma 6. *Sei S endliche Summe unabhängiger Indikator-Zufallsvariablen mit Erwartungswert $\mathbb{E}(S) = \mu$, dann gilt für alle $0 < \gamma < \frac{1}{2}$ und $\mu \geq \mu_0$*

$$\mathbb{P} \left[|S - \mu| \geq \mu^{1/2+\gamma} \right] \leq 2 \exp \left\{ -\frac{1}{4} \mu^{2\gamma} \right\}.$$

3.1. VORRAUSSETZUNGEN

Beweis. Im Folgenden bezeichnen wir mit φ_X die charakteristische Funktion der Zufallsvariablen X . Wir setzen $S = \sum_{j=1}^n \mathbb{1}_{A_j}$ mit $\mathbb{P}[A_j] = p_j$.

Wir betrachten den Erwartungswert

$$\mathbb{E} [e^{\lambda(S-\mu)}] = \mathbb{E} \left[e^{\lambda(\sum_{j=1}^n \mathbb{1}_{A_j} - \mu)} \right] = \varphi_\mu(i\lambda) \cdot \prod_{j=1}^n \varphi_{\mathbb{1}_{A_j}}(-i\lambda),$$

berechnen die dabei auftretenden charakteristischen Funktionen

$$\begin{aligned} \varphi_\mu(t) &= \mathbb{E} [e^{it\mu}] = e^{it\mu} = e^{it \sum_{j=1}^n p_j} = \prod_{j=1}^n e^{itp_j} \\ \varphi_{\mathbb{1}_{A_j}}(t) &= \mathbb{E} [e^{it\mathbb{1}_{A_j}}] = 1 - p_j + e^{it}p_j \end{aligned}$$

und erhalten somit

$$\mathbb{E} [e^{\lambda(S-\mu)}] = \prod_{j=1}^n e^{-\lambda p_j} \cdot (1 - p_j + e^{\lambda} p_j) = \prod_{j=1}^n [(1 - p_j)e^{-\lambda p_j} + p_j e^{\lambda(1-p_j)}].$$

Mit Hilfe der Ungleichung von Chebyshev erhält man für $\lambda > 0$ die Abschätzung

$$\begin{aligned} \mathbb{P} [S - \mu \geq \mu^{1/2+\gamma}] &= \mathbb{P} [e^{\lambda(S-\mu)} \geq \exp\{\lambda\mu^{1/2+\gamma}\}] \\ &\leq \exp\{-\lambda\mu^{1/2+\gamma}\} \cdot \mathbb{E} [e^{\lambda(S-\mu)}] \\ &= \exp\{-\lambda\mu^{1/2+\gamma}\} \cdot \prod_{j=1}^n [(1 - p_j)e^{-\lambda p_j} + p_j e^{\lambda(1-p_j)}]. \end{aligned} \tag{3.10}$$

Mittels Kurvendiskussion ergibt sich daraus für $\lambda < \frac{3}{2}$ die Abschätzung

$$(1 - p_j)e^{-\lambda p_j} + p_j e^{\lambda(1-p_j)} \leq 1 + \lambda^2 p_j. \tag{3.11}$$

Wir setzen nun (3.11) in (3.10) ein

$$\begin{aligned} \mathbb{P} [S - \mu \geq \mu^{1/2+\gamma}] &\leq \exp \{-\lambda \mu^{1/2+\gamma}\} \cdot \exp \left\{ \sum_{j=1}^n \log(1 + \lambda^2 p_j) \right\} \\ &\leq \exp \left\{ -\lambda \mu^{1/2+\gamma} + \lambda^2 \sum_{j=1}^n p_j \right\} \\ &= \exp \{-\lambda \mu^{1/2+\gamma} + \lambda^2 \mu\} \end{aligned}$$

Zum Abschluss wird noch in (3.11) $\lambda = \frac{1}{2} \mu^{\gamma-1/2}$ gesetzt. Da $\gamma < \frac{1}{2}$ nach Voraussetzung des Lemmas, ist dieser Wert für genügend großes μ kleiner als $\frac{3}{2}$. Die Ungleichung (3.10) lässt sich also anwenden, und wir erhalten als finale Abschätzung

$$\mathbb{P} [S - \mu \geq \mu^{1/2+\gamma}] \leq \exp \left\{ -\frac{1}{4} \mu^{2\gamma} \right\}.$$

Eine analoge Abschätzung von $\mathbb{E} [e^{-\lambda(S-\mu)}]$ liefert

$$\mathbb{P} [\mu - S \geq \mu^{1/2+\gamma}] \leq \exp \left\{ -\frac{1}{4} \mu^{2\gamma} \right\},$$

und damit die Aussage des Lemmas. □

Zum Beweis der oberen Schranke von Satz 1 benötigen wir noch ein paar Resultate aus der diskreten Potentialtheorie. Dazu zunächst einige grundlegende Definitionen.

Definition 8. Seien $p(x, y)$ die Übergangswahrscheinlichkeiten einer Irrfahrt S_n in \mathbb{Z}^d , und sei $f : \mathbb{Z}^d \rightarrow \mathbb{R}$, dann wird der diskrete Laplace-Operator Δ definiert durch

$$\Delta f(x) = \left(\sum_{y \in \mathbb{Z}^d} p(x, y) f(y) \right) - f(x).$$

Bemerkung 8. Es gibt einen einfachen Zusammenhang zwischen dem Laplace-Operator und der Irrfahrt S_n , durch

$$\Delta f(x) = \mathbb{E}_x [f(S_1) - f(S_0)].$$

Definition 9. Eine Funktion $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ heißt *harmonisch* auf $A \subset \mathbb{Z}^d$, wenn

$$\Delta f(x) = 0$$

für alle $x \in A$ gilt.

Beispiel 1. Die Greensche Funktion $G(z) = G(0, z)$ ist harmonisch in $\mathbb{Z}^d \setminus \{0\}$. Dazu müssen wir zeigen, dass für alle $x \neq 0$

$$G(x) = \sum_{y \in \mathbb{Z}^d} p(x, y)G(y)$$

gilt. Für die rechte Seite gilt

$$\begin{aligned} \sum_{y \in \mathbb{Z}^d} p(x, y)G(y) &= \sum_{y \in \mathbb{Z}^d} p(x, y) \sum_{t=0}^{\infty} p^{(t)}(0, y) \\ &= \sum_{t=0}^{\infty} \sum_{y \in \mathbb{Z}^d} p(x, y) \cdot p^{(t)}(0, y) \\ &= \sum_{t=0}^{\infty} p^{(t+1)}(0, x). \end{aligned}$$

Für $x \neq 0$ ist $p^{(0)}(0, x) = 0$ und somit

$$\sum_{y \in \mathbb{Z}^d} p(x, y)G(y) = \sum_{t=0}^{\infty} p^{(t)}(0, x) = G(x).$$

Das folgende Lemma aus [7] zeigt eine Verbindung zwischen harmonischen Funktionen und Martingalen.

Lemma 7. Sei $f : \mathbb{Z}^d \rightarrow \mathbb{R}$ beschränkt und harmonisch auf der Menge $A \subset \mathbb{Z}^d$, und die Stoppzeit τ gegeben durch

$$\tau = \inf\{j \geq 0 : S_j \notin A\},$$

dann ist der Prozess $M_n = f(S_{n \wedge \tau})$ ein Martingal bezüglich der kanonischen Filtration \mathcal{F}_n von S_n .

Beweis. Sei $x = S_0$, und S' eine Irrfahrt mit den selben Übergangswahrscheinlichkeiten wie S und Startpunkt $S'_0(\omega) = S_n(\omega)$ für ein fixes $n \in \mathbb{N}$.

Sei $n \geq 1$, dann gilt aufgrund der Markoveigenschaft für den folgenden bedingten Erwartungswert

$$\begin{aligned} \mathbb{E}_x[f(S_{n+1})|\mathcal{F}_n] &= \mathbb{E}_x[f(S_n + X_{n+1})|\mathcal{F}_n] = \mathbb{E}_{S_n} \left[\mathbb{E}_x[f(S'_1)|\mathcal{F}_n] \right] \\ &= \mathbb{E}_{S_n} [f(S'_1)] = f(S_n) + \Delta f(S_n). \end{aligned} \quad (3.12)$$

Die letzten beiden Gleichungen gelten wegen der \mathcal{F}_n -Messbarkeit von S'_1 und Bemerkung 8.

Die Gleichung (3.12) ist auch für $n = 0$ erfüllt, da

$$\mathbb{E}_x[f(S_1)|\mathcal{F}_0] = \mathbb{E}_x[f(S_1)],$$

und mit Bemerkung 8

$$f(S_0) + \Delta f(S_0) = f(S_0) + \mathbb{E}_x[f(S_1)] - \mathbb{E}_x[f(S_0)] = \mathbb{E}_x[f(S_1)]$$

gilt.

Wir betrachten nun das Ereignis $B_n = [\tau > n]$. B_n ist \mathcal{F}_n -messbar, und für $\omega \in B_n^c$ gilt $M_{n+1}(\omega) = M_n(\omega)$.

$$\begin{aligned} \mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}[M_{n+1} \cdot \mathbb{1}_{B_n}|\mathcal{F}_n] + \mathbb{E}[M_{n+1} \cdot \mathbb{1}_{B_n^c}|\mathcal{F}_n] \\ &= \mathbb{E}[f(S_{n+1}) \cdot \mathbb{1}_{B_n}|\mathcal{F}_n] + \mathbb{E}[M_n \cdot \mathbb{1}_{B_n^c}|\mathcal{F}_n] \\ &= \mathbb{E}[f(S_{n+1})|\mathcal{F}_n] \mathbb{1}_{B_n} + M_n \cdot \mathbb{1}_{B_n^c} \end{aligned}$$

da B_n und M_n \mathcal{F}_n -messbar sind. Mit (3.12) folgt weiters

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] = (f(S_n) + \Delta f(S_n)) \mathbb{1}_{B_n} + M_n \cdot \mathbb{1}_{B_n^c}.$$

Nun ist aber die Funktion f harmonisch auf A . Damit gilt $\Delta f(S_n) = 0$ für das Ereignis B_n , und somit

$$\begin{aligned} \mathbb{E}[M_{n+1}|\mathcal{F}_n] &= f(S_n) \cdot \mathbb{1}_{B_n} + M_n \cdot \mathbb{1}_{B_n^c} \\ &= M_n. \end{aligned}$$

□

3.2 Die untere Schranke

In diesem Abschnitt beweisen wir den ersten Teil von Theorem 1. Wir zeigen hierzu die folgende zu (3.2) äquivalente Aussage.

Satz 3. Für $\varepsilon > 0$ existiert ein n_ε sodass gilt

$$\mathbb{P}[\mathcal{B}_{n(1-\varepsilon)} \subseteq A(\omega_d n^d(1+\varepsilon)), \forall n > n_\varepsilon] = 1.$$

Zur Vereinfachung der Notation geben wir zunächst folgende

Definition 10. Wir bezeichnen mit $E_z(n)$ das Ereignis, dass der Knoten z nach den ersten n Partikeln noch nicht zum Aggregationscluster gehört. Also

$$E_z(n) := [z \notin A(n)].$$

Bemerkung 9. Für jede Knotenmenge $B \subset \mathbb{Z}^d$ gilt

$$[B \not\subset A(n)] = \bigcup_{z \in B} E_z(n).$$

Im ersten Schritt des Beweises von Satz 3 verwenden wir das folgende bekannte

Lemma 8 (Lemma von Borel-Cantelli). Sei $\{A_n\}_{n=0}^\infty$ eine Folge von Ereignissen. Die Wahrscheinlichkeit dass unendlich viele Ereignisse A_n eintreffen ist 0 wenn gilt

$$\sum_{n=0}^{\infty} \mathbb{P}[A_n] < \infty.$$

Aufgrund des Lemmas von *Borel-Cantelli* folgt aus

$$\sum_{n \geq 0} \mathbb{P} [\mathcal{B}_{n(1-\varepsilon)} \not\subset A(\omega_d n^d(1+\varepsilon))] < \infty \quad (3.13)$$

die Aussage

$$\mathbb{P} [\mathcal{B}_{n(1-\varepsilon)} \not\subset A(\omega_d n^d(1+\varepsilon)) \text{ für unendlich viele } n] = 0.$$

Dies ist äquivalent zu Satz 3 und somit zum ersten Teil von Satz 1.

Andererseits gilt aufgrund von Bemerkung 9 für (3.13)

$$\begin{aligned} \sum_{n \geq 0} \mathbb{P} [\mathcal{B}_{n(1-\varepsilon)} \not\subset A(\omega_d n^d(1+\varepsilon))] &= \sum_{n \geq 0} \mathbb{P} \left[\bigcup_{z \in \mathcal{B}_{n(1-\varepsilon)}} E_z(\omega_d n^d(1+\varepsilon)) \right] \leq \\ &\leq \sum_{n \geq 0} \sum_{z \in \mathcal{B}_{n(1-\varepsilon)}} \mathbb{P} [E_z(\omega_d n^d(1+\varepsilon))]. \end{aligned}$$

Für die untere Schranke genügt es also die Aussage

$$\sum_{n \geq 0} \sum_{z \in \mathcal{B}_{n(1-\varepsilon)}} \mathbb{P} [E_z(\omega_d n^d(1+\varepsilon))] < \infty \quad (3.14)$$

zu zeigen. Dazu benötigen wir zunächst die folgenden Stoppzeiten.

Definition 11.

$$\begin{aligned} \sigma^i &= \min \{t \geq 0 : X^i(t) \notin A(i-1)\} \\ \tau_z^i &= \min \{t \geq 0 : X^i(t) = z\} \\ \tau_n^i &= \min \{t \geq 0 : X^i(t) \notin \mathcal{B}_n\} \end{aligned}$$

Bemerkung 10.

- σ^i bezeichnet die Anzahl der Schritte, die das i -te Partikel benötigt um den Cluster $A(i-1)$ zu verlassen.

3.2. DIE UNTERE SCHRANKE

- τ_z^i bezeichnet die Anzahl der Schritte, die das i -te Partikel benötigt um zum Knoten z zu gelangen.
- τ_n^i bezeichnet die Anzahl der Schritte, die das i -te Partikel benötigt um \mathcal{B}_n (die Kugel mit Radius n) zu verlassen.
- τ_z^i und τ_n^i hängen nicht vom Aggregationscluster ab und sind somit für alle i identisch verteilt.
- $E_z(n) = \bigcap_{i \leq n} [\sigma^i < \tau_z^i]$.

Für den Rest dieses Abschnittes sei nun n und $z \in \mathcal{B}_{n(1-\varepsilon)}$ beliebig, aber fix gewählt.

Definition 12. Wir definieren folgende Zufallsvariablen

$$\begin{aligned}
 N &= \sum_{i=0}^{\omega_d n^d(1+\varepsilon)} \mathbb{1}_{\{\tau_z^i < \sigma^i\}} \\
 M &= \sum_{i=0}^{\omega_d n^d(1+\varepsilon)} \mathbb{1}_{\{\tau_z^i < \tau_n^i\}} \\
 L &= \sum_{i=0}^{\omega_d n^d(1+\varepsilon)} \mathbb{1}_{\{\sigma^i \leq \tau_z^i < \tau_n^i\}}
 \end{aligned}$$

Bemerkung 11.

- N ist die Anzahl der Partikel, die z besuchen bevor sie den aktuellen Cluster $A(i-1)$ verlassen.
- M ist die Anzahl der Partikel, die z besuchen bevor sie \mathcal{B}_n , die Kugel mit Radius n verlassen.
- L ist die Anzahl der Partikel, die zuerst den Cluster $A(i-1)$ verlassen, dann z besuchen und erst danach \mathcal{B}_n verlassen.

Bemerkung 12. Für die Zufallsvariablen N , M und L gelten folgende Ungleichungen

$$\begin{aligned}
 M &\geq L \\
 N &\geq M - L.
 \end{aligned}$$

Die erste Ungleichung ist klar da $[\sigma^i \leq \tau_z^i < \tau_n^i] \subseteq [\tau_z^i < \tau_n^i]$ ist.

Die zweite Ungleichung sieht man wie folgt: Wenn ein Partikel i zur rechten Seite $M - L$ einen Beitrag liefert, muss $\tau_z^i < \tau_n^i$ gelten, aber $\sigma^i \leq \tau_z^i < \tau_n^i$ darf nicht erfüllt sein. Aus diesen beiden Voraussetzungen folgt dann aber dass $\tau_z^i < \sigma^i$ ist. Das heißt, das i -te Partikel liefert auch einen Beitrag zur Summe N .

Als nächstes schätzen wir die Wahrscheinlichkeiten aus (3.14) mit Hilfe von N , M und L ab. Für jedes beliebige $a > 0$ gilt Aufgrund der Ungleichungen aus Bemerkung 12

$$\begin{aligned} \mathbb{P}[E_z(\omega_d n^d(1 + \varepsilon))] &= \mathbb{P}[z \notin A(\omega_d n^d(1 + \varepsilon))] = \mathbb{P}[N = 0] \\ &\leq \mathbb{P}[M - L = 0] = 1 - \mathbb{P}[M - L > 0] \\ &\leq 1 - \mathbb{P}[M > a \wedge L < a] = \mathbb{P}[M \leq a \vee L \geq a] \\ &\leq \mathbb{P}[M \leq a] + \mathbb{P}[L \geq a]. \end{aligned} \tag{3.15}$$

Bemerkung 13. Unserer weiters Ziel ist es, eine Wahl für a zu finden, sodass der Ausdruck $\mathbb{P}[M \leq a] + \mathbb{P}[L \geq a]$ klein genug wird, um die Bedingung aus (3.14) zu erfüllen.

Da n und $z \in \mathcal{B}_{n(1-\varepsilon)}$ jeweils fix gewählt sind, sind die Summanden $\mathbb{1}_{\{\tau_z^i < \tau_n^i\}}$ von M vom Index i unabhängig und identisch verteilt. Somit ergibt sich für den Erwartungswert von M

$$\mathbb{E}[M] = \lfloor \omega_n n^d(1 + \varepsilon) \rfloor \mathbb{P}[\tau_z < \tau_n]. \tag{3.16}$$

Wir schreiben hier und im Folgenden, um die Notation zu vereinfachen, die Stoppzeiten ohne einen Index i , wenn ihre Verteilung nicht von dem speziellen Partikel abhängt.

Im Gegensatz zu den Summanden von M hängen die Summanden $\mathbb{1}_{\{\sigma^i \leq \tau_z^i < \tau_n^i\}}$ von L von der Gestalt der Cluster $A(i)$ ab, und sind somit nicht voneinander unabhängig. Allerdings trägt das Partikel X^i nur dann zu L bei, wenn $X^i(\sigma^i) \in \mathcal{B}_n$ ist. Also nur wenn es innerhalb der Kugel \mathcal{B}_n stoppt. Nun gibt es jedes $y \in \mathcal{B}_n$ höchstens ein i so dass $X^i(\sigma^i) = y$ ist. Das heißt

$$L \leq \tilde{L} = \sum_{y \in \mathcal{B}_n} \mathbb{1}_{\{\tau_z < \tau_n\}}^y.$$

Die Indikatorfunktionen $\mathbb{1}^y$ beziehen sich dabei auf in y startende Irrfahrten.

3.2. DIE UNTERE SCHRANKE

Die Summanden von \tilde{L} sind unabhängig vom Index i , und somit lässt sich der Erwartungswert von \tilde{L} berechnen:

$$\mathbb{E}[\tilde{L}] = \sum_{y \in \mathcal{B}_n} \mathbb{P}_y[\tau_z < \tau_n]. \quad (3.17)$$

Da $L \leq \tilde{L}$ gilt, folgt aus (3.15)

$$\mathbb{P}[E_z(\omega_d n^d(1 + \varepsilon))] \leq \mathbb{P}[\tilde{L} \geq a] + \mathbb{P}[M \leq a]. \quad (3.18)$$

Um die Ungleichung (3.14) zu zeigen genügt es also eine Konstante a zu finden, für die sowohl $\mathbb{P}[\tilde{L} \geq a]$ als auch $\mathbb{P}[M \leq a]$ exponentiell in n fallen. Das nächste Lemma liefert eine geeignete Wahl für die Konstante a .

Lemma 9. *Sei $\varepsilon > 0$, $n > n_0$ und $a = (1 + \frac{\varepsilon}{4})\mathbb{E}[\tilde{L}]$, dann gilt*

$$\begin{aligned} \mathbb{P}[\tilde{L} \geq a] &\leq e^{-cn} \\ \mathbb{P}[M \leq a] &\leq e^{-cn} \end{aligned}$$

$\forall z \in \mathcal{B}_{n(1-\varepsilon)}$ und einer geeigneten Konstante $c = c_d(\varepsilon) > 0$.

Beweis. Es gilt

$$G_n(y, z) = \mathbb{P}_y[\tau_z < \tau_n] G_n(z, z).$$

Daraus folgt mit (3.16), Lemma 5 und (3.17)

$$\begin{aligned} \mathbb{E}[M] &= \lfloor \omega_d n^d(1 + \varepsilon) \rfloor \mathbb{P}[\tau_z < \tau_n] = \lfloor \omega_d n^d(1 + \varepsilon) \rfloor \frac{G_n(0, z)}{G_n(z, z)} \geq \\ &\geq \left(1 + \frac{\varepsilon}{2}\right) \sum_{y \in \mathcal{B}_n} \frac{G_n(y, z)}{G_n(z, z)} = \left(1 + \frac{\varepsilon}{2}\right) \sum_{y \in \mathcal{B}_n} \mathbb{P}_y[\tau_z < \tau_n] = \\ &= \left(1 + \frac{\varepsilon}{2}\right) \mathbb{E}[\tilde{L}]. \end{aligned} \quad (3.19)$$

Für den Erwartungswert von \tilde{L} erhält man aus (3.17) und Proposition 3

$$\mathbb{E}[\tilde{L}] = \sum_{y \in \mathcal{B}_n} \mathbb{P}_y[\tau_z < \tau_n] = \sum_{y \in \mathcal{B}_n} \frac{G_n(y, z)}{G_n(z, z)} = \frac{\mathbb{E}_z[\tau_n]}{G_n(z, z)}. \quad (3.20)$$

Unter Anwendung von Lemma 3 und Proposition 1 erhält man daraus

$$\mathbb{E}[\tilde{L}] \geq \frac{n^2 - \|z\|^2}{G_{2n}(0,0)}$$

und mit den Abschätzungen aus Lemma 2(b) und Proposition 1

$$\mathbb{E}[\tilde{L}] \geq \beta'_d n^d \geq \beta_d n^3 \tag{3.21}$$

für eine geeignete Konstante β_d . Mit (3.19) erhalten wir die selbe untere Schranke für M , also

$$\mathbb{E}[M] \geq \beta_d n^3. \tag{3.22}$$

Nun sind sowohl M als auch \tilde{L} Summen von unabhängigen Indikatorzufallsvariablen, es ist somit möglich Lemma 6 anzuwenden, und man erhält

$$\begin{aligned} \mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}] + \mathbb{E}[\tilde{L}]^{1/2+\gamma}] &\leq 2 \exp\{-\tfrac{1}{4}\mathbb{E}[\tilde{L}]^{2\gamma}\} \\ \mathbb{P}[M \leq \mathbb{E}[M] - \mathbb{E}[M]^{1/2+\gamma}] &\leq 2 \exp\{-\tfrac{1}{4}\mathbb{E}[M]^{2\gamma}\} \end{aligned} \tag{3.23}$$

für $z \in \mathcal{B}_{n(1-\varepsilon)}$ und $n \geq n_\varepsilon$. Einsetzen von $\gamma = \frac{1}{3}$ ergibt für \tilde{L}

$$\mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}] + \mathbb{E}[\tilde{L}]^{5/6}] \leq 2 \exp\{-\tfrac{1}{4}\mathbb{E}[\tilde{L}]^{2/3}\}$$

und schließlich mit Verwendung der Abschätzung (3.21) für $\mathbb{E}[\tilde{L}]$

$$\mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}] + \mathbb{E}[\tilde{L}]^{5/6}] \leq 2 \exp\{-\tfrac{1}{4}\beta_d^{2/3} n^2\} \leq e^{-c_d n},$$

für eine geeignete Konstante c_d . Weiters ist

$$\begin{aligned} \mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}] + \mathbb{E}[\tilde{L}]^{5/6}] &= \mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}](1 + \mathbb{E}[\tilde{L}]^{-1/6})] \\ &\geq \mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}](1 + \tilde{\beta}_d n^{-1/2})] \\ &\geq \mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}](1 + \tfrac{\varepsilon}{4})], \end{aligned}$$

da nach (3.21) für jedes $\varepsilon > 0$ und n entsprechend groß gilt:

$$\frac{\varepsilon}{4} \geq \tilde{\beta}_d n^{-1/2} \geq \mathbb{E}[\tilde{L}]^{-1/6} \quad \text{mit } \tilde{\beta}_d = \beta_d^{-1/6}$$

3.2. DIE UNTERE SCHRANKE

Es bleibt nur noch die beiden letzten Ungleichungen zu kombinieren, und man erhält die Aussage des Lemmas für die Zufallsvariable \tilde{L}

$$\mathbb{P}[\tilde{L} \geq \mathbb{E}[\tilde{L}](1 + \frac{\varepsilon}{4})] \leq e^{-c_d n}.$$

Für die Zufallsvariable M folgt aus (3.23) mit $\gamma = \frac{1}{3}$

$$\mathbb{P}[M \leq \mathbb{E}[M](1 - \mathbb{E}[M]^{-1/6})] \leq e^{-c_d n}.$$

Andererseits gilt aufgrund der unteren Schranke (3.19) für $\mathbb{E}[M]$

$$\mathbb{P}[M \leq (1 + \frac{\varepsilon}{2})\mathbb{E}[\tilde{L}](1 - \mathbb{E}[M]^{-1/6})] \leq \mathbb{P}[M \leq \mathbb{E}[M](1 - \mathbb{E}[M]^{-1/6})].$$

Zum Beweis des Lemmas für M bleibt also noch die folgende Ungleichung zu zeigen

$$\left(1 + \frac{\varepsilon}{2}\right) \mathbb{E}[\tilde{L}](1 - \mathbb{E}[M]^{-1/6}) \geq \left(1 + \frac{\varepsilon}{4}\right) \mathbb{E}[\tilde{L}]. \quad (3.24)$$

Dies ist äquivalent zu

$$\frac{\varepsilon}{4} \geq \mathbb{E}[M]^{-1/6} \left(1 + \frac{\varepsilon}{2}\right)$$

Abschätzung (3.22) $\mathbb{E}[M] \geq \beta_d n^3$ liefert daraus die folgende hinreichende Bedingung für (3.24):

$$\frac{\varepsilon}{4} \geq \tilde{\beta}_d n^{-1/2} \left(1 + \frac{\varepsilon}{2}\right) \quad (3.25)$$

Hier ist wieder $\tilde{\beta}_d = \beta_d^{-1/6}$.

Die Ungleichung (3.25) gilt nun offensichtlich für n genügend groß. Damit ist (3.24) und somit auch die Aussage des Lemmas für M erfüllt. \square

Wir haben nun alle Voraussetzungen zum Beweis von Satz 3. Lemma 9 angewendet auf (3.18) liefert

$$\mathbb{P}[E_z(\omega_d n^d(1 + \varepsilon))] \leq 2e^{-cn}$$

für $n > n_0$ geeignet, und weiters

$$\sum_{n \geq n_0} \sum_{z \in \mathcal{B}_{n(1-\varepsilon)}} \mathbb{P}[E_z(\omega_d n^d(1 + \varepsilon))] \leq \sum_{n \geq n_0} 2\omega_d n^d e^{-cn} < \infty.$$

Damit ist (3.14) und somit Satz 3 bewiesen.

3.3 Die obere Schranke

Zum Beweis der oberen Schranke (3.3) von Satz 1 genügt es

$$\mathbb{P}[A(\omega_d n^d) \subset \mathcal{B}_{n(1+K\varepsilon^{1/d})} \text{ für } n \geq n_\varepsilon] \geq 1 - \varepsilon \quad (3.26)$$

für jedes $1 > \varepsilon > 0$ und ein $K = K_d < \infty$ zu zeigen.

Die untere Schranke (3.2) von Satz 1 besagt nun, dass

$$\mathbb{P}[\mathcal{B}_{n(1-\varepsilon)} \subset A(\omega_d n^d), \forall n \geq n_\varepsilon] \geq 1 - \varepsilon$$

gilt.

Das heißt der größte Teil der Elemente von $A(\omega_d n^d)$ liegt in der Kugel $\mathcal{B}_{n(1-\varepsilon)}$. Es können also nur noch wenige Partikel außerhalb liegen. Genauer

$$\mathbb{P}[|A(\omega_d n^d) \cap \mathcal{B}_n^c| \leq |A(\omega_d n^d)| - |\mathcal{B}_{n(1-\varepsilon)}|, \forall n \geq n_\varepsilon] \geq 1 - \varepsilon.$$

Für diese Schranke gilt die Abschätzung

$$\begin{aligned} |A(\omega_d n^d)| - |\mathcal{B}_{n(1-\varepsilon)}| &= \omega_d n^d - \omega_d n^d (1 - \varepsilon)^d = \omega_d n^d \left(1 - \sum_{i=0}^d \binom{d}{i} (-\varepsilon)^i \right) \\ &= \omega_d n^d \varepsilon \cdot \sum_{i=0}^{d-1} \binom{d}{i+1} (-\varepsilon)^i = \omega_d n^d \varepsilon \cdot \sum_{i=0}^{d-1} \frac{d}{i+1} \binom{d-1}{i} (-\varepsilon)^i \\ &\leq \omega_d n^d d \varepsilon (1 - \varepsilon)^{d-1} \leq K_d \varepsilon n^d. \end{aligned}$$

Es existiert also eine Konstante $K_d < \infty$, sodass für alle $1 > \varepsilon > 0$ gilt

$$\mathbb{P}[|A(\omega_d n^d) \cap \mathcal{B}_n^c| \leq K_d \varepsilon n^d, \forall n \geq n_\varepsilon] \geq 1 - \varepsilon. \quad (3.27)$$

Definition 13. Seien $A, B, C \subset \mathbb{Z}^d$, dann bezeichnet die Stoppzeit

$$H_A := \min\{t \geq 0 : X(t) \in A\}$$

die Zeit des ersten Eintreffens der Irrfahrt in der Menge A , und

$$h_z(B, C) := \mathbb{P}_z[H_B = H_{C \cap B}]$$

ist die Wahrscheinlichkeit, dass bei Start in z die Menge B zuerst in einem Punkt der Menge C getroffen wird.

Definition 14. Eine Menge der Form

$$\mathcal{S}_k := \{x \in \mathbb{Z}^d : k \leq \|x\| < k + 1\}$$

heißt *Schale* mit Radius $k \in \mathbb{N}_0$.

Bemerkung 14.

- Zwei aufeinanderfolgende Schalen sind ineinander verschachtelt. Das heißt $\forall k$ gelten die Relationen

$$\mathcal{B}_k \cap \mathcal{S}_k = \emptyset \quad \text{und} \quad \mathcal{B}_k \cup \mathcal{S}_k = \mathcal{B}_{k+1}.$$

- Schalen haben keine Lücken. Das heißt eine Irrfahrt kann die Kugel \mathcal{B}_k nicht verlassen, ohne die Schale \mathcal{S}_k zu betreten.

Die folgende Ungleichung aus [7] wird zum Beweis des zentralen Lemmas dieses Abschnitts benötigt.

Lemma 10 (Harnack Ungleichung). $\forall r < 1$ existiert eine Konstante $\mathcal{J}_r < \infty$, sodass für jede auf der Kugel B_n harmonische Funktion $h : \overline{B_n} \rightarrow [0, \infty)$ die folgende Bedingung erfüllt ist:

$$h(x) \leq \mathcal{J}_r h(y) \quad \text{für } |x|, |y| \leq rn.$$

Beweis. Siehe [7] Theorem 1.7.2. □

Lemma 11. Sei $T_k := \min\{t \geq 1 : X(t) \in \mathcal{S}_k\}$, dann existiert eine von der Dimension d abhängige Konstante $\mathcal{J} < \infty$, sodass für $j < k$, $\Delta = k - j$ und die Punkte $z \in \mathcal{S}_k$ und $y \in \mathcal{S}_j$ gilt

(a) $\mathbb{P}_z[T_j < T_k] \leq \mathcal{J} \Delta^{-1}$

(b) $h_y(\mathcal{S}_k, B) \leq \mathcal{J} |B| \Delta^{1-d}$ für eine Menge $B \subset \mathcal{S}_k$.

Beweis.

- (a) Zum Beweis des ersten Teils betrachten wir die Menge $A = \{x : j < \|x\| < k\}$, und die Stoppzeit $\tau = \min\{i \geq 0 : X(i) \notin A\}$. τ ist der Zeitpunkt in dem die Irrfahrt die Menge A zum ersten Mal verlässt.

Es gilt für jedes $z' \in A$

$$\mathbb{P}_z [T_j < T_k] \leq \mathbb{P}_{z'} [|X(\tau)| \leq j]. \quad (3.28)$$

Sei nun $M_i = G(X(i \wedge \tau))$. Die Funktion G ist harmonisch auf A , da $0 \notin A$, und somit ist nach Lemma 7 der Prozess M_i ein Martingal. Wir können das Optional-Stopping Theorem anwenden, und erhalten

$$\mathbb{E}_{z'} [M_\tau] = \mathbb{E}_{z'} [M_0] = G(z'). \quad (3.29)$$

Andererseits ist $[|X(\tau)| \leq j] \cup [|X(\tau)| \geq k] = \Omega$ eine disjunkte Zerlegung des Wahrscheinlichkeitsraumes, und somit erhalten wir

$$\mathbb{E}_{z'} [M_\tau] = \mathbb{E}_{z'} [M_\tau \cdot \mathbb{1}_{\{|X(\tau)| \leq j\}}] + \mathbb{E}_{z'} [M_\tau \cdot \mathbb{1}_{\{|X(\tau)| \geq k\}}].$$

Die beiden Faktoren in diesen Erwartungswerten sind nicht unabhängig, sie lassen sich allerdings mit Hilfe von bedingten Erwartungen aufspalten. Somit ergibt sich zusammen mit (3.29)

$$\begin{aligned} \mathbb{E}_{z'} [M_\tau] &= \mathbb{P}_{z'} [|X(\tau)| \leq j] \cdot \mathbb{E}_{z'} [G(X(\tau)) \mid |X(\tau)| \leq j] + \\ &\quad (1 - \mathbb{P}_{z'} [|X(\tau)| \leq j]) \cdot \mathbb{E}_{z'} [G(X(\tau)) \mid |X(\tau)| \geq k], \end{aligned}$$

und durch Auflösen dieser Gleichung weiters

$$\mathbb{P}_{z'} [|X(\tau)| \leq j] = \frac{G(z') - \mathbb{E}_{z'} [G(X(\tau)) \mid |X(\tau)| \geq k]}{\mathbb{E}_{z'} [G(X(\tau)) \mid |X(\tau)| \leq j] - \mathbb{E}_{z'} [G(X(\tau)) \mid |X(\tau)| \geq k]}.$$

Nun gilt aufgrund von Bemerkung 7

$$G(z) = a_d \|z\|^{2-d} + o(\|z\|^{1-d})$$

sowie

$$\begin{aligned}\mathbb{E}_{z'}[G(X(\tau)) \mid |X(\tau)| \leq j] &= a_d j^{2-d} + O(j^{1-d}) \\ \mathbb{E}_{z'}[G(X(\tau)) \mid |X(\tau)| \geq k] &= a_d k^{2-d} + O(k^{1-d}).\end{aligned}$$

Mit (3.28) folgt daraus

$$\mathbb{P}_z[T_j < T_k] \leq \frac{\|z'\|^{2-d} + o(\|z'\|^{1-d}) - j^{2-d} - O(j^{1-d})}{j^{2-d} - k^{2-d} + O(j^{1-d}) - O(k^{1-d})} \quad (3.30)$$

für alle z' , für die $j < \|z'\| < k$ gilt.

Wählen wir jetzt speziell $z' \in \mathcal{S}_{k-1}$, dann ergibt sich für eine Konstante \mathcal{C}_d aus (3.30)

$$\mathbb{P}_z[T_j < T_k] \leq \mathcal{C}_d \frac{k^{1-d}}{j^{2-d} - k^{2-d}}. \quad (3.31)$$

Um (3.31) zu zeigen, müssen wir die folgende Ungleichung überprüfen:

$$\frac{(k-1)^{2-d} + o((k-1)^{1-d}) - j^{2-d} - O(j^{1-d})}{j^{2-d} - k^{2-d} + O(j^{1-d}) - O(k^{1-d})} \leq \mathcal{C}_d \frac{k^{1-d}}{j^{2-d} - k^{2-d}}.$$

Dies ist äquivalent zu

$$(k-1)^{2-d} + o((k-1)^{1-d}) - j^{2-d} - O(j^{1-d}) \leq \mathcal{C}_d k^{1-d} \left(1 + \frac{O(j^{1-d}) - O(k^{1-d})}{j^{2-d} - k^{2-d}} \right).$$

Der Ausdruck in der Klammer auf der rechten Seite ist beschränkt, und somit impliziert

$$(k-1) \left(\frac{k-1}{k} \right)^{1-d} + \mathcal{C}' \left(\frac{k-1}{k} \right)^{1-d} - j \left(\frac{j}{k} \right)^{1-d} - \mathcal{C}'' \left(\frac{j}{k} \right)^{1-d} \leq \tilde{\mathcal{C}}_d$$

die Ungleichung (3.31) für geeignet gewählte Konstanten \mathcal{C}' , \mathcal{C}'' und $\tilde{\mathcal{C}}_d$. Diese Ungleichung vereinfacht sich zu

$$\left(\frac{k-1}{k} \right)^{1-d} (k-1 + \mathcal{C}') - \left(\frac{j}{k} \right)^{1-d} (j + \mathcal{C}'') \leq \tilde{\mathcal{C}}_d. \quad (3.32)$$

Als Funktion in j betrachtet, ist die linke Seite von (3.32) monoton wachsend.

Es gilt $j < k$ und somit genügt es

$$\left(\frac{k-1}{k}\right)^{1-d} (k-1 + \mathcal{C}') - \left(\frac{k-1}{k}\right)^{1-d} (k-1 + \mathcal{C}'') \leq \tilde{\mathcal{C}}_d$$

zu betrachten. Dies ist äquivalent zu

$$\left(\frac{k-1}{k}\right)^{1-d} (\mathcal{C}' + \mathcal{C}'') \leq \tilde{\mathcal{C}}_d,$$

und ist wegen $\lim_{k \rightarrow \infty} \left(\frac{k-1}{k}\right)^{1-d} = 1$ für große k erfüllt. Damit ist 3.31 gezeigt. Durch elementares Umformen erhält man daraus

$$\mathbb{P}_z[T_j < T_k] \leq \frac{\mathcal{C}_d}{k \left[\left(\frac{j}{k}\right)^{2-d} - 1 \right]} \leq \frac{\mathcal{C}_d}{\Delta \left[\left(\frac{k}{j}\right)^{d-2} - 1 \right]}.$$

Nun ist für den Fall $\Delta \geq j$ der Quotient $\frac{k}{j} \geq 2$, und damit

$$\mathbb{P}_z[T_j < T_k] \leq \mathcal{J}_d \Delta^{-1} \quad \text{für} \quad \mathcal{J}_d = \frac{\mathcal{C}_d}{2^{d-2} - 1}.$$

Für $\Delta < j$ zeigen wir, dass $\Delta \cdot \mathbb{P}_z[T_j < T_k]$ nach oben durch eine Konstante beschränkt ist. Wir betrachten dazu die Funktion

$$f_k(j) = \frac{k^{1-d}(k-j)}{j^{2-d} - k^{2-d}}.$$

Nach 3.31 gilt $\Delta \cdot \mathbb{P}_z[T_j < T_k] \leq f_k(j)$. Es lässt sich leicht sehen, dass $\frac{\partial}{\partial y} f_k(j) \geq 0$ für $\Delta < j$. Die Funktion $f_k(j)$ ist in diesem Bereich also monoton wachsend, und da $j < k$ und $j, k \in \mathbb{N}$ gilt, genügt es die Beschränktheit von $f_k(k-1)$ zu zeigen. Nun ist

$$f_k(k-1) = \left(\frac{(k-1)^{2-d}}{k^{1-d}} - k \right)^{-1},$$

und es gilt $\lim_{k \rightarrow \infty} f_k(k-1) = \frac{1}{d-2}$. Damit ist das Lemma auch für den Fall $\Delta < j$ gezeigt.

Dieser Beweis verwendet die Abschätzung der Greenfunktion aus Satz 1 und ist somit nur für Dimensionen $d \geq 3$ gültig. Für \mathbb{Z}^2 lässt sich der Beweis über eine zu (3.30) äquivalente Abschätzung mit Hilfe des Potentialkerns führen. Siehe

dazu [9] und Aufgabe 1.6.8 in [7].

- (b) Die Funktion $h_z(\bullet, \bullet)$ ist additiv in der zweiten Komponente, somit genügt es den Beweis für $B = \{z\}$, $z \in \mathcal{S}_k$ zu führen. Sei

$$\rho = \rho_{k,y} = \inf\{t \geq 1 : X(t) \in \mathcal{S}_k \cup \{y\}\},$$

dann gilt für die Wahrscheinlichkeit $h_y(\mathcal{S}_k, z)$, dass bei Start der Irrfahrt in y die Schale \mathcal{S}_k zuerst im Punkt z getroffen wird

$$h_y(\mathcal{S}_k, z) = \mathbb{P}_y[X(T_k) = z] = G_k(y, y) \cdot \mathbb{P}_y[X(\rho) = z]$$

und Aufgrund der Reversibilität von $X(t)$

$$h_y(\mathcal{S}_k, z) = G_k(y, y) \cdot \mathbb{P}_z[X(\rho) = y]. \quad (3.33)$$

Sei nun m der Radius der Schale mit genau gleichem Abstand von \mathcal{S}_j und \mathcal{S}_k , also $m = m_{k,\Delta} = \lfloor k - \frac{\Delta}{2} \rfloor$. Dann gilt $y \notin \mathcal{S}_m$ wenn Δ genügend groß ist. Wir betrachten nun die bedingte Wahrscheinlichkeit

$$\mathbb{P}_z[X(\rho) = y | T_m < T_k] = \frac{\mathbb{P}_z[X(\rho) = y, T_m < T_k]}{\mathbb{P}_z[T_m < T_k]}$$

Da auf dem Weg von $z \in \mathcal{S}_k$ nach $y \in \mathcal{S}_j$ auf jeden Fall \mathcal{S}_m durchquert wird, ist $[X(\rho) = y] \subset [T_m < T_k]$, und somit gilt

$$\begin{aligned} \mathbb{P}_z[X(\rho) = y] &= \mathbb{P}_z[T_m < T_k] \cdot \mathbb{P}_z[X(\rho) = y | T_m < T_k] \leq \\ &\leq \mathbb{P}_z[T_m < T_k] \cdot \sup_{x \in \mathcal{S}_m} \mathbb{P}_x[X(\rho) = y]. \end{aligned} \quad (3.34)$$

Aufgrund von Teil (a) des Lemmas und (3.33), (3.34) folgt

$$\begin{aligned} h_y(\mathcal{S}_k, z) &\leq \frac{\mathcal{J}}{\Delta_{m,k}} \cdot G_k(y, y) \cdot \sup_{x \in \mathcal{S}_m} \mathbb{P}_x[X(\rho) = y] = \\ &= \frac{\mathcal{J}}{\Delta_{m,k}} \cdot \sup_{x \in \mathcal{S}_m} G_k(y, y) \cdot \mathbb{P}_x[X(\rho) = y] = \\ &= \frac{\mathcal{J}}{\Delta_{m,k}} \cdot \sup_{x \in \mathcal{S}_m} G_k(x, y) \end{aligned}$$

mit $\Delta_{m,k} := k - m = \lceil \frac{\Delta}{2} \rceil$.

Somit genügt es zum Beweis des Lemmas folgende Abschätzung zu zeigen:

$$\sup_{x \in \mathcal{S}_m} G_k(x, y) \leq \mathcal{J} \Delta^{2-d} \quad (3.35)$$

Dazu betrachten wir die Kugel V um den Punkt y .

$$V = V_{y,\Delta} := \{w \in \mathbb{Z}^d \mid \|w - y\| < \frac{\Delta}{4}\}$$

Für $x \in \mathcal{S}_m$ und Δ genügend groß ist die Funktion

$$g_x(w) := G_k(x, w) = \mathbb{E}_x \left[\sum_{t=0}^{\tau_k-1} \mathbb{1}_{\{X(t)=w\}} \right]$$

harmonisch in einer Menge $A := \{w \in \mathbb{Z}^d : \|w - y\| < \frac{3\Delta}{8}\} \supset V_{y,\Delta}$.

Zur Überprüfung der Harmonie ist $g_x(w) = \sum_{z \in \mathbb{Z}^d} p(w, z) \cdot g_x(z)$, für $w \in A$, zu zeigen.

$$\begin{aligned} \sum_{z \in \mathbb{Z}^d} p(w, z) \cdot g_x(z) &= \sum_{z \in \mathbb{Z}^d} p(w, z) \cdot \mathbb{E}_x \left[\sum_{t=0}^{\tau_k-1} \mathbb{1}_{\{X(t)=z\}} \right] \\ &= \sum_{z \in \mathbb{Z}^d} p(w, z) \cdot \sum_{i=0}^{\infty} \mathbb{P}_x[\tau_k = i] \cdot \sum_{t=0}^{i-1} \mathbb{P}_x[X(t) = z \mid \tau_k = i] \\ &= \sum_{i=0}^{\infty} \mathbb{P}_x[\tau_k = i] \cdot \sum_{t=0}^{i-1} \sum_{z \in \mathbb{Z}^d} p(z, w) \cdot \mathbb{P}_x[X(t) = z \mid \tau_k = i] \\ &= \sum_{i=0}^{\infty} \mathbb{P}_x[\tau_k = i] \cdot \sum_{t=0}^{i-1} \mathbb{P}_x[X(t+1) = w \mid \tau_k = i] \end{aligned}$$

Nun ist $\mathbb{P}_x[X(i) = w \mid \tau_k = i] = 0$, da $w \in \mathcal{B}_k$ ist. Weiters gilt $\mathbb{P}_x[X(0) = w] = 0$, da $x \notin A$ und $w \in A$. Somit ergibt sich

$$\sum_{z \in \mathbb{Z}^d} p(w, z) \cdot g_x(z) = \sum_{i=0}^{\infty} \mathbb{P}_x[\tau_k = i] \cdot \sum_{t=0}^{i-1} \mathbb{P}_x[X(t) = w \mid \tau_k = i] = g_x(w).$$

Die Anwendung der *Harnack-Ungleichung* (Lemma 10) auf $g_x(w)$ liefert die Ab-

schätzung

$$G_k(x, y) \leq \mathcal{J} G_k(x, w)$$

für alle $w \in V$, $x \in \mathcal{S}_m$ und eine Konstante $\mathcal{J} < \infty$. Summation über alle Elemente von V ergibt

$$|V| \cdot G_k(x, y) \leq \mathcal{J} \sum_{w \in V} G_k(x, w)$$

und mit $\left(\frac{\Delta}{4}\right)^d < |V|$

$$G_k(x, y) \leq \mathcal{J}' \Delta^{-d} \sum_{w \in V} G_k(x, w).$$

Mittels Supremumbildung ergibt sich daraus eine Abschätzung für (3.35).

$$\begin{aligned} \sup_{x \in \mathcal{S}_m} G_k(x, y) &\leq \mathcal{J}' \Delta^{-d} \cdot \sup_{x \in \mathcal{S}_m} \sum_{w \in V} G_k(x, w) \\ &= \mathcal{J}' \Delta^{-d} \cdot \sup_{x \in \mathcal{S}_m} \mathbb{E}_x[Y] \end{aligned}$$

mit der Zufallsvariablen

$$Y = \sum_{t=0}^{\tau_k-1} \mathbb{1}_{\{X(t) \in V\}}.$$

$\mathbb{E}_x[Y]$ ist die Erwartete Anzahl der Besuche in der Menge V bevor die Irrfahrt die Kugel \mathcal{B}_k verlässt. Ist dieser Erwartungswert positiv, dann gibt es ein $w \in V$ mit $X(t) = w$ für ein $t < \tau_k$. Und es gilt

$$\mathbb{E}_x[Y] \leq \sup_{w \in V} \mathbb{E}_w[Y],$$

für alle $x \in \mathcal{S}_m$.

Zum Beweis von (3.35) bleibt somit noch

$$\sup_{w \in V} \mathbb{E}_w[Y] \leq \mathcal{J} \Delta^2 \tag{3.36}$$

zu zeigen.

Aus dem Zentralen Grenzwertsatz folgt, dass eine von k unabhängige Konstante

$\Phi = \Phi_{d,\Delta} > 0$ existiert, so dass für alle $w \in V$ gilt

$$\mathbb{P}_w[\tau_k \leq \Delta^2] \geq \Phi. \quad (3.37)$$

Daraus ergibt sich folgende Abschätzung für die Anzahl der Besuche in der Menge V . Für ganzzahlige $l \geq 0$ gilt

$$\sup_{w \in V} \mathbb{P}_w[Y > l\Delta^2] \leq (1 - \Phi)^l. \quad (3.38)$$

Der Beweis von (3.38) erfolgt durch vollständige Induktion.

Für $l = 0$ ist die Aussage offensichtlich. Für $l = 1$ gilt

$$\mathbb{P}_w[Y > \Delta^2] \leq 1 - \mathbb{P}_w[\tau_k \leq \Delta^2] \leq 1 - \Phi,$$

da die Irrfahrt V nur dann öfters als Δ^2 -mal Besuchen kann, wenn sie nach Δ^2 Schritten die Kugel B_k noch nicht verlassen hat.

Für den Induktionsschritt betrachten wir den Zeitpunkt k des $(l\Delta^2 + 1)$ -ten Besuches in V . Also $k \in \mathbb{N}$ mit $X(k) \in V$ und

$$|\{0 \leq i \leq k : X(i) \in V\}| = l\Delta^2 + 1.$$

So ein k existiert wenn $Y > l\Delta^2$ ist. Da sich die Irrfahrt zum Zeitpunkt k irgendwo in der Menge V befindet, lässt sich die gesuchte Wahrscheinlichkeit aufspalten:

$$\begin{aligned} \mathbb{P}_w[Y > (l+1)\Delta^2] &= \mathbb{P}_w[Y > l\Delta^2] \cdot \sup_{w \in V} \mathbb{P}_w[Y > \Delta^2] \\ &\leq (1 - \Phi)^{l+1}. \end{aligned}$$

Diese Abschätzung gilt $\forall w \in V$, und bleibt somit auch bei Bildung des Supremums gültig. Damit ist (3.38) gezeigt.

Zum Abschluss verwenden wir noch (3.38) zur Abschätzung von $\sup_{w \in V} \mathbb{E}_w[Y]$.

$$\begin{aligned}
 \sup_{w \in V} \mathbb{E}_w[Y] &= \sup_{w \in V} \sum_{i=1}^{\infty} i \cdot \mathbb{P}_w[Y = i] \\
 &= \sup_{w \in V} \sum_{l=1}^{\infty} \sum_{j=1}^{\Delta^2} ((l-1)\Delta^2 + j) \cdot \mathbb{P}_w[Y = (l-1)\Delta^2 + j] \\
 &\leq \sup_{w \in V} \sum_{l=1}^{\infty} l\Delta^2 \sum_{j=1}^{\Delta^2} \mathbb{P}_w[Y = (l-1)\Delta^2 + j] \\
 &\leq \Delta^2 \sup_{w \in V} \sum_{l=1}^{\infty} l \cdot \mathbb{P}_w[(l-1)\Delta^2 < Y \leq l\Delta^2] \\
 &\leq \Delta^2 \sum_{l=1}^{\infty} l \cdot \sup_{w \in V} \mathbb{P}_w[Y > (l-1)\Delta^2] \\
 &\leq \Delta^2 \sum_{l=1}^{\infty} l(1 - \Phi)^{l-1} = \frac{\Delta^2}{\Phi^2} = \mathcal{J}\Delta^2
 \end{aligned}$$

Dies beweist (3.36) und somit (3.35). □

Mit Hilfe von Lemma 11 können wir nun abschätzen, wie sich die, außerhalb von $\mathcal{B}_{n(1-\varepsilon)}$ liegenden, Elemente von $A(\omega_d n^d)$ in die einzelnen Schalen verteilen. Zur besseren Übersicht geben wir allerdings zunächst die folgende Definition.

Definition 15. Sei für $k \geq 1$ und $k_0 = \lfloor n(1 + \varepsilon^{\frac{1}{d}}) \rfloor + 1$. Die Zufallsvariable

$$Z_k(l) = |\tilde{A}(l) \cap \mathcal{S}_{k_0+k}|$$

bezeichnet die Anzahl der Partikel in der Schale \mathcal{S}_{k_0+k} , nachdem l Partikel die Kugel \mathcal{B}_n verlassen haben.

Damit die Aggregationscluster Kugelgestalt annehmen können, müssen die $K_k(l)$, mit zunehmendem Schalenradius k , sehr schnell gegen 0 streben. Das nächste Lemma liefert dafür eine hinreichend genaue Abschätzung.

Lemma 12. Für eine von der Dimension d abhängigen Konstante $\mathcal{J} < \infty$ gilt

$$\mathbb{E}[Z_k(j)] \leq n^{(1-k)(d-1)} \left[\mathcal{J} \frac{j}{k} \varepsilon^{\frac{1-d}{d}} \right]^k.$$

Beweis. Wir definieren die Stoppzeit $\tau_n^l := \min\{t \geq 1 : Y^l(t) \notin \mathcal{B}_n\}$ und $x_n^l := Y^l(\tau_n^l)$, den Punkt in dem das l -te Partikel die Kugel \mathcal{B}_n verlässt.

Die Wahrscheinlichkeit, dass Y^{l+1} in der Schale \mathcal{S}_{k_0+k} stoppt ist gegeben durch

$$h_{x_n^{l+1}}(\tilde{A}^c(l), \mathcal{S}_{k_0+k}),$$

die Wahrscheinlichkeit, dass das Komplement des Clusters zuerst in der Schale \mathcal{S}_{k_0+k} getroffen wird.

Der Erwartungswert von $Z_k(l+1)$ lässt sich damit aufspalten:

$$\mathbb{E}[Z_k(l+1)] = \mathbb{E}[Z_k(l)] + \mathbb{E}[h_{x_n^{l+1}}(\tilde{A}^c(l), \mathcal{S}_{k_0+k})]. \quad (3.39)$$

Nun ist aber die Wahrscheinlichkeit dass sich Y^{l+1} in der Schale \mathcal{S}_{k_0+k} anlagert kleiner als die Wahrscheinlichkeit dass Y^{l+1} beim ersten Eintreffen in der Schale \mathcal{S}_{k_0+k-1} nicht stoppt. Also

$$\begin{aligned} \mathbb{E}[h_{x_n^{l+1}}(\tilde{A}^c(l), \mathcal{S}_{k_0+k})] &\leq \mathbb{E}[h_{x_n^{l+1}}(\mathcal{S}_{k_0+k-1}, \tilde{A}(l))] \\ &\leq \max_{y \in \mathcal{S}_n} \mathbb{E}[h_y(\mathcal{S}_{k_0+k-1}, \tilde{A}(l))]. \end{aligned} \quad (3.40)$$

Auf $h_y(\mathcal{S}_{k_0+k-1}, \tilde{A}(l))$ lässt sich nun Lemma 11(b) anwenden und man erhält für $y \in \mathcal{S}_n$

$$\begin{aligned} h_y(\mathcal{S}_{k_0+k-1}, \tilde{A}(l)) &\leq \mathcal{J} \cdot |\mathcal{S}_{k_0+k-1} \cap \tilde{A}(l)| \cdot \Delta^{1-d} \\ &\leq \mathcal{J} \cdot Z_{k-1}(l) \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \end{aligned} \quad (3.41)$$

wegen $\Delta = k_0 + k - 1 - n = \lfloor n(1 + \varepsilon^{\frac{1}{d}}) \rfloor + k - n \geq n\varepsilon^{\frac{1}{d}} + k - 1 \geq n\varepsilon^{\frac{1}{d}}$.

3.3. DIE OBERE SCHRANKE

Einsetzen von (3.40) und (3.41) in (3.39) liefert

$$\mathbb{E}[Z_k(l+1)] - \mathbb{E}[Z_k(l)] \leq \mathcal{J} \cdot \mathbb{E}[Z_{k-1}(l)] \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1}.$$

Daraus erhalten wir durch Summation und dem Auflösen der sich ergebenden Teleskopsumme

$$\mathbb{E}[Z_k(j)] \leq \mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \cdot \sum_{l=1}^{j-1} \mathbb{E}[Z_{k-1}(l)]. \quad (3.42)$$

Mittels vollständiger Induktion lässt sich mit Hilfe von (3.42) nun die Ungleichung

$$\mathbb{E}[Z_k(j)] \leq \frac{j^k}{k!} \left[\mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \right]^{k-1} \quad (3.43)$$

zeigen.

Die Induktionsbasis $\mathbb{E}[Z_1(j)] \leq j$ ist offensichtlich. Für den Induktionsschritt führt (3.42)

$$\mathbb{E}[Z_k(j)] \leq \mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \cdot \sum_{l=1}^{j-1} \mathbb{E}[Z_{k-1}(l)]$$

einsetzen der Induktionsannahme

$$\leq \mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \cdot \left[\mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \right]^{k-2} \sum_{l=1}^{j-1} \frac{l^{k-1}}{(k-1)!}$$

und die Verwendung der elementaren Ungleichung $\sum_{l=1}^{j-1} l^{k-1} \leq \frac{j^k}{k}$

$$\leq \frac{j^k}{k!} \left[\mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \right]^{k-1}$$

zum gewünschten Ergebnis.

Es bleibt nur noch $k! \geq k^k e^{-k}$ in (3.43) einzusetzen

$$\begin{aligned} \mathbb{E}[Z_k(j)] &\leq \left(\frac{j}{k}\right)^k e^k \left[\mathcal{J} \left(\frac{1}{n\varepsilon^{1/d}} \right)^{d-1} \right]^{k-1} \\ &= n^{(1-d)(k-1)} \mathcal{J}^{k-1} e^k \left(\frac{j}{k}\right)^k \left(\varepsilon^{\frac{1-d}{d}}\right)^{k-1} \\ &\leq n^{(1-d)(k-1)} \left[\tilde{\mathcal{J}} \frac{j}{k} \varepsilon^{\frac{1-d}{d}} \right]^k, \end{aligned}$$

und wir erhalten die gesuchte Abschätzung. \square

Damit haben wir nun alle Voraussetzungen zum Beweis von (3.26). Wir bezeichnen mit F das Ereignis aus (3.27), also

$$F = [A(\omega_d n^d) \cap \mathcal{B}_n^c < K_d \varepsilon n^d].$$

Wenn die Schale $\mathcal{S}_{n(1+K\varepsilon^{1/d})}$ nicht leer ist, dann ist der aktuelle Cluster $A(t)$ keine Teilmenge der Kugel $\mathcal{B}_{n(1+K\varepsilon^{1/d})}$.

$$\begin{aligned} \mathbb{P}[A(\omega_d n^d) \not\subset \mathcal{B}_{n(1+K\varepsilon^{1/d})}, F] &\leq \mathbb{P}[Z_{k'}(\lfloor K_d \varepsilon n^d \rfloor) \geq 1] \\ &\leq \mathbb{E}[Z_{k'}(\lfloor K_d \varepsilon n^d \rfloor)] \end{aligned}$$

für $k' = \lceil (K-1)\varepsilon^{1/d}n - 1 \rceil$.

Setzen wir nun $k = k'$ und $j = \lfloor K_d \varepsilon n^d \rfloor$ in Lemma 12 ein, so ergibt sich

$$\mathbb{E}[Z_{k'}(\lfloor K_d \varepsilon n^d \rfloor)] \leq n^{d-1} \left[\mathcal{J} \frac{\lfloor K_d \varepsilon n^d \rfloor}{\lceil (K-1)\varepsilon^{1/d}n - 1 \rceil} \cdot \varepsilon^{\frac{1-d}{d}} \cdot n^{1-d} \right]^{k'}$$

In der Klammer heben sich nun die Potenzen von ε und n auf, und für eine neue Konstante \mathcal{J} gilt

$$\mathbb{P}[A(\omega_d n^d) \not\subset \mathcal{B}_{n(1+K\varepsilon^{1/d})}, F] \leq n^{d-1} \left(\frac{\mathcal{J}}{K} \right)^{k'} \quad (3.44)$$

Wählt man nun $K > \mathcal{J}$ und $n > n_0$ dann geht die rechte Seite von (3.44) exponentiell gegen 0. Also existiert ein $q < 1$ sodass

$$\sum_{n \geq n_0} \mathbb{P}[A(\omega_d n^d) \not\subset \mathcal{B}_{n(1+K\varepsilon^{1/d})}, F] \leq \sum_{n \geq n_0} q^n < \infty.$$

Mit dem Lemma von Borel-Cantelli folgt daraus

$$\mathbb{P}[\limsup_{n \geq n_0} [A(\omega_d n^d) \not\subset \mathcal{B}_{n(1+K\varepsilon^{1/d})}, F]] = 0$$

3.4 Weitere Ergebnisse

Da nun die Grenzstruktur bekannt ist, stellt sich die Frage nach der Geschwindigkeit mit der sich der Aggregationscluster $A(\omega_d n^d)$ den Kugeln \mathcal{B}_n annähert.

Dazu gibt es theoretische Resultate von Lawler [10] und Blachère [3]. Moore und Machta liefern in [16] für den Fall $d = 2$ ein numerisches Ergebnis.

Definition 16. Die maximalen inneren bzw. äußeren Abweichungen $\delta_I(n)$, bzw. $\delta_O(n)$ des Aggregationsclusters A von der Kugel \mathcal{B}_n sind gegeben durch

$$\begin{aligned} \delta_I(n) &:= n - \inf\{|z| : z \notin A(\omega_d n^d)\} \\ \delta_O(n) &:= \sup\{|z| : z \in A(\omega_d n^d)\} - n. \end{aligned}$$

Die ersten Abschätzungen der Abweichungen $\delta_I(n)$ und $\delta_O(n)$ lieferte Lawler in [10].

Satz 4 (Lawler). *Für $d \geq 2$ gilt fast sicher*

$$(a) \quad \lim_{n \rightarrow \infty} \frac{\delta_I(n)}{n^{\frac{1}{3}} \log^2 n} = 0$$

$$(b) \quad \lim_{n \rightarrow \infty} \frac{\delta_O(n)}{n^{\frac{1}{3}} \log^4 n} = 0.$$

Dieses Ergebnis wurde später von Blachère [3] verbessert, indem er zeigte dass die Abweichungen höchstens logarithmisch im Durchmesser des Clusters sind.

Satz 5 (Blachère). *Es gibt eine von der Dimension d abhängige Konstante \mathcal{C}_d , sodass fast sicher gilt*

$$(a) \quad \delta_I(n) \sim \begin{cases} \mathcal{C}_d \log^{1/2} n & \text{für } d \geq 3 \\ \mathcal{C}_2 \log n \log(\log n)^{1/2} & \text{für } d = 2 \end{cases}$$

$$(b) \quad \delta_O(n) \sim \begin{cases} \mathcal{C}_d \log^{3/2} n & \text{für } d \geq 4 \\ \mathcal{C}_3 \log^{5/3} n & \text{für } d = 3 \\ \mathcal{C}_2 \log^2 n & \text{für } d = 2. \end{cases}$$

Kapitel 4

Internal Diffusion-Limited Aggregation als Computing Modell

In diesem Kapitel beschäftigen wir uns mit internal Diffusion-Limited Aggregation aus dem Blickwinkel der theoretischen Informatik. Das Konzept der Berechenbarkeit ist sehr robust bezüglich der konkret gewählten „Maschine“ auf der die Berechnungen durchgeführt werden. Sehr unterschiedliche Berechenbarkeitsmodelle wie Turingmaschinen oder der λ -Kalkül haben die Fähigkeit, exakt die selbe Klasse von Funktionen zu berechnen. Dies ist die Aussage der berühmten Church-Turing These.

Wir definieren uns ein „Computer“-Modell auf Basis von iDLA, und untersuchen seine Eigenschaften. Um dies zu bewerkstelligen benötigen wir zunächst eine Reihe von Definitionen. Einführungen in die Theorie der Berechenbarkeit sowie der Komplexitätstheorie finden sich in [19] und [18].

Definition 17. Ein Bool'scher Schaltkreis ist ein kreisfreies Netzwerk von logischen Gattern, die durch „Drähte“ verbunden werden. Es gibt drei Typen von Gattern: **AND**, **OR** mit jeweils zwei Eingängen, sowie **NOT**-Gatter. Die „Drähte“ leiten Bool'sche Werte zwischen den einzelnen Gattern weiter. Die Gatter sind kleine Prozessoren, die die Funktionen \vee , \wedge sowie \neg berechnen. Die Eingänge des Schaltkreises bezeichnen wir mit x_1, \dots, x_n . Weiters wird ein Gatter als Ausgang des Schaltkreises gewählt.

Ein Bool'scher Schaltkreis C berechnet eine Boolsche Funktion $f_C : \{0, 1\}^n \rightarrow \{0, 1\}$.

Bemerkung 15. Wir betrachten später Schaltkreise mit k Ausgangs-Gattern. Diese berechnen Funktionen mit Bildbereich $\{0, 1\}^k$.

Definition 18. Die Größe eines Bool'schen Schaltkreises ist definiert als die Anzahl der Gatter im Schaltkreis. Die Länge des längsten Weges von einem Eingang zum Ausgangsgatter, bezeichnen wir als die Tiefe des Schaltkreises.

Definition 19. Eine Menge von Schaltkreisen $\mathcal{C} = \{C_n\}_{n \geq 0}$, wobei C_n jeweils genau n Eingänge besitzt, heißt Schaltkreis Familie. \mathcal{C} akzeptiert eine Sprache $A \subseteq \{0, 1\}^*$ genau dann, wenn für jeden binären String w der Länge n

$$w \in A \iff f_{C_n}(w) = 1$$

gilt.

Definition 20. Die Komplexitätsklasse \mathbf{P} ist definiert als die Klasse von Sprachen, die von einer Familie \mathcal{C} von Bool'schen Schaltkreisen akzeptiert werden, wobei die Größe der Schaltkreise C_n polynomiell in n , der Größe der Eingabedaten wächst.

Bemerkung 16. Diese Definition der Klasse \mathbf{P} ist äquivalent zur klassischen Definition mit Hilfe von Turingmaschinen mit polynomieller Zeitkomplexität. Für einen Beweis siehe zum Beispiel [19] Theorem 9.25.

Eine Teilklasse von \mathbf{P} ist die Klasse \mathbf{NC} , die Klasse der effizient parallelisierbaren Probleme. Eine Sprache ist in \mathbf{NC} wenn sie durch eine Familie von Bool'schen Schaltkreisen¹ mit einer polynomiellen Anzahl von Gattern sowie einer polylogarithmischen² Tiefe berechnet werden kann.

Definition 21 (L-Reduzierbarkeit). Seien A, B Sprachen aus \mathbf{P} , dann heißt A \mathbf{L} -reduzierbar auf B - geschrieben $A \preceq_{\mathbf{L}} B$ - wenn eine Funktion $\psi \in \mathbf{L}$ existiert, sodass für alle w

$$w \in A \iff \psi(w) \in B$$

gilt. \mathbf{L} ist dabei die Klasse, der unter Verwendung von $O(\log|w|)$ Speicher berechenbaren Funktionen.

¹Genauer wird eine *uniforme* Klasse von Schaltkreisen benötigt. Für Details siehe [18].

²Eine Funktion $f(n)$ heißt polylogarithmisch wenn $f(n) = o(\log^k(n))$ für eine Konstante k . Also wenn $\lim_{n \rightarrow \infty} \frac{f(n)}{\log^k(n)} = 0$ gilt.

Definition 22. Eine Sprache $p \in \mathbf{P}$ heißt \mathbf{P} -vollständig, wenn jede Sprache q aus der Klasse \mathbf{P} \mathbf{L} -reduzierbar auf p ist, das heißt

$$q \preceq_{\mathbf{L}} p \quad \text{für alle } q \in \mathbf{P}.$$

Definition 23 (Circuit-Value Problem). Gegeben sei die Beschreibung eines Bool'schen Schaltkreises, sowie die an den Eingängen des Schaltkreises anliegenden Wahrheitswerte. Gesucht sind die Wahrheitswerte an den Ausgängen des Schaltkreises.

Bemerkung 17. Das Circuit-Value Problem ist \mathbf{P} -vollständig.

Es ist nicht bekannt ob \mathbf{NC} eine echte Teilmenge der Komplexitätsklasse \mathbf{P} ist. Analog zum $\mathbf{P}=\mathbf{NP}$ Problem würde es genügen, die effiziente Parallelisierbarkeit eines einzelnen \mathbf{P} -vollständigen Problems zu zeigen oder zu widerlegen, um diese Frage zu beantworten. Es wird allerdings vermutet, dass \mathbf{P} -vollständige Probleme inhärent sequentiell sind. Das heißt, selbst mittels einer parallelen Berechnung unter Verwendung von polynomiell vielen Prozessoren ließe sich die Lösung eines \mathbf{P} -vollständigen Problems nicht wesentlich beschleunigen.

Das Problem der Vorhersage einer Reihe physikalischer Modelle, wenn die benötigten Zufallsbits mit den Eingabedaten mitgeliefert werden, ist \mathbf{P} -vollständig auf \mathbb{Z}^d mit $d \geq 3$. Darunter fällt zum Beispiel das Ising-Modell [15], Zelluläre Automaten und auch Diffusion-Limited Aggregation.

Da die Struktur der iDLA-Cluster im Vergleich zu externer DLA viel einfacher ist, liegt die Vermutung nahe, dass die Vorhersage des Verhaltens von iDLA in eine kleinere Komplexitätsklasse als \mathbf{P} fällt. Diese Frage wurde von Moore und Machta [16] beantwortet. Dazu benötigen wir zunächst die passende Komplexitätsklasse.

4.1 Die Klasse CC

Ein Komparator-Schaltkreis ist ein spezieller Bool'scher Schaltkreis, der ausschließlich aus sogenannten Komparator-Gattern besteht, und für den außerdem keine Verzweigungen in der Verdrahtung erlaubt sind. Das heißt jeder Ausgang eines Gatters wird

mit höchstens einem Eingang eines Gatters aus der nächsten Schicht des Schaltkreises verbunden.

Ein Komparator-Gatter besitzt zwei Eingänge a und b , sowie zwei Ausgänge. Die Logikwerte der Ausgänge sind $a \vee b$ und $a \wedge b$.

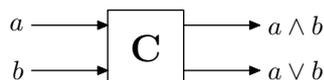


Abbildung 4.1: Komparator-Gatter

Die Komplexitätsklasse \mathbf{CC} ist definiert als die Menge aller Algorithmen, die von einem Komparator-Schaltkreis mit $O(n^a)$ Gattern berechnet werden können. n bezeichnet dabei wieder die Größe des Inputs.

Da Komparator-Schaltkreise aufgrund ihrer eingeschränkten Verdrahtungsmöglichkeiten keine Information kopieren können, wird vermutet, dass das Circuit-Value Problem in der Klasse \mathbf{CC} nicht \mathbf{P} -vollständig ist.

4.2 iDLA Vorhersage

Wir definieren nun das Problem der Vorhersage von iDLA-Clustern. Damit das Problem deterministisch bleibt, folgen die Partikel hier nicht Irrfahrten, sondern es werden die Schrittfolgen der einzelnen Teilchen in den Eingabedaten des Problems spezifiziert. Dies ist keine Einschränkung, da man einfach vor Beginn der Auswertung des Problems eine genügend große Anzahl von Zufallszahlen generieren kann, die dann als Eingabedaten kodiert werden.

Definition 24 (iDLA-Vorhersage). Gegeben sei eine Knotenmenge S , sowie eine Familie von Partikeln \mathcal{I} . Für jeden Zeitpunkt $0 \leq t < T$ ist als Input eine Folge von Übergängen (t, i, s) mit $i \in \mathcal{I}$ und $s \in S$ gegeben. Das Trippel (t, i, s) bedeutet, dass zum Zeitpunkt t das Partikel i zum Knoten s springt und dabei s als besetzt markiert. Dieser Übergang wird nur unter der Voraussetzung durchgeführt, dass das Partikel i zu diesem Zeitpunkt noch aktiv ist. Aktiv ist ein Partikel dann, wenn es zu jedem Zeitpunkt $t' < t$ ausschließlich bereits besetzte Knoten besucht hat. Zum Zeitpunkt $t = 0$ sind alle Partikel aktiv, und alle Knoten noch unbesetzt. Damit die Übergänge

wohldefiniert sind, darf kein Partikel zum selben Zeitpunkt an zwei verschiedene Orte geschickt werden. Das heißt es dürfen im Input keine zwei Übergangstrippel der Form (t, i, s) und (t, i, s') mit $s \neq s'$ enthalten sein.

Das Problem der iDLA-Vorhersage besteht nun darin, aus diesen Eingabedaten die besetzten Knoten, sowie die noch aktiven Partikel, vorherzusagen.

Bemerkung 18. Diese Definition ist allgemeiner als das bisher verwendete Modell, da zu jedem Zeitpunkt mehr als ein aktives Partikel erlaubt ist. Es wird auch keine Einschränkung bezüglich der zugrunde liegenden Graphenstruktur und der verwendeten Irrfahrt gemacht. So sind nichtlokale Sprünge der Partikel durchaus zulässig. Wir verwenden diese verallgemeinerte Definition, da dadurch die Untersuchung vereinfacht wird. Außerdem stellt sich heraus, dass kein Komplexitätsunterschied zu iDLA mit nur einem aktiven Partikel besteht.

Die folgenden beiden Lemmata von Moore und Machta [16] zeigen den engen Zusammenhang zwischen der Komplexitätsklasse **CC** und iDLA.

Satz 6. *iDLA-Vorhersage ist in der Klasse **CC**.*

Beweis. Für jeden Zeitpunkt t führen wir die Bool'schen Variablen $\text{aktiv}_t(i)$, für alle $i \in \mathcal{I}$, sowie $\text{besetzt}_t(s)$, für alle $s \in S$, ein. $\text{aktiv}_t(i) = 1$ genau dann, wenn das Partikel i zum Zeitpunkt t aktiv ist. Analog ist $\text{besetzt}_t(s) = 1$ genau dann, wenn der Knoten s zum Zeitpunkt t besetzt ist.

Es gilt $\text{aktiv}_0(i) = 1$, für alle $i \in \mathcal{I}$, sowie $\text{besetzt}_0(s) = 0$, für alle $s \in S$. Dies sind die Eingabewerte des zu konstruierenden Schaltkreises.

Aufgrund der Regeln für iDLA ergeben sich die folgenden Zustandsgleichungen für den Übergang (t, i, s)

$$\begin{aligned}\text{aktiv}_{t+1}(i) &= \text{besetzt}_t(s) \wedge \text{aktiv}_t(i) \\ \text{besetzt}_{t+1}(s) &= \text{besetzt}_t(s) \vee \text{aktiv}_t(i).\end{aligned}$$

Also ein Partikel bleibt nur dann aktiv, wenn es zum vorhergehenden Zeitpunkt noch aktiv war, und es einen bereits besetzten Knoten besucht hat. Analog ist ein Knoten

besetzt, wenn er zum vorhergehenden Zeitpunkt bereits besetzt gewesen ist, oder er von einem noch aktiven Partikel besucht wurde.

Diese Zustandsgleichungen haben nun aber genau die Struktur eines Komparatorgatters (Abbildung 4.1). Es lässt sich also jeder Übergang (t, i, s) durch ein Komparatorgatter mit den Eingabewerten $\text{aktiv}_t(i)$, $\text{besetzt}_t(s)$ und Ausgabe $\text{aktiv}_{t+1}(i)$, $\text{besetzt}_{t+1}(s)$ darstellen. Dass der so definierte Schaltkreis keine Verzweigungen besitzt, wird Aufgrund der oben genannten Konsistenzbedingung der Übergangstrippel gewährleistet. Jedem Übergangstrippel entspricht genau ein Gatter, das heißt die Schaltkreisgröße ist polynomiell in der Inputgröße und somit liegt iDLA-Vorhersage in **CC**. \square

Die Verbindung zwischen iDLA und der Komplexitätsklasse ist allerdings noch enger. iDLA-Vorhersage ist, selbst wenn wir uns auf ein einzelnes aktives Partikel in jedem Zeitpunkt beschränken, **CC**-vollständig, das heißt jede Funktion in **CC** lässt sich durch iDLA mit einer geeignet gewählten Schrittfolge berechnen.

Satz 7. *iDLA-Vorhersage für \mathbb{Z}^d ($d \geq 2$) ist **CC**-vollständig.*

Beweis. Zum Beweis müssen wir einen beliebigen Komparator Schaltkreis mittels iDLA simulieren. Das grundlegende Element dafür ist die in Abbildung 4.2 beschriebene Schrittfolge zur Simulation eines einzelnen Komparator-Gatters. Dabei sind alle Knoten in der ersten Zeile besetzt, und der Knoten γ ist nicht besetzt. Die Knoten α , β und γ stellen Wahrheitswerte dar („besetzt“ $\cong 1$ und „nicht besetzt“ $\cong 0$). Wir benötigen nur 3 Knoten zur Simulation aller 4 Ein- und Ausgänge eines Komparator-Gatters, da die Knoten vor und nach der Schrittfolge unterschiedliche Variablen repräsentieren.

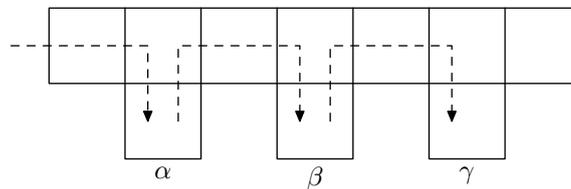


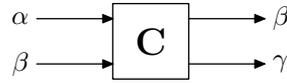
Abbildung 4.2: Schrittfolge zur Simulation eines Komparator-Gatters

Der Knoten γ ist nach dieser Schrittfolge nur besetzt, wenn sowohl α als auch β vorher schon besetzt waren, insgesamt ergibt sich die folgende Situation.

Sei s ein Zeitpunkt vor der Durchführung der Schrittfolge aus Abbildung 4.2 und t ein Zeitpunkt danach, dann gelten für die Knoten α , β und γ die folgenden Übergangsgleichungen:

$$\begin{aligned} \text{besetzt}_t(\alpha) &= 1 \\ \text{besetzt}_t(\beta) &= \text{besetzt}_s(\alpha) \vee \text{besetzt}_s(\beta) \\ \text{besetzt}_t(\gamma) &= \text{besetzt}_s(\alpha) \wedge \text{besetzt}_s(\beta) \end{aligned}$$

Die obige Konstruktion simuliert also das Gatter:



Wenn wir mehrere solche Komponenten verbinden, können wir beliebige Komparator-Schaltkreise simulieren. Ein Schaltkreis mit n Gattern hat höchstens $4n$ Verbindungsdrähte, deren Wahrheitswerte unter Verwendung von $3n$ Knoten simulieren können. Wenn wir diese Knoten alle auf einer horizontalen Linie benachbart wählen, benötigen wir höchstens $9n + 5$ Schritte um 3 von diesen Knoten nacheinander zu besuchen.

Das heißt zur Simulation eines Komparator-Schaltkreises wird ein Aggregationscluster mit Größe $O(n)$, sowie $O(n^2)$ Simulationsschritte benötigt. \square

Der Zusammenhang zwischen iDLA und der Klasse **CC** besteht in der Unfähigkeit beider Modelle, vorhandene Informationen zu kopieren. Wenn ein iDLA Partikel den Zustand eines Knotens abfragt, wird dieser dadurch verändert. Bei den Komparatorgattern folgt die Unmöglichkeit, Information zu kopieren aus dem Verbot, dass sich Verbindungsdrähte verzweigen.

Wie man aus dem Beweis zu Satz 7 sieht, wird zur Simulation von Komparatorschaltkreisen, nicht das gesamte Gitter benötigt. Es reicht eine kammartige Struktur mit einem Verbindungsweg und daran angehängte Knoten zur Informationsspeicherung aus. Wird der Verbindungsweg auf einen einzelnen Knoten kontrahiert, so funktioniert die Konstruktion auch mit einem Stern.

Für \mathbb{Z} ist iDLA-Vorhersage allerdings nicht mehr **CC**-vollständig, da mit den beiden Randpunkten des Clusters zu jedem Zeitpunkt nur zwei Knoten zur Informationsspeicherung zur Verfügung stehen. iDLA-Vorhersage für \mathbb{Z} liegt auch in der Klasse der effizient parallelisierbaren Funktion **NC** [16].

Kapitel 5

Internal-Diffusion-Limited Aggregation auf Bäumen

Dieser Abschnitt widmet sich der Definition von *Bäumen mit endlich vielen Kegelklassen* (manchmal auch *periodische Bäume* genannt). Irrfahrten auf Bäumen mit endlich vielen Kegelklassen wurden im Detail von Nagnibeda und Woess [17] untersucht.

5.1 Bäume mit endlich vielen Kegelklassen

Definition 25. Sei T ein lokalendlicher Baum mit Wurzel \mathcal{O} , und x, y Knoten von T . Wir schreiben $x \rightarrow y$ wenn x der unmittelbare Vorgänger von y ist, und $x \xrightarrow{*} y$ wenn x auf dem kürzesten Weg von \mathcal{O} nach y liegt.

Die Menge

$$C_x := \{y \in T : x \xrightarrow{*} y\}$$

heißt der von x aufgespannte *Kegel*.

Definition 26. Ein lokalendlicher Baum T heißt *Baum mit endlich vielen Kegelklassen*, wenn die Anzahl der Isomorphieklassen der Kegel C_x , $x \in T$ endlich ist. Für die Isomorphie werden dabei die C_x als Bäume mit Wurzel x betrachtet.

Ein Baum mit endlich vielen Kegelklassen lässt sich mit Hilfe eines endlichen Graphen beschreiben, dessen Knoten im Wesentlichen den Isomorphieklassen der Kegel entsprechen. Wenn wir nur Irrfahrten betrachten, die an die Kegelstruktur des Baumes adaptiert sind genügt es überdies, die Übergangswahrscheinlichkeiten nur auf den Kanten des dem Baum zugehörigen Graphen festzulegen.

Sei nun T ein Baum mit endlich vielen Kegelklassen. Wir bezeichnen mit $[C_x]$ die Isomorphieklasse des Kegels C_x und mit \mathcal{I} die Menge aller Isomorphieklassen $[C_x]$ von T , für $x \neq \mathcal{O}$. Die Abbildung

$$\begin{aligned} \iota : T &\longrightarrow \mathcal{I} \cup \{\mathcal{O}\} \\ \mathcal{O} &\longrightarrow \mathcal{O} \\ x &\longrightarrow [C_x] \text{ für } x \neq \mathcal{O} \end{aligned}$$

weist jedem Knoten seine Kegelklasse zu.

Sei $G_T = (V_T, E_T)$ ein gerichteter Graph mit Knotenmenge $V_T = \mathcal{I} \cup \{\mathcal{O}\}$. Es existiert eine Kante $(i, j) \in E_T$ genau dann, wenn eine Kante $x \rightarrow y$ in T existiert, mit $\iota(x) = i$ und $\iota(y) = j$. Außerdem gibt es noch Kanten $(\mathcal{O}, i) \in E_T$, wenn $\mathcal{O} \rightarrow x$ in T mit $\iota(x) = i$.

Für $(i, j) \in E_T$ bezeichnen die Labels $\mathbf{d}(i, j)$ die „Vielfachheit“ der Kanten in T . Sei also ein $x \in T$ mit $\iota(x) = i$ fest gewählt, dann ist

$$\mathbf{d}(i, j) = |\{y \in T : x \rightarrow y, \iota(y) = j\}|.$$

Diese Definition ist eindeutig da die Anzahl und der Typ der Nachbarn von der Wahl der Repräsentanten der einzelnen Klassen unabhängig sind.

Der Baum T lässt sich eindeutig aus G_T rekonstruieren. Dabei entspricht jedem Knoten in T ein im Knoten \mathcal{O} beginnender, gerichteter Weg in G_T . Wobei $(i, j) \in E_T$ als $\mathbf{d}(i, j)$ unterschiedliche Kanten angesehen werden. Die Wurzel von T entspricht dabei dem Weg der Länge 0.

Zur Definition einer Irrfahrt auf T wählen wir nun Gewichte $\mathbf{p}(i, j)$ für alle Kanten $(i, j) \in E_T$, sodass gilt

$$\sum_{j \in \mathcal{I}} \mathbf{d}(\mathcal{O}, j) \mathbf{p}(\mathcal{O}, j) = 1,$$

und

$$\sum_{j \in \mathcal{I}} \mathbf{d}(i, j) \mathbf{p}(i, j) \leq 1 \quad \forall i \in \mathcal{I}.$$

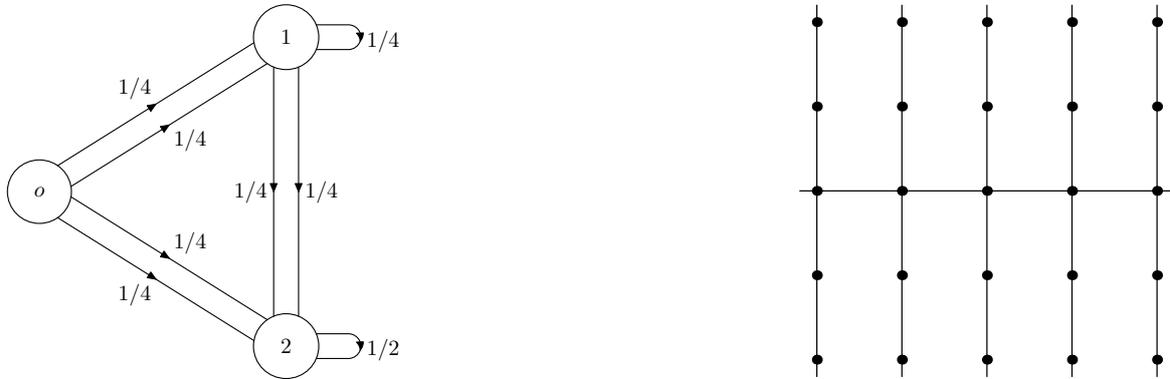
Wenn y Nachfolger von x in T ist, mit $\iota(x) = i$ und $\iota(y) = j$, dann sind die Einschnitt Übergangswahrscheinlichkeiten definiert als

$$\begin{aligned} p(x, y) &:= \mathbf{p}(i, j) \\ p(y, x) &:= \mathbf{p}_{-1}(i) = 1 - \sum_{j \in \mathcal{I}} \mathbf{d}(i, j) \mathbf{p}(i, j). \end{aligned} \quad (5.1)$$

Bemerkung 19. Man erhält die *einfache Irrfahrt* auf T für die folgende spezielle Wahl der Gewichte

$$\mathbf{p}(i, j) := \begin{cases} (\sum_{k \in \mathcal{I}} \mathbf{d}(\mathcal{O}, k))^{-1}, & \text{für } i = \mathcal{O} \\ (1 + \sum_{k \in \mathcal{I}} \mathbf{d}(i, k))^{-1}, & \text{sonst.} \end{cases}$$

Beispiel 2. Der zweidimensionale Kamm mit der einfachen Irrfahrt.



Aus den Labels \mathbf{d} und \mathbf{p} erhalten wir auch eine Bedingung für die Rekurrenz bzw. Transienz der oben definierten Irrfahrten [17].

Satz 8. Sei die Matrix A definiert als

$$A = (a(i, j))_{i, j \in \mathcal{I}}, \quad \text{mit} \quad a(i, j) = \frac{\mathbf{d}(i, j) \mathbf{p}(i, j)}{\mathbf{p}_{-1}(i)},$$

und sei $\lambda(A)$ der größte positive Eigenwert von A , dann ist die in (5.1) definierte

Irrfahrt

$$\begin{aligned} \text{positiv rekurrent} &\iff \lambda(A) < 1 \\ \text{null rekurrent} &\iff \lambda(A) = 1 \\ \text{transient} &\iff \lambda(A) > 1. \end{aligned}$$

Bemerkung 20. Für die einfache Irrfahrt ist $a(i, j) = \mathbf{d}(i, j)$.

5.2 Das Simulationsprogramm

Im Rahmen dieser Arbeit wurde ein Programm erstellt, das die Simulation von Aggregationsclustern auf beliebigen Bäumen mit endlich vielen Kegeltypen (Siehe Definition 26) erlaubt.

Der Baum mit den korrespondierenden Übergangswahrscheinlichkeiten wird über eine einfache Textdatei spezifiziert. Das Dateiformat folgt der untenstehenden Grammatik. Zusätzlich werden Zeilen, die mit einem „%“-Zeichen beginnen, als Kommentar ignoriert.

Grammatik 1.

$$\begin{aligned} & ::= \text{nodetypes root edges} \\ \text{nodetypes} & ::= \text{nodetypes: list_of_nodes} \\ \text{list_of_nodes} & ::= \text{node; | node , list_of_nodes} \\ \text{node} & ::= \mathbf{literal} \\ \text{root} & ::= \text{root: node} \\ \text{edges} & ::= \text{edges: list_of_forward_edges list_of_back_edges} \\ \text{list_of_forward_edges} & ::= \text{forward_edge | forward_edge list_of_forward_edges} \\ \text{forward_edge} & ::= \text{node} \rightarrow \text{node} [\mathbf{numexpr}] \text{vis_parameter}_{\text{opt}} \\ \text{vis_parameter} & ::= \{ \mathbf{numexpr} \text{a}_{\text{opt}}, \mathbf{numexpr} \text{a}_{\text{opt}}, \mathbf{numexpr} \text{a}_{\text{opt}} \} \\ \text{list_of_back_edges} & ::= \text{back_edge | back_edge list_of_back_edges} \\ \text{back_edge} & ::= \text{node} \leftarrow [\mathbf{numexpr}] \end{aligned}$$

Die Grammatik erlaubt auch die Spezifikation einer Einbettung des resultierenden Baumes in die Ebene. Das für Vorwärtskanten optionale Nonterminal *vis_parameter* ist dabei für die eigentliche Simulation unerheblich und enthält die für die Einbettung des Graphen benötigten Parameter $\{\varphi, r, s\}$. $\varphi \in [0, 2)$ bezeichnet dabei den Winkel der Kante, r einen Skalierungsfaktor für den Winkel und s einen Skalierungsfaktor für die Länge der Kante. Alle diese Werte sind relativ zum Vorgänger der Kante, außer wenn die Variante mit einem angehängtem a gewählt wird, in diesem Fall bezeichnen die Parameter absolute Werte.

Abbildung 5.1 zeigt einige mit diesem System erzeugte Visualisierungen von Aggregationsclustern.

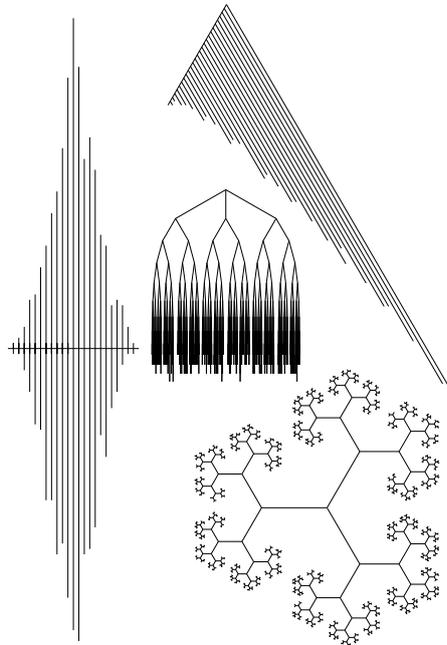


Abbildung 5.1: Beispiele von Cluster-Visualisierungen

Die Visualisierungsparameter werden, außer zur Erzeugung von Bildern, auch von einem Teil der Analyse-Funktionen verwendet um Informationen zur Clusterverteilung zu erhalten. So lassen sich durch eine geeignete Einbettung, wie zum Beispiel mit den Visualisierungsparametern $\{1, 0, 1\}$ für jeden Kantentyp, auch Tests bezüglich der Graphenmetrik realisieren.

Beispiel 3. Das untenstehende Listing definiert den zweidimensionalen Kamm mit der einfachen Irrfahrt aus Beispiel 2. Die angegebenen Visualisierungsparameter realisieren

die Standardeinbettung des Kammes in das zweidimensionale Gitter.

```
nodetypes:
k_0, k_1, k_2;

root:
k_0

edges:
% Ausgangsknoten -> Zielknoten [Übergangswahrscheinlichkeit]
% {Winkel in Radianten/pi, Winkelskalierung, Längenskalierung}
k_0 -> k_1 [1/4] { 0, 1, 1}
k_0 -> k_2 [1/4] {1/2, 1, 1}
k_0 -> k_1 [1/4] { 1, 1, 1}
k_0 -> k_2 [1/4] {3/2, 1, 1}

k_1 -> k_1 [1/4] { 0, 1, 1}
k_1 -> k_2 [1/4] {1/2, 1, 1}
k_1 -> k_2 [1/4] {3/2, 1, 1}

k_2 -> k_2 [1/2] { 0, 1, 1}

% Wurzelknoten <- [Übergangswahrscheinlichkeit]
k_1 <- [1/4]
k_2 <- [1/2]
```

Eine naive Vorgangsweise für die Simulation des Aggregationsclusters ist für Bäume, in denen Knoten mit jeweils nur einem Nachfolger existieren (wie zum Beispiel im Baum von Beispiel 2), leider nicht praktikabel. Da die einfache Irrfahrt auf linearen Ketten rekurrent ist, ist der Zeitaufwand zur Simulation selbst moderater Clustergrößen enorm (mehrere Tage für Clustergrößen von ≈ 100.000 Partikeln). Glücklicherweise ist im Fall der linearen Ketten, also für Knoten aus einer Klasse $k \in \mathcal{I}$ für die nur eine einzige Ableitungsregel vom Typ

```
k -> k [p]
```

definiert ist, eine explizite Berechnung der Wahrscheinlichkeit dass der Cluster verlassen wird mittels der *Gambler's-Ruin-Formel* möglich.

Sei der Knoten v mit $\iota(v) = k$ derart, dass für den Vorgänger u von v gilt dass $\iota(u) \neq k$. Bezeichne l die Anzahl der Nachfolger von v im aktuellen Aggregationscluster. Die Wahrscheinlichkeit, dass ein Partikel welches sich in v befindet den Cluster verlässt ohne in weiterer Folge u zu betreten, ist gleich der Wahrscheinlichkeit dass ein Spieler mit einem Startkapital von 0 bei einer Folge von unabhängigen Glücksspielen mit Gewinnwahrscheinlichkeit p einen Reingewinn von $l + 1$ erzielt, ohne vorher jemals ins Minus zu geraten.

Nach der *Gambler's-Ruin-Formel* ist diese Wahrscheinlichkeit gleich

$$\frac{1}{l+2} \quad \text{für } p = \frac{1}{2},$$

und

$$\frac{1-t}{t^{-(l+1)}-t} \quad \text{mit } t = \frac{p}{1-p} \quad \text{für } p \neq \frac{1}{2}.$$

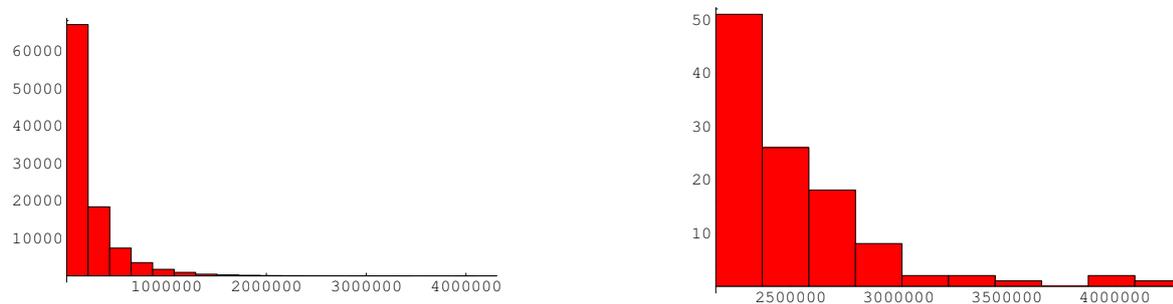
Mit obiger Wahrscheinlichkeit wird der Cluster im nächsten Simulationsschritt verlassen, und somit die Kette am Knoten v um einen Nachfolger verlängert, andernfalls führt man die Simulation im Vorgänger des Knotens v fort.

Eine Einbeziehung dieses Spezialfalles verringert die Simulationszeit für den oben genannten Fall auf wenige Minuten und ermöglicht erst die Simulation von Clustern in der Größenordnung von einigen Millionen Partikeln.

Trotz dieser Maßnahme werden für die Simulation sehr lange Folgen von Zufallszahlen benötigt. Um Artefakte zu minimieren wird der Pseudozufallszahlgenerator „Mersenne-Twister“¹ von Makoto Matsumoto und Takuji Nishimura [14] mit einer Zykluslänge von $2^{19937} - 1$ verwendet.

¹Verwendet wird die Implementierung aus der GNU Scientific Library (GSL). Falls benötigt kann zur Laufzeit auch ein anderer Zufallszahlgenerator ausgewählt werden. Siehe dazu die Dokumentation der GSL unter <http://www.gnu.org/software/gsl>.

Direkte Simulation



Explizite Berechnung linearer Ketten

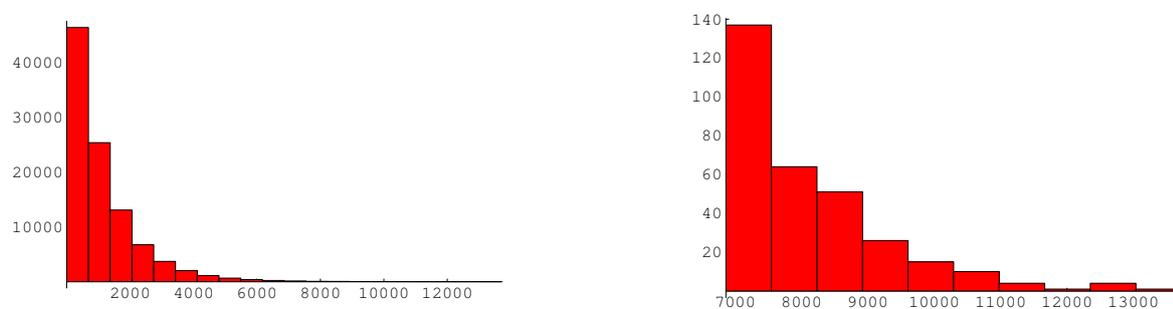


Abbildung 5.2: Anzahl der Simulationsschritte

Kapitel 6

Simulationsergebnisse

In diesem Kapitel werden einige Simulationsresultate vorgestellt, die mit Hilfe des in Abschnitt 5.2 vorgestellten Simulationsprogrammes gewonnen wurden.

6.1 Homogene Bäume

In [9], Abschnitt 7 wird ein zu Satz 1 analoges unpubliziertes Resultat von Etheridge und Lawler für den Fall des regulären Baumes vom Grad 3 (\mathbb{T}_3) und der einfachen Irrfahrt erwähnt. Blachère und Brofferio erhalten ein etwas schwächeres Resultat für beliebige reguläre Bäume \mathbb{T}_q ($q \geq 3$) als Spezialfall des Hauptresultates in [4]. Siehe dazu auch Abschnitt 7.1. Weitere Resultate zu iDLA auf homogenen Bäumen finden sich ausserdem in der Dissertation von Blachère [2].

Satz 9. *Sei $\{A(i)\}_{i \geq 1}$ eine Folge von iDLA-Clustern des Baumes \mathbb{T}_q ($q \geq 3$) mit der einfachen Irrfahrt, dann existieren Konstanten C_1, C_2 und $n_0 \in \mathbb{N}$ sodass gilt*

$$\mathbb{P}[\mathcal{B}(n - C_1 \log n) \subset A(|\mathcal{B}(n)|) \subset \mathcal{B}(n + C_2 \sqrt{n}), \forall n \geq n_0] = 1.$$

Dabei bezeichnet $\mathcal{B}(n)$ die „Kugel“ mit Radius n bezüglich der Graphenmetrik. Für das Volumen und den Rand der Kugeln gilt

$$|\mathcal{B}(n)| = \frac{q(q-1)^n - 2}{q-2} \quad \text{sowie} \quad |\partial\mathcal{B}(n)| = q(q-1)^n.$$

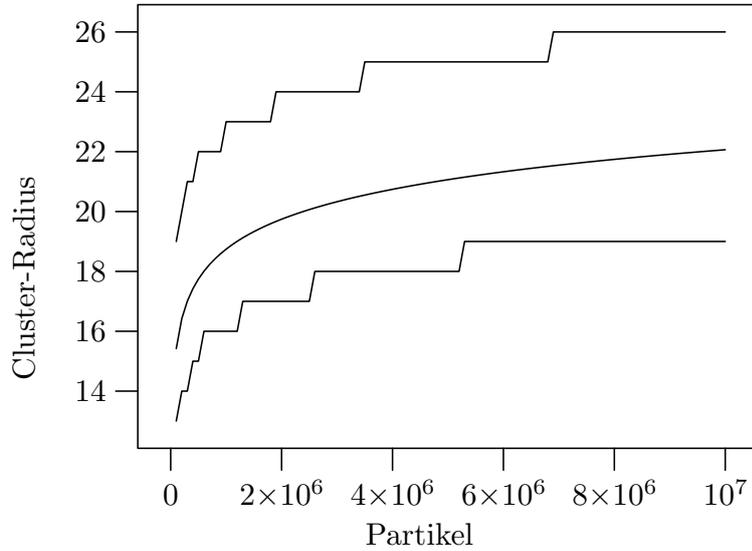


Abbildung 6.1: Radius der Aggregationscluster von \mathbb{T}_3

Weiters gelten die folgenden unteren Schranken für die inneren und äußeren Fluktuationen, also den Abweichungen der iDLA-Cluster von der idealen Kugelform.

Satz 10. *Unter den gleichen Voraussetzungen wie im vorhergehenden Satz gelten die Aussagen:*

(a) Für jede Konstante $\mathcal{C}_0 < \left(\frac{\ln(q-1)}{\ln q}\right)^{1/2}$ gilt:

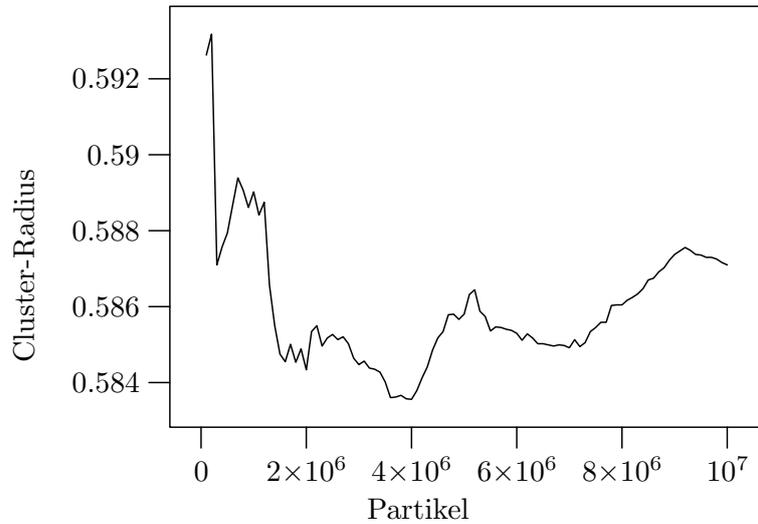
$$\mathbb{P}[A(|\mathcal{B}(n)|) \cap \partial\mathcal{B}(n + \mathcal{C}_0\sqrt{n}) = \emptyset \text{ fast sicher}] = 0$$

(b) Für jede Konstante $\mathcal{C}_1 < \frac{1}{2\ln(q-1)}$ gilt:

$$\mathbb{P}[A(|\mathcal{B}(n)|)^c \cap \mathcal{B}(n - \mathcal{C}_1 \ln n) \neq \emptyset \text{ fast sicher}] = 1$$

Homogene Bäume sind, aufgrund des exponentiellen Wachstums der Kugeln $\mathcal{B}(n)$, nicht besonders gut für Simulationen geeignet. Es folgt hier deshalb auch nur kurz ein Resultat einer Simulation mit 10 Millionen Partikeln.

Abbildung 6.1 zeigt den minimalen und den maximalen Clusterradius, sowie die mittlere Entfernung der Randpunkte des Clusters von der Wurzel, in Abhängigkeit von der

Abbildung 6.2: Varianzen Clusterrandpunkte von \mathbb{T}_3

Clustergröße, für die einfache Irrfahrt auf \mathbb{T}_3 . In Abbildung 6.2 sind die Varianzen der Radien der Clusterrandpunkte angegeben. Wie erwartet widersprechen die ermittelten Daten nicht den Sätzen 9 und 10, sie sind allerdings, da nur eine Baumtiefe von ≈ 20 erreicht wird, auch nicht besonders aussagekräftig.

Interessantere Ergebnisse lassen sich für Bäume mit geringerem Wachstum erzielen. Der Rest dieses Kapitels behandelt eine solche Klasse von periodischen Bäumen.

6.2 Der Kamm

Der d -dimensionale Kamm K_d ist ein spannender Baum des d -dimensionalen Gitters, also ein Baum, der alle Knoten von \mathbb{Z}^d enthält. K_d wird rekursiv definiert. $K_1 = \mathbb{Z}$ und K_d erhält man aus K_{d-1} , indem man an jedem Knoten von K_{d-1} zwei unendliche Pfade (also eine Kopie von \mathbb{Z}) anheftet.

K_d ist ein Baum mit d Kegelklassen. Für den erzeugenden Graphen $G_{K_d} = (V, E)$ ergibt sich die Knotenmenge $V = \{0, \dots, d\}$, die Kantenmenge

$$E_d = \{(i, j) : \forall i < j, i = 0, \dots, d-1\} \cup \{(i, i) : i = 1, \dots, d\},$$

und die Kantenlabels

$$\begin{aligned} \mathbf{d}(i, j) &= 2 \quad \text{für } \forall i < j, i = 0, \dots, d-1 \\ \mathbf{d}(i, i) &= 1 \quad \text{für } i = 1, \dots, d. \end{aligned}$$

Für die einfache Irrfahrt auf K_d erhalten wir, nach Bemerkung 19, als Gewichte für die Übergangswahrscheinlichkeiten:

$$\begin{aligned} \mathbf{p}(0, j) &= \frac{1}{2d} && \text{für } j = 1, \dots, d \\ \mathbf{p}(i, j) &= \frac{1}{2(d-i+1)} && \text{für } i = 1, \dots, d \text{ und } i < j \\ \mathbf{p}_{-1}(i) &= \frac{1}{2(d-i+1)} && \text{für } i = 1, \dots, d \end{aligned}$$

Für die einfache Irrfahrt ist die Matrix A aus Satz 8, aufgrund von Bemerkung 20, eine obere Dreiecksmatrix mit Hauptdiagonalelementen gleich 1, und allen Einträgen oberhalb der Hauptdiagonale gleich 2. Also ist $\lambda(A) = 1$ und somit ist, nach Satz 8, die einfache Irrfahrt auf dem Kamm K_d *null-rekurrent*.

Für eine detaillierte Untersuchung von Irrfahrten auf K_2 siehe [1].

6.2.1 Einfache Irrfahrt

In diesem Abschnitt präsentieren wir Simulationsergebnisse für iDLA auf dem *zweidimensionalen Kamm* (siehe Beispiel 2). Als zugrundeliegende Irrfahrt verwenden wir dabei die einfache Irrfahrt. Die Steuerungsdatei für das Simulationsprogramm ist dabei in Beispiel 3 angegeben.

Abbildung 6.3 zeigt Simulationscluster für den *zweidimensionalen Kamm*.

Diese Bilder legen die Vermutung nahe, dass der Rand des Clusters durch eine Gleichung der Form

$$\left(\frac{x}{h_x}\right)^\alpha + \left(\frac{y}{h_y}\right)^\beta = 1 \tag{6.1}$$

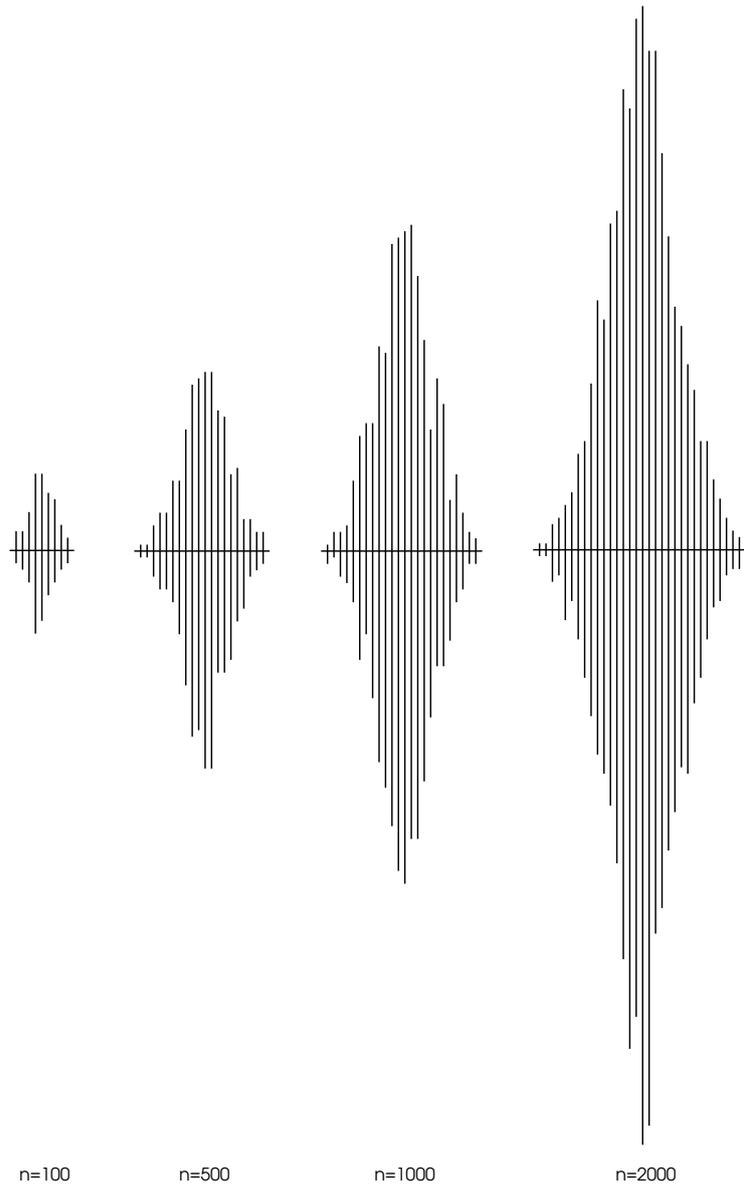


Abbildung 6.3: Aggregationscluster des zweidimensionalen Kammes

<i>keine Beschleunigung</i>	<i>Beschleunigung</i>
$\bar{\xi} = 0.344421$ $\text{Var}(\xi) = 1.11265 \times 10^{-5}$	$\bar{\xi} = 0.343377$ $\text{Var}(\xi) = 1.94179 \times 10^{-5}$
$\bar{\eta} = 0.674097$ $\text{Var}(\eta) = 5.84893 \times 10^{-5}$	$\bar{\eta} = 0.66668$ $\text{Var}(\eta) = 5.06347 \times 10^{-5}$
$\bar{k} = 1.13468$ $\text{Var}(k) = 1.822 \times 10^{-3}$	$\bar{k} = 1.14747$ $\text{Var}(k) = 3.0898 \times 10^{-3}$
$\bar{l} = 0.528354$ $\text{Var}(l) = 2.06781 \times 10^{-3}$	$\bar{l} = 0.57294$ $\text{Var}(l) = 2.34053 \times 10^{-3}$

Tabelle 6.1: Vergleich der Simulationen am Kamm (100000 Partikel)

beschrieben wird. Die Halbachsen h_x und h_y sind dabei Funktionen von n , der Anzahl der Partikel des Clusters. Aus den Simulationsdaten ergibt sich weiters die Vermutung von Halbachsen der Form

$$h_x(n) = k \cdot n^\xi \quad h_y(n) = l \cdot n^\eta$$

für die reellen Parameter ξ , η , k und l .

In Tabelle 6.1 vergleichen wir die Simulationsergebnisse bei direkter Simulation von 100000 Partikeln mit den Ergebnissen unter Verwendung von *Gambler's-Ruin*-Beschleunigung (siehe vorhergehenden Abschnitt). Angegeben sind die Schätzwerte für die beiden Halbachsen h_x und h_y . Es wurden jeweils 10 Durchgänge gerechnet. Die Rechenzeit bei direkter Simulation betrug ≈ 12 Stunden, im Vergleich zu 30 Minuten bei Verwendung des beschleunigten Algorithmus.

Die Histogramme in Abbildung 5.2 zeigen die Verteilung der für jedes einzelne Teilchen benötigten Simulationsschritte (die Anzahl der benötigten Zufallszahlen). Im Falle der direkten Simulation stimmt die Anzahl der Simulationsschritte mit der Anzahl der Schritte, die das Partikel benötigt um den Aggregationscluster zu verlassen, überein.

Insgesamt erhält man eine gute Übereinstimmung zwischen den beiden Simulationen, der beschleunigte Algorithmus erweist sich also als durchaus praktikabel.

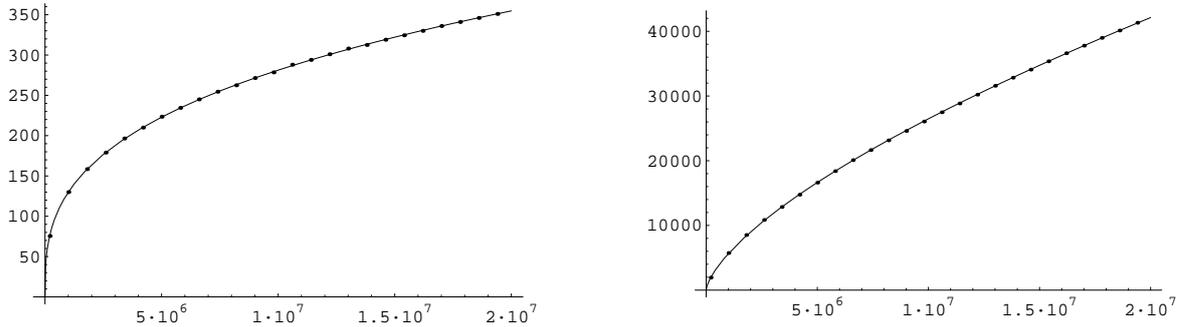


Abbildung 6.4: Halbachsen des zweidimensionalen Kamms

Abbildung 6.4 zeigt Graphen der Halbachsen h_x und h_y in Abhängigkeit von der Clustergröße n . Die Punkte bezeichnen dabei durch Simulation ermittelte Werte, die durchgängigen Linien sind die Schätzer $\bar{h}_x(n)$ und $\bar{h}_y(n)$. In Abbildung 6.5 wird der Rand des Clusters dargestellt. Aufgrund der Symmetrie beschränkt sich die Abbildung auf den ersten Quadranten. Die durchgängige Linie ist dabei der Graph der Kurve 6.1 mit den aus den Daten ermittelten Parametern.

In der folgenden Tabelle schließlich sind die Ergebnisse¹ von fünf Simulationsläufen mit einer Clustergröße von jeweils 20 Millionen Partikeln angegeben.

$\bar{\xi} = 0.333244$	$\text{Var}(\xi) = 1.064 \times 10^{-7}$
$\bar{k} = 1.31324$	$\text{Var}(k) = 5.319 \times 10^{-5}$
$\bar{\eta} = 0.665707$	$\text{Var}(\eta) = 4.466 \times 10^{-6}$
$\bar{l} = 0.582816$	$\text{Var}(l) = 4.059 \times 10^{-4}$
$\bar{\alpha} = 0.997102$	$\text{Var}(\alpha) = 1.363 \times 10^{-5}$
$\bar{\beta} = 0.500393$	$\text{Var}(\beta) = 1.808 \times 10^{-6}$

Für α und β ergeben sich also aus den Daten die Schätzwerte

$$\alpha \approx 1 \quad \beta \approx 0.5,$$

und für die Exponenten der Halbachsen

$$\xi \approx \frac{1}{3} \quad \eta \approx \frac{2}{3}.$$

¹Die Schätzungen der Parameter erfolgte in Mathematica mittels der Funktion *NonlinearFit* aus dem *Statistics* Paket.

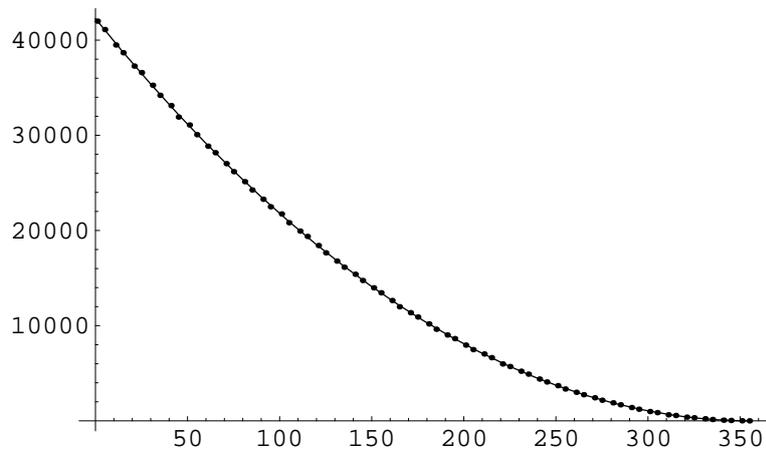
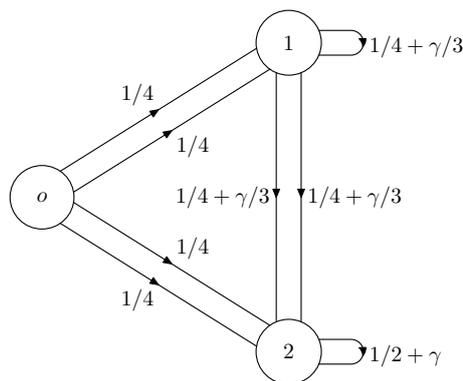


Abbildung 6.5: Aggregationscluster von K_2

Diese Vermutung lässt sich für Kämmen in höheren Dimensionen verallgemeinern, siehe dazu Abschnitt 6.4.

6.2.2 Inhomogene Irrfahrten

Bis jetzt wurden ausschließlich Beispiele mit der einfachen Irrfahrt betrachtet. In diesem Abschnitt untersuchen wir iDLA-Cluster für eine Klasse von inhomogenen Irrfahrten auf dem zweidimensionalen Kamm. Für $\gamma \in (-\frac{1}{2}, \frac{1}{4}]$ definiert der Multigraph



eine Irrfahrt auf dem zweidimensionalen Kamm (Für Werte von $\gamma \notin (-\frac{1}{2}, \frac{1}{4}]$ ergeben sich negative Wahrscheinlichkeiten beziehungsweise Wahrscheinlichkeiten > 1).

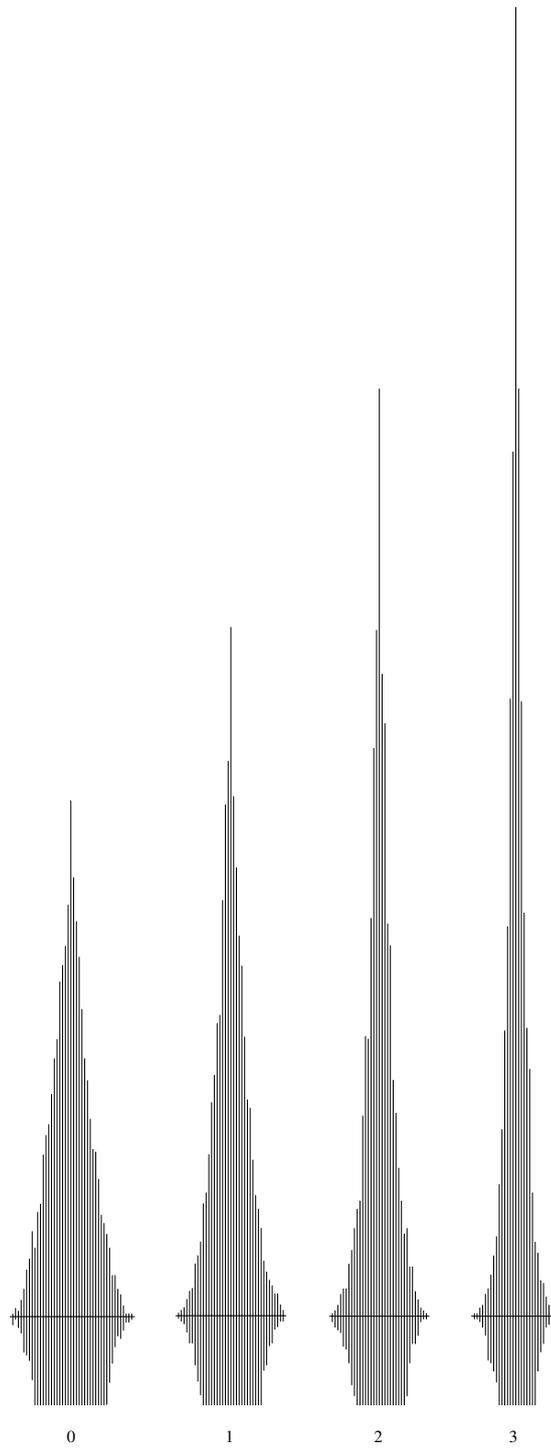
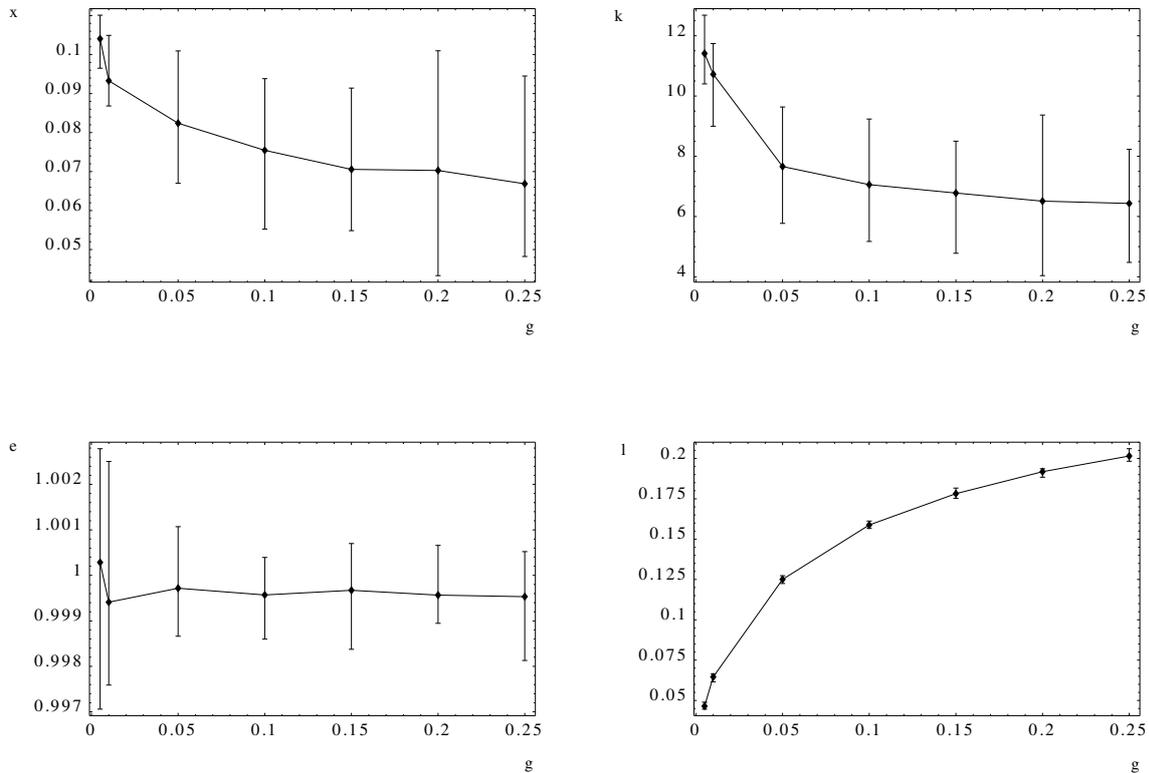


Abbildung 6.6: iDLA-Cluster auf K_2 mit inhomogener Irrfahrt

Abbildung 6.7: Halbachsenparameter für Biased-Irrfahrt auf K_2

Für $\gamma > 0$ ist diese Irrfahrt transient, im Gegensatz zum rekurrenten Fall der einfachen Irrfahrt $\gamma = 0$. Für iDLA bedeutet das einen Phasenübergang mit dem kritischen Wert $\gamma_c = 0$. Bei positivem γ fließt beinahe die gesamte Masse des Aggregationsclusters in die „Zähne“ des Kamms. Das heißt für die Halbachse $h_y(n) = l \cdot n^\eta$ erhält man den Parameter $\eta = 1$.

Abbildung 6.7 zeigt die geschätzten Parameter der beiden Halbachsen für Simulationen mit $0.005 \leq \gamma \leq 0.25$. Berechnet wurden für jeden Wert von γ jeweils 10 Aggregationscluster mit jeweils 10 Millionen Partikeln. Aufgetragen sind jeweils die Minimal-, Durchschnitts- und Maximalwerte der Parameter ξ , l , η und k .

Da die Cluster im transienten Fall kaum in Richtung der x -Achse wachsen, sind auch die Schätzwerte der Parameter ξ und k mit großen Fehlern behaftet und nicht besonders aussagekräftig. Aus dem selben Grund (nur wenige vorhandene Stützstellen auf der x -Achse mit zugleich sehr großen Werten auf der y -Achse) lässt sich eine Hypothese für die Form des Clusters aus den Simulationsdaten nicht mehr zufriedenstellend ableiten. Es ist allerdings anzunehmen, dass das Modell in Gleichung 6.1 Aufgrund des

Phasenüberganges nicht mehr anwendbar ist, da nicht nur ein quantitativer sondern auch ein qualitativer Unterschied zur einfachen Irrfahrt besteht.

6.3 Sternförmige Kämme

Das obige Ergebnis bleibt auch für eine größere Klasse von Bäumen gültig. Wir definieren k -symmetrische *Stern Kämme* S_k als Verallgemeinerung des 2-dimensionalen Kammes K_2 .

S_k besteht aus einem aus k einseitig unendlichen Strahlen zusammengesetzten Stern. Und an jeden Punkt dieses Sternes wird analog zum Kamm eine Kopie von \mathbb{Z} geheftet.

Der erzeugende Graph $G_{S_k} = (V, E)$ unterscheidet sich nur durch die Kantenlabels vom erzeugenden Graphen des 2-dimensionalen Kammes. Es gilt

$$\begin{aligned} \mathbf{d}(0, 1) &= k \\ \mathbf{d}(1, 2) &= 2 \\ \mathbf{d}(i, i) &= 1 \quad \text{für } i = 1, \dots, 2. \end{aligned}$$

Abbildung 6.8 zeigt internal DLA Cluster für die Graphen S_4 und S_7 .

Wird nun für jeden Randknoten des entstehenden Aggregationsclusters, die Anzahl seiner Vorgänger, die auf dem sternförmigen „Rücken“ des Graphen liegen, auf die x -Achse aufgetragen, und die Anzahl der Vorgänger auf den „Zähnen“ des Graphen auf die y -Achse, so ergibt sich das selbe Bild wie im Falle von K_2 . Tabelle 6.2 zeigt die geschätzten Parameter für S_3 und S_7 bezüglich des Modells 6.1. Wie erwartet bestehen die einzigen Unterschiede im Vergleich mit dem Standardkamm in den Größen der Skalierungsfaktoren k und l .

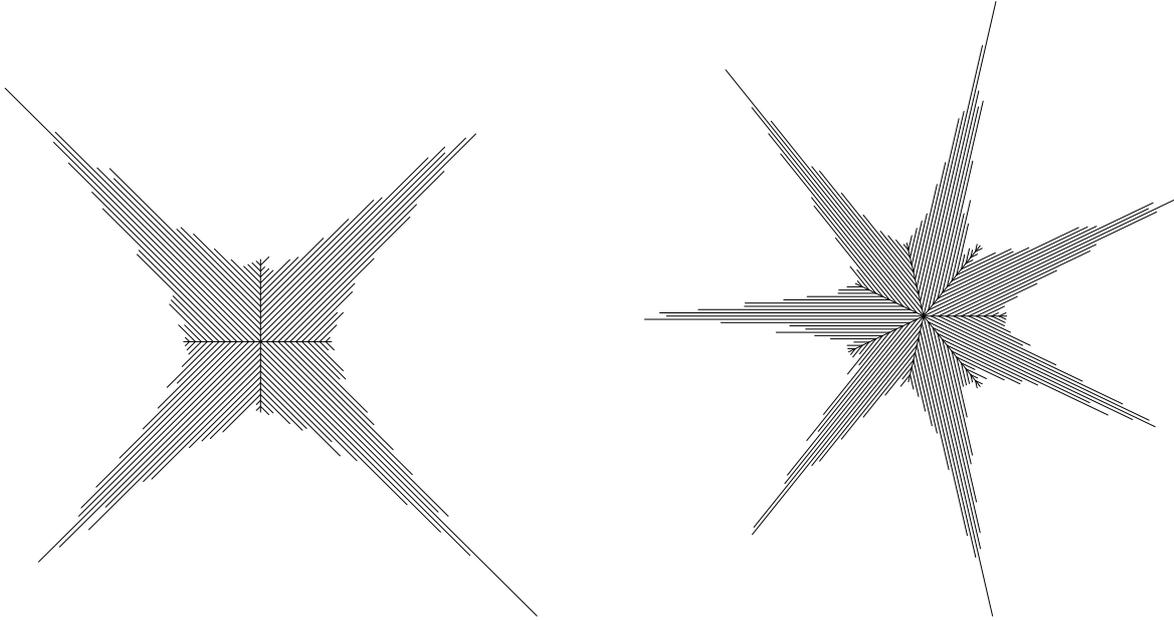


Abbildung 6.8: iDLA-Cluster auf S_4 und S_7

6.4 Kämmе in höheren Dimensionen

Im folgenden betrachten wir iDLA auf Kämmen in Dimensionen $d \geq 3$, bezüglich der einfachen Irrfahrt. Kämmе sind rekursiv definiert, so besteht K_3 aus einer Folge von 2-dimensionalen Kämmen, die jeweils an ihrer Wurzel miteinander verbunden sind. Es ist zu erwarten, dass die iDLA-Cluster diese Struktur widerspiegeln. Zum Beispiel zeigt wirklich jeder K_2 -Teilbaum von K_3 das gleiche Wachstumsverhalten, das wir schon am 2-dimensionalen Kamm gesehen haben.

Wir überprüfen zunächst die Annahme, ob sich die Struktur der iDLA-Cluster von K_2 einfach auf höhere Dimensionen verallgemeinern lässt. Sei $x_1, \dots, x_d \in \mathbb{R}^d$, dann wird

S_3		S_7	
$\alpha = 0.991$	$\beta = 0.502$	$\alpha = 0.986$	$\beta = 0.502$
$\xi = 0.333$	$\eta = 0.666$	$\xi = 0.332$	$\eta = 0.664$
$k = 1.156$	$l = 0.445$	$k = 0.883$	$l = 0.26$

Tabelle 6.2: Clusterparameter von S_3 und S_7

durch die algebraische Gleichung

$$\sum_{i=1}^d \left(\frac{x_i}{h_i(n)} \right)^{\alpha_i} \leq 1. \quad (6.2)$$

eine Teilmenge von \mathbb{R}^d , und daraus durch Restriktion auch eine Teilmenge von \mathbb{Z}^d definiert.

Dabei sind die Halbachsen h_i von der Form

$$h_i(n) = k_i \cdot n^{\xi_i}, \quad i = 1, \dots, d.$$

In Tabelle 6.3 sind die geschätzten Werte der Halbachsenparameter für die Kämmen K_2, \dots, K_5 angegeben. Es wurden jeweils 5 Simulationen mit jeweils 20 Millionen Partikeln durchgeführt. Angegeben sind die Mittelwerte und Varianzen für die Schätzer der Parameter ξ_i .

Dabei werden die Schätzwerte für kleine Indices i mit zunehmender Dimension, aufgrund des extrem langsamen Wachstums der Halbachsen, immer unsicherer. So gilt zum Beispiel für K_5 und die erste Halbachse $h_1(2 \times 10^7) = 11$. Besonders bei den Parametern ξ_1 und ξ_2 von K_4 und K_5 beeinträchtigen numerische Instabilitäten die Schätzung signifikant.

Dessen ungeachtet lässt sich doch eine gewisse Gesetzmäßigkeit erkennen. So gilt für die Summen der einzelnen Exponenten $\bar{\xi}_i$:

$$\begin{aligned} K_2: \quad & \bar{\xi}_1 + \bar{\xi}_2 = 0.998951 \\ K_3: \quad & \bar{\xi}_1 + \bar{\xi}_2 + \bar{\xi}_3 = 1.01357 \\ K_4: \quad & \bar{\xi}_1 + \bar{\xi}_2 + \bar{\xi}_3 + \bar{\xi}_4 = 1.0584 \\ K_5: \quad & \bar{\xi}_1 + \bar{\xi}_2 + \bar{\xi}_3 + \bar{\xi}_4 + \bar{\xi}_5 = 1.08581 \\ K_6: \quad & \bar{\xi}_1 + \bar{\xi}_2 + \bar{\xi}_3 + \bar{\xi}_4 + \bar{\xi}_5 + \bar{\xi}_6 = 1.14322 \end{aligned}$$

Weiters sind die Quotienten aufeinanderfolgender $\bar{\xi}_i$ für K_2

$$\frac{\bar{\xi}_2}{\bar{\xi}_1} = 1.99766,$$

K_2	
$\bar{\xi}_1 = 0.333244$	$\text{Var}(\xi_1) = 1.064 \times 10^{-7}$
$\bar{\xi}_2 = 0.665707$	$\text{Var}(\xi_2) = 4.466 \times 10^{-6}$
K_3	
$\bar{\xi}_1 = 0.156321$	$\text{Var}(\xi_1) = 7.981 \times 10^{-5}$
$\bar{\xi}_2 = 0.286142$	$\text{Var}(\xi_2) = 3.508 \times 10^{-6}$
$\bar{\xi}_3 = 0.571109$	$\text{Var}(\xi_3) = 2.160 \times 10^{-5}$
K_4	
$\bar{\xi}_1 = 0.103099$	$\text{Var}(\xi_1) = 2.885 \times 10^{-4}$
$\bar{\xi}_2 = 0.1512$	$\text{Var}(\xi_2) = 9.721 \times 10^{-5}$
$\bar{\xi}_3 = 0.268878$	$\text{Var}(\xi_3) = 1.013 \times 10^{-6}$
$\bar{\xi}_4 = 0.535228$	$\text{Var}(\xi_4) = 3.140 \times 10^{-5}$
K_5	
$\bar{\xi}_1 = 0.075992$	$\text{Var}(\xi_1) = 2.483 \times 10^{-4}$
$\bar{\xi}_2 = 0.085606$	$\text{Var}(\xi_2) = 5.759 \times 10^{-5}$
$\bar{\xi}_3 = 0.140836$	$\text{Var}(\xi_3) = 3.299 \times 10^{-6}$
$\bar{\xi}_4 = 0.263322$	$\text{Var}(\xi_4) = 6.005 \times 10^{-7}$
$\bar{\xi}_5 = 0.520049$	$\text{Var}(\xi_5) = 2.188 \times 10^{-5}$

Tabelle 6.3: Halbachsenparameter von K_2, \dots, K_5

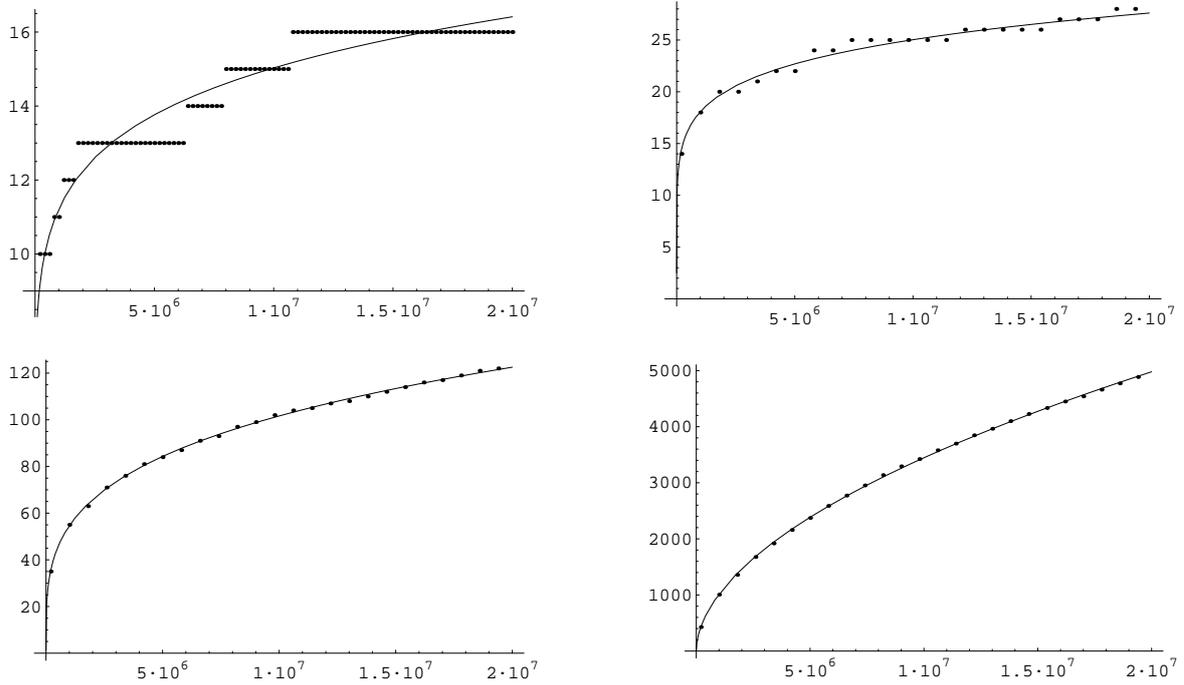


Abbildung 6.9: Halbachsen des 4-dimensionalen Kammes K_4

für den Kamm K_3

$$\frac{\bar{\xi}_2}{\bar{\xi}_1} = 1.83047 \quad \frac{\bar{\xi}_3}{\bar{\xi}_2} = 1.9959,$$

für K_4 analog

$$\frac{\bar{\xi}_2}{\bar{\xi}_1} = 1.46656 \quad \frac{\bar{\xi}_3}{\bar{\xi}_2} = 1.77829 \quad \frac{\bar{\xi}_4}{\bar{\xi}_3} = 1.9906.$$

Und schließlich für K_5

$$\frac{\bar{\xi}_2}{\bar{\xi}_1} = 1.12651 \quad \frac{\bar{\xi}_3}{\bar{\xi}_2} = 1.64516 \quad \frac{\bar{\xi}_4}{\bar{\xi}_3} = 1.8697 \quad \frac{\bar{\xi}_5}{\bar{\xi}_4} = 1.97495.$$

Aus diesen Daten ergibt sich für die Exponenten ξ_i die folgende Vermutung:

$$\sum_{i=1}^d \xi_i \approx 1 \quad \text{und} \quad \frac{\xi_i}{\xi_{i-1}} \approx 2, \quad i = 2, \dots, d \quad (6.3)$$

Der erste Teil dieser Vermutung passt recht gut zu den Daten. Die Vermutung für die Quotienten in (6.3) ist allerdings nur eine einfache Extrapolation des Sachverhaltes im

	$ \bar{\xi}_1 - \xi'_1 $	$ \bar{\xi}_2 - \xi'_2 $	$ \bar{\xi}_3 - \xi'_3 $	$ \bar{\xi}_4 - \xi'_4 $	$ \bar{\xi}_5 - \xi'_5 $	$ \bar{\xi}_6 - \xi'_6 $
K_2	0.00009	0.00096				
K_3	0.013	0.0004	0.0003			
K_4	0.036	0.017	0.002	0.002		
K_5	0.044	0.021	0.012	0.005	0.004	
K_6	0.038	0.041	0.033	0.013	0.005	0.014

Tabelle 6.4: Absolute Fehler der Halbachsenexponenten von K_2, \dots, K_6

zweidimensionalen Fall, und nicht vollständig aus den Simulationsdaten zu rechtfertigen. Aus den Daten ist nicht schlüssig ersichtlich, ob die zu Beobachtende Abnahme des Quotienten $\frac{\xi_i}{\xi_{i-1}}$ das wahre Verhalten widerspiegelt, oder nur ein Artefakt des langsamen Wachstums der kleinen Halbachsen ist.

Allerdings ist die absolute Abweichung der $\bar{\xi}_i$ von den exakten Lösungen ξ'_i der Gleichung (6.3) weit geringer als es die Fehler der Quotienten vermuten lassen würden. In Tabelle 6.4 sind die absoluten Fehler $|\bar{\xi}_i - \xi'_i|$ angegeben.

Die Schwierigkeit, die Parameter zu schätzen, wird in Abbildung 6.9 im Graphen links oben deutlich. Dieses Bild zeigt das Wachstum der Halbachse $h_1(n)$ im iDLA-Cluster von K_4 . Um die Länge dieser Halbachse zu verdoppeln, und damit bessere Schätzwerte zu erreichen, müsste eine Simulation mit rund 2×10^{10} Partikeln durchgeführt werden. Das ist das 1000-fache der bisher simulierten Clustergrößen. Für Käme höherer Dimensionen ist das Problem noch schwieriger.

Um diese Vermutung zu erhärten oder zu widerlegen, müssten also weitere Simulationen mit um einige Größenordnungen größeren Clustern durchgeführt werden, um bessere Schätzwerte für die langsam wachsenden Halbachsen zu erhalten. Die Rechenzeit stellt für Käme höherer Dimension, durch die stärkere Verzweigung und den damit verbundenen kürzeren linearen Ketten, nicht mehr den limitierenden Faktor dar. So benötigt die Simulation eines iDLA Clusters von K_2 mit 20 Millionen Partikeln auf einem Rechner mit zwei 3GHz-Prozessoren 22 Stunden, wogegen die selbe Simulation für den Kamm K_6 nur rund 40 Minuten benötigt, wobei noch ein signifikanter Teil der Rechenzeit auf die Ausgabe der Daten entfällt. Allerdings wird die

maximal erreichbare Clustergröße im vorliegenden Programm effektiv durch den vorhandenen Arbeitsspeicher begrenzt, und der ist bei den verfügbaren Rechnern (2GB RAM) mit 25 Millionen Partikeln ausgeschöpft.

Ein speziell für die Untersuchung von Kämmen geschriebenes Programm könnte den Cluster allerdings viel effizienter speichern. Wird die natürliche Einbettung von K_d in das Zahlengitter \mathbb{Z}^d zugrundegelegt, so lässt sich jeder Knoten von K_d , unabhängig von seiner Entfernung von der Wurzel, mittels d Zahlen adressieren. Da der Cluster $A(n)$ zusammenhängend ist, könnte man sich darauf beschränken den Rand von $A(n)$ zu speichern. Damit ließen sich weit größere iDLA-Cluster berechnen, bevor der Arbeitsspeicher ausgeschöpft ist.

Diese verbesserte Methode wäre für all jene Bäume möglich, für die eine Adressierung konstanter Länge der einzelnen Knoten existiert. Der zugrundeliegende Graph darf dafür zum Beispiel nicht stark zusammenhängend sein.

Während sich also das Modell für die Halbachsen für höhere Dimensionen verallgemeinern lässt, ist die Situation für die Parameter α_i aus der Gleichung 6.2 komplizierter. Aufgrund der rekursiven Struktur der Kämmen und der Markoveigenschaft der Irrfahrten ist anzunehmen, dass sich diese rekursive Struktur auch auf die Aggregationscluster überträgt. Das heißt wir erwarten, dass eine geeignete Projektion des Aggregationsclusters von K_d auf einen Unterraum niedrigerer Dimension d' die selbe Gestalt annimmt, wie der Aggregationscluster des Kammes $K_{d'}$.

Im folgenden stellen wir uns K_d wieder in \mathbb{Z}^d eingebettet vor. Die Koordinatenachsen von \mathbb{Z}^d bezeichnen wir mit $\epsilon_1, \dots, \epsilon_d$, wobei die Nummerierung der Koordinatenachsen der Konstruktionsreihenfolge im Kamm folgt. Jeder Unterbaum $K_{d'}$ von K_d ist als spannender Baum in das, von den Achsen $\epsilon_1, \dots, \epsilon_{d'}$ aufgespannte, Gitter $\mathbb{Z}^{d'}$ eingebettet.

Bezeichne $A_d(n)$ die Folge von Aggregationsclustern auf dem Graphen K_d . Eine Projektion des Aggregationsclusters $A_d(n)$ auf das von den Achsen ϵ_{d-1} und ϵ_d aufgespannte Gitter ergibt, wie nach den obigen Überlegungen zu erwarten, wieder eine Menge die durch die Gleichung 6.1 mit den Parametern $\alpha \approx 1$ und $\beta \approx \frac{1}{2}$ beschrieben wird.

Wir betrachten nun beliebige zweidimensionale Schnitte von $A_d(n)$ durch den Ursprung. Es zeigt sich, dass die einfache Verallgemeinerung des zweidimensionalen Modells aus Gleichung (6.2) nicht zutrifft. Allerdings lässt sich der Rand der Schnitte

6.4. KÄMME IN HÖHEREN DIMENSIONEN

K_2	K_3	K_4	K_5	K_6	
				(1,2) 0.88 (1,3) 0.796 (1,4) 0.537 (1,5) 0.307 (1,6) 0.138	$\frac{8}{10}$ $\frac{6}{10}$ $\frac{4}{10}$ $\frac{2}{10}$ $\frac{1}{10}$
			(1,2) 0.84 (1,3) 0.588 (1,4) 0.318 (1,5) 0.15	(2,3) 0.773 (2,4) 0.576 (2,5) 0.317 (2,6) 0.146	$\frac{6}{8}$ $\frac{4}{8}$ $\frac{2}{8}$ $\frac{1}{8}$
		(1,2) 0.612 (1,3) 0.342 (1,4) 0.154	(2,3) 0.685 (2,4) 0.376 (2,5) 0.182	(3,4) 0.661 (3,5) 0.373 (3,6) 0.178	$\frac{4}{6}$ $\frac{2}{6}$ $\frac{1}{6}$
	(1,2) 0.499 (1,3) 0.263	(2,3) 0.539 (2,4) 0.267	(3,4) 0.545 (3,5) 0.269	(4,5) 0.53 (4,6) 0.257	$\frac{2}{4}$ $\frac{1}{4}$
(1,2) 0.5	(2,3) 0.53	(3,4) 0.497	(4,5) 0.501	(5,6) 0.486	$\frac{1}{2}$

Tabelle 6.5: Zweidimensionale Schnitte von K_2, \dots, K_6

wieder durch eine Gleichung der Form 6.1 beschreiben. Das heißt für einen Schnitt bezüglich der Achsen ϵ_i und ϵ_j mit $i > j$ gilt

$$\left(\frac{x_i}{h_i(n)}\right)^{\eta_{i,j}} + \left(\frac{x_j}{h_j(n)}\right)^{\xi_{i,j}} = 1.$$

h_i bezeichnet hier wieder die Halbachse von $A_d(n)$ bezüglich der Koordinatenachse ϵ_i . Für Parameter $\eta_{i,j}$ gilt dabei jeweils

$$\eta_{i,j} \approx 1.$$

Die Werte für $\xi_{i,j}$ sind aus Tabelle 6.5 zu entnehmen. Die Einträge sind dabei von der Form $(i, j) \quad \xi_{i,j}$. Dabei werden die Schätzwerte nach rechts oben (also für Kämmen höherer Dimension sowie Schnitte mit kleineren Koordinaten Indices) jeweils immer unsicherer. In der letzten Spalte ist eine Vermutung für die $\xi_{i,j}$ angegeben.

Für allgemeinere Schnitte, die nicht durch den Ursprung verlaufen, ist es schwieriger eine schlüssige Aussage zu treffen. Wie schon erwähnt sind beliebige Schnitte durch die Achsen ϵ_{d-1} und ϵ_d isomorph zum zweidimensionalen Kamm, und das Wachstum der Aggregationscluster verhält sich dementsprechend auch gleich (siehe dazu Abschnitt 6.2.1). Schnitte durch andere Achsenpaare sind allerdings nicht einmal mehr zusammenhängende Graphen. Hinzu kommt als weitere Schwierigkeit wiederum das geringe Wachstum des Clusters in Richtungen ϵ_i für kleine Indices i .

Die Abbildung 6.10 zeigt einige $(\epsilon_1, \epsilon_3) = (x, z)$ -Schnitte des Aggregationsclusters des Graphen K_3 , mit Werten für y zwischen 0 und 120. Als Modell des Randes des Clusters nehmen wir für jedes y eine Gleichung der Form:

$$\left(\frac{x}{h_x(y)}\right) + \left(\frac{z}{h_z(y)}\right)^{\xi_y} = 1 \tag{6.4}$$

Die Halbachsen sind dabei schon bekannt und ergeben sich aus ausprechenden Schnitten durch den Ursprung. Der Exponent ξ_y scheint mit wachsendem y zuzunehmen. Die

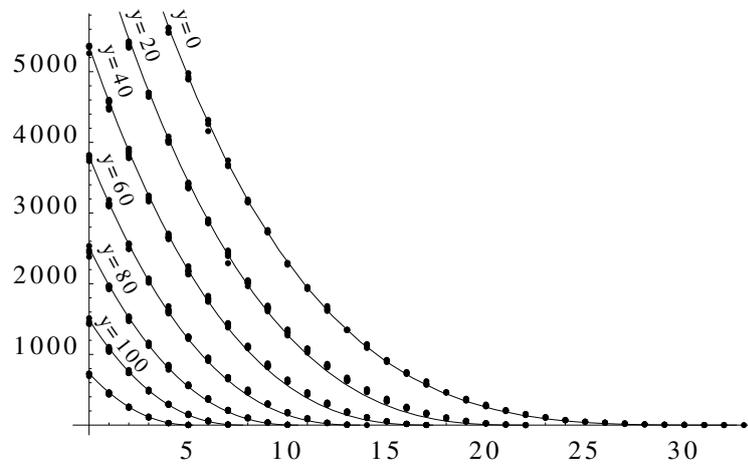


Abbildung 6.10: (x, z) -Schnitte von K_3

Werte aus Abbildung 6.10 lauten wie folgt:

$$\xi_0 = 0.264$$

$$\xi_{20} = 0.35$$

$$\xi_{40} = 0.38$$

$$\xi_{60} = 0.4$$

$$\xi_{80} = 0.42$$

$$\xi_{100} = 0.44$$

$$\xi_{120} = 0.46$$

Allerdings ist nicht klar, inwiefern sich diese Werte schon stabilisiert haben. Es wäre auch hier wünschenswert, Simulationsdaten weit größerer Cluster zur Verfügung zu haben, um gesichertere Aussagen treffen zu können.

Kapitel 7

Jüngere Entwicklungen

Nach dem Erfolg von Lawler, Bramson und Griffeath wurde, neben den bereits erwähnten Verfeinerungen von Satz 1, iDLA auch für andere Graphen als \mathbb{Z}^d untersucht.

7.1 Gruppen mit exponentiellem Wachstum

Blachère und Brofferio erhalten in [4] ein zu Satz 1 analoges Resultat für iDLA auf dem Cayley Graphen von Gruppen mit exponentiellem Wachstum. Zunächst benötigen wir allerdings eine Reihe von Definitionen.

Sei Γ eine endlich erzeugte Gruppe, und S eine endliche, symmetrische Menge von Erzeugern von Γ (d.h.: $\Gamma = \langle S \rangle$ und $S = S^{-1}$). Dann lässt sich jedes Element von Γ als Wort über dem Alphabet S schreiben.

Definition 27. Sei Γ eine endlich erzeugte Gruppe mit Erzeugermenge S . Der Graph mit Knotenmenge Γ und den Kanten

$$x \sim y \iff x^{-1}y \in S$$

für $x, y \in \Gamma$, heißt *Cayley-Graph* von Γ bezüglich der erzeugender Menge S .

Definition 28. Die *Wortmetrik* $|x|_{w,S}$ auf Γ ist definiert als die Länge des kürzesten Wortes über S , das x darstellt:

$$|x|_{w,S} = \min_n \{ \exists x_1, \dots, x_n \in S \text{ mit } x = x_1 x_2 \cdots x_n \}$$

Bemerkung 21. Die Wortmetriken bezüglich verschiedener erzeugenden Mengen sind vergleichbar. Sind also S und T endliche, symmetrische Erzeuger von Γ , dann existiert eine Konstante C sodass gilt

$$\frac{1}{C} |x|_{w,T} \leq |x|_{w,S} \leq C |x|_{w,T}.$$

Diese Bedingung ist erfüllt, da sich jedes Element in S als Wort in den Buchstaben von T schreiben lässt und umgekehrt. Als Konstante C wählt man die maximale Länge dieser Darstellungen.

Im Folgenden bezeichnen wir die Wortmetrik nur noch mit $|x|_w$, wenn keine Verwechslungsgefahr für die erzeugende Menge gegeben ist.

Die Kugeln bezüglich der Wortmetrik bezeichnen wir mit

$$\mathcal{B}_w(n) := \{x \in \Gamma : |x|_w \leq n\}.$$

Definition 29. Sei Γ eine endlich erzeugte Gruppe. Γ ist von *exponentiellem Wachstum*, wenn eine Konstante $c > 0$ existiert, sodass gilt

$$|\mathcal{B}_w(n)| \geq \exp(cn), \quad \forall n \geq 0.$$

Bemerkung 22. Aufgrund der Vergleichbarkeit der einzelnen Wortmetriken ist exponentielles Wachstum eine Gruppeneigenschaft, und nicht von der Wahl einer speziellen erzeugenden Menge abhängt.

Definition 30. Gegeben sei ein Wahrscheinlichkeitsmaß μ auf Γ mit $\text{supp}(\mu) = S$ und $\mu(s) = \mu(s^{-1})$, und eine Folge Y_i von unabhängigen identisch μ -verteilten Zufallsvariablen, dann definiert der Prozess $X(0) = x$ und

$$X(k) = x Y_1 Y_2 \cdots Y_k, \quad k \geq 1$$

eine irreduzible, symmetrische Irrfahrt mit Startpunkt x auf der Gruppe Γ .

Auf dieser Struktur lässt sich nun analog zum Zahlengitter internal Diffusion-Limited Aggregation studieren. Wir setzen dazu den Aggregationskeim $A(0) = \{e\}$ und betrachten eine Folge von unabhängigen identisch verteilten Irrfahrten X^i mit Startpunkt $x = e$. e bezeichnet dabei das neutrale Element in der Gruppe Γ .

Sei die Stopzeit τ_y die Zeit des ersten Eintreffens in y , also

$$\tau_y = \inf\{k \geq 0 : X(k) = y\}.$$

Die Trefferwahrscheinlichkeit von y bei Start in x bezeichnen wir als

$$F(x, y) = \mathbb{P}_x[\tau_y < \infty].$$

Mit Hilfe der Funktion F können wir nun eine an die iDLA-Cluster adaptierte Metrik definieren.

Definition 31. Seien $x, y \in \Gamma$, dann definiert

$$d(x, y) := -\log F(x, y)$$

eine Metrik auf Γ . $d(\bullet, \bullet)$ heißt *Trefferdistanz*.

Wir müssen noch überprüfen ob die Funktion $d(\bullet, \bullet)$ tatsächlich die Eigenschaften einer Metrik besitzt.

Lemma 13. *Die Trefferdistanz $d(\bullet, \bullet)$ ist eine linksinvariante Metrik auf Γ .*

Beweis. $F(\bullet, \bullet)$ ist eine Wahrscheinlichkeit, und somit gilt $F(x, y) \leq 1$. Das bedeutet aber, dass $d(x, y)$ nicht negativ ist.

Die Symmetrie und Linksinvarianz von $d(\bullet, \bullet)$ folgt direkt aus der Symmetrie der Irrfahrt und der Linksinvarianz von $F(\bullet, \bullet)$.

Sei $\tau'_y = \inf\{k \geq 1 : X(k) = y\}$. Die Gruppe Γ hat exponentielles Wachstum, also ist die Irrfahrt X transient, und es gilt $\forall x \neq y$

$$1 > \mathbb{P}_x[\tau'_x < \infty] \geq \mathbb{P}_x[\tau_y < \infty]\mathbb{P}_y[\tau_x < \infty] = F(x, y)F(y, x) = F(x, y)^2.$$

Offensichtlich gilt $F(x, x) = 1$, also ist

$$x = y \iff F(x, y) = 1 \iff d(x, y) = 0.$$

Schließlich folgt die Dreiecksungleichung aus

$$\mathbb{P}_x[\tau_z < \infty] \geq \mathbb{P}_x[\tau_y < \infty]\mathbb{P}_y[\tau_z < \infty].$$

Das ist äquivalent zu

$$\log \mathbb{P}_x[\tau_z < \infty] \geq \log \mathbb{P}_x[\tau_y < \infty] + \log \mathbb{P}_y[\tau_z < \infty],$$

und somit zur Dreiecksungleichung

$$d(x, z) \leq d(x, y) + d(y, z).$$

□

Mit Hilfe der Trefferdistanz definieren wir nun neue Kugeln

$$\mathcal{B}_x(n) = \{y \in \Gamma : d(x, y) \leq Kn\}, \tag{7.1}$$

wobei $K = \max\{d(e, s) : s \in S\}$ gesetzt wird.

Die Konstante K wird benötigt um sicherzustellen dass die Kugeln $\mathcal{B}_x(n)$ und $\mathcal{B}_x(n+1)$ passend ineinander verschachtelt sind.

Wenn wir den *äußeren Rand* von $\mathcal{B}_x(n)$ definieren als

$$\partial\mathcal{B}_x(n) = \{z \notin \mathcal{B}_x(n) : \exists y \in \mathcal{B}_x(n) \text{ mit } y^{-1}x \in S\},$$

dann gilt $\partial\mathcal{B}_e(n) \subset \mathcal{B}_e(n+1)$. Das ist leicht einzusehen, da für $x \in \partial\mathcal{B}_e(n)$ und $y \in \mathcal{B}_e(n)$ mit $y^{-1}x \in S$ gilt

$$\begin{aligned} d(e, x) &\leq d(e, y) + d(y, x) = d(e, y) + d(e, y^{-1}x) \\ &\leq Kn + \max\{d(e, z) : z \in S\} \leq K(n+1). \end{aligned}$$

Bemerkung 23. Da die Funktion $F(e, x)$ proportional zur Greenschen Funktion $G(e, x)$ ist, gilt für die Kugeln

$$\mathcal{B}_e(n) = \{y \in \Gamma : G(e, y) \geq N\}$$

für eine geeignete Konstante N .

Nach diesen Vorarbeiten lässt sich nun das Resultat aus [4] formulieren.

Satz 11. *Sei $\{A(n)\}_{n \geq 0}$, $A(0) = \{e\}$ Folge von internal DLA Clustern auf einer endlich erzeugten Gruppe Γ mit exponentiellem Wachstum, bezüglich einer symmetrischen Irrfahrt mit einer endlichen Menge von Inkrementen. Dann gilt für alle Konstanten $\mathcal{C}_1 > 3/K$ und $\mathcal{C}_0 > 2$,*

$$\mathbb{P}[\mathcal{B}_e(n - \mathcal{C}_1 \ln n) \subset A(V(n)) \subset \mathcal{B}_e(n + \mathcal{C}_0 \sqrt{n}), \text{ für unendlich viele } n] = 1.$$

Dabei sind $\mathcal{B}_e(n)$ und K wie in (7.1) definiert, und

$$V(n) = |\mathcal{B}_e(n)|$$

ist das Volumen der Kugeln bezüglich der Trefferdistanz.

7.2 Rotor Router

Jim Propp definierte mit dem Rotor-Router Modell ein deterministisches Analogon zu klassischen Irrfahrten. Hierbei wandert ein Partikel jeweils in eine vom aktuell besetzten Knoten spezifizierte Richtung (den Rotor). Nach jedem dieser Routingschritte wird die Rotor-Richtung deterministisch modifiziert. Damit die Eigenschaften von einfachen Irrfahrten weitestgehend erhalten bleiben, durchläuft der Routing-Vektor in jedem Knoten der Grundmenge zyklisch alle möglichen Richtungen.

Levine [11] beschrieb ein zu iDLA analoges Wachstumsmodell auf Basis von Rotor-Routing.

Seien e_1, \dots, e_d die standard Basisvektoren von \mathbb{Z}^d . Dann definiert die Menge $\mathcal{E}_d = \{\pm e_1, \dots, \pm e_d\}$ die natürlichen Fortbewegungsrichtungen im d -dimensionalen Gitter. Wir definieren weiters eine beliebige Totalordnung $\epsilon_0 < \epsilon_1 < \dots < \epsilon_{2d-1}$, $\epsilon_i \in \mathcal{E}_d$.

Jeder Zustand des Rotor-Router Automaten lässt sich dann als Tripel (A, x, v) beschreiben. Dabei ist die Teilmenge $A \subset \mathbb{Z}^d$ (endlich und zusammenhängend) der aktuelle Cluster, $x \in \mathbb{Z}^d$ beschreibt die Position des Partikels und $v : \mathbb{Z}^d \rightarrow \mathcal{E}_d$ definiert die Routing-Richtung in jedem Gitterpunkt.

Der Anfangszustand des Automaten wird festgelegt mit $A = \{0\}$, $x = 0$ und $v(y) = \epsilon_0$, $\forall y \in \mathbb{Z}^d$.

Die Abbildung $t(A, x, v) = (A', x', v')$ definiert die Übergangstransformation des Automaten. Mit

$$A' = \begin{cases} A & \text{für } x + v(x) \in A \\ A \cup \{x + v(x)\} & \text{sonst,} \end{cases}$$

$$x' = \begin{cases} x + v(x) & \text{für } x + v(x) \in A \\ 0 & \text{sonst,} \end{cases}$$

und die neuen Routing-Vektoren v' werden mittels

$$v'(y) = \begin{cases} \epsilon_{(k+1) \bmod 2d} & \text{für } x = y \text{ und } v(x) = \epsilon_k \\ v(y) & \text{sonst,} \end{cases}$$

berechnet.

Durch Iteration der Übergangsfunktion t erhalten wir eine aufsteigende Folge von Teilmengen A'_n . Beschränken wir uns nur auf die Zeitpunkte n_k , an denen das Partikel den jeweils aktuellen Cluster verlässt, so ist die dadurch induzierte Teilfolge $A_k = A'_{n_k}$ streng monoton wachsend.

Simulationen legen für diese Folge, wie im Falle von iDLA, eine Konvergenz gegen euklidische Kugeln nahe. Levine und Peres beweisen diese Vermutung in [12] für \mathbb{Z}^d

und ein etwas verallgemeinertes Modell, und geben auch eine Abschätzung für die Konvergenzgeschwindigkeit.

Sei $R \triangle S$ die symmetrische Differenz der Mengen R und S . Weiters, für eine Teilmenge $A \subset \mathbb{Z}^d$, bezeichne A^\square die Vereinigung von d -dimensionalen Einheitswürfeln mit Mittelpunkten in den Punkten von A . Wir schreiben λ_d für das Lebesgue-Maß auf \mathbb{R}^d .

Satz 12. *Sei $\{A_n\}_{n \geq 1}$ eine Folge von Rotor-Router Aggregationsclustern in \mathbb{Z}^d , für eine beliebige Anfangskonfiguration von Rotoren. Dann gilt*

$$\lambda_d(n^{-1/d}A_n^\square \triangle \mathcal{B}) \longrightarrow 0 \quad \text{für } n \rightarrow \infty.$$

Hierbei bezeichnet \mathcal{B} die Einheitskugel in \mathbb{R}^d mit Zentrum im Ursprung.

Bemerkung 24. Dieses Ergebnis ist schwächer als Satz 1 für iDLA, da zum Beispiel „Löcher“ im Cluster oder dünne Filamente am Rand nicht ausgeschlossen werden, solange ihr Volumen klein ist im Vergleich zu n . Bei Computersimulationen des Modells zeigt der Rotor-Router-Cluster allerdings nur sehr geringe Abweichungen von der Kugelform. Es wird deshalb angenommen, dass es möglich sein sollte Satz 12 noch weiter zu verbessern.

7.3 Ausblick

Obwohl internal Diffusion-Limited Aggregation von Diaconis und Fulton [5] eigentlich nur als Beiprodukt beim Studium algebraischer Strukturen eingeführt wurde, hat sich das Modell in den letzten Jahren für die statistische Physik als fruchtbar erwiesen [6].

Die Struktur der durch iDLA erzeugten Aggregationscluster hat sich in allen bisher untersuchten Beispielen (Gitter, homogene Bäume und Cayleygraphen) für theoretische Untersuchungen zugänglich erwiesen. Dies steht im Gegensatz zu klassischer (externer) DLA, deren fraktale Struktur, trotz intensiver Untersuchungen in den letzten 25 Jahren, bis jetzt noch nicht bewiesen werden konnte.

Dieser Umstand mag auch mit der geringeren Komplexität von iDLA als Berechenbarkeitsmodell zusammenhängen, da wie zum Beispiel Moore und Machta [16] vermuten,

sich daraus auch Rückschlüsse auf die Komplexität eines physikalischen Systems ziehen lassen.

Die Form aller bisher bekannten iDLA-Cluster entspricht Kugeln bezüglich der in Definition 31 eingeführten Trefferdistanz. Diese ist für die einfache Irrfahrt auf \mathbb{Z}^d proportional zur euklidischen Metrik, und auf den homogenen Bäumen \mathbb{T}_q proportional zur Graphenmetrik.

Ob dies auch auf weniger reguläre Strukturen zutrifft, ließe sich durch Untersuchung der Trefferdistanz auf Kämmen und Vergleich mit den Simulationsresultaten aus Abschnitt 6.2 klären. Dazu wäre eine explizite Berechnung der Greenfunktion $G(x, y)$ für beliebige Paare $x, y \in K_2$ erforderlich. Ähnliche Berechnungen, allerdings nicht für beliebige Paare, wurden, zur Abschätzung der asymptotischen Fluchtrate von Irrfahrten auf den Kämmen, von Bertacchi und Zucca [1] durchgeführt.

Abbildungsverzeichnis

2.1	DLA Simulation mit 10000 Teilchen in \mathbb{Z}^2	9
3.1	iDLA Simulation von 10000 Partikel in \mathbb{Z}^2	14
4.1	Komparator-Gatter	54
4.2	Schrittfolge zur Simulation eines Komparator-Gatters	56
5.1	Beispiele von Cluster-Visualisierungen	63
5.2	Anzahl der Simulationsschritte	66
6.1	Radius der Aggregationscluster von \mathbb{T}_3	68
6.2	Varianzen Clusterrandpunkte von \mathbb{T}_3	69
6.3	Aggregationscluster des zweidimensionalen Kammes	71
6.4	Halbachsen des zweidimensionalen Kammes	73
6.5	Aggregationscluster von K_2	74
6.6	iDLA-Cluster auf K_2 mit inhomogener Irrfahrt	75
6.7	Halbachsenparameter für Biased-Irrfahrt auf K_2	76
6.8	iDLA-Cluster auf S_4 und S_7	78
6.9	Halbachsen des 4-dimensionalen Kammes K_4	81
6.10	(x, z) -Schnitte von K_3	86

Tabellenverzeichnis

6.1	Vergleich der Simulationen am Kamm (100000 Partikel)	72
6.2	Clusterparameter von S_3 und S_7	78
6.3	Halbachsenparameter von K_2, \dots, K_5	80
6.4	Absolute Fehler der Halbachsenexponenten von K_2, \dots, K_6	82
6.5	Zweidimensionale Schnitte von K_2, \dots, K_6	84

Literaturverzeichnis

- [1] BERTACCHI, D. und ZUCCA, F. (2003). Uniform asymptotic estimates of transition probabilities on combs. *J. Aust. Math. Soc.* **75** 325-353.
- [2] BLACHÈRE, S. (2000). Agrégation limitée par diffusion interne et temps de coupure sur les groupes discrets à croissance polynomiale. *Doctorat De L'Université Paul Sabatier*.
- [3] BLACHÈRE, S. (2001). Logarithmic fluctuations for the Internal Diffusion Limited Aggregation. *arXiv:math/0111253*.
- [4] BACHÈRE, S. und BROFFERIO, S. (2005). Internal Diffusion Limited Aggregation on discrete groups having exponential growth. *arXiv:math/0507582*.
- [5] DIACONIS, P. und FULTON, W. (1991). A growth model, a game, an algebra, Lagrange inversion, and characteristic classes. *Rend. Sem. Mat. Univ. Pol. Torino* **49** 95-119.
- [6] GRAVNER, J. und QUASTEL, J. (2000). Internal DLA and the Stefan problem. *Ann. Probab.* **28**, no. 4 1528-1562.
- [7] LAWLER, G. (1991). Intersections of Random Walks. *Birkäuser, Boston*.
- [8] LAWLER, G. (1995). Introduction to stochastic processes. *Chapman & Hall*.
- [9] LAWLER, G., BRAMSON, M. und GRIFFEATH, D. (1992). Internal diffusion limited aggregation. *Ann. Probab.* **20** 2117-2140.
- [10] LAWLER, G. (1995). Subdiffusive fluctuations for internal diffusion limited aggregation. *Ann. Probab.* **23** 71-86.

- [11] LEVINE, L. (2002). The rotor-router model. *Harvard University senior thesis*, *arXiv:math/0409407*.
- [12] LEVINE, L. und PERES, Y. (2005). Spherical Asymptotics for the Rotor-Router Model in \mathbb{Z}^d . *arXiv:math/0503251*.
- [13] LYONS, R. und PERES, Y. Probability on trees and networks. (*Preprint*) <http://mypage.iu.edu/~rdlyons/prbtree/prbtree.html>
- [14] MAKOTO MATSUMOTO und TAKUJI NISHIMURA (1998) Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation* **Vol. 8, No. 1** 3-30.
- [15] MOORE, C. (1997) Majority-Vote Cellular Automata, Ising Dynamics, and P-Completeness. *J. Stat. Phys.* **88 (3/4)** 795-805.
- [16] MOORE, C. und MACHTA, J. (2000). Internal Diffusion-Limited Aggregation: Parallel Algorithms and Complexity. *J. Stat. Phys.* **99** 661-690.
- [17] NAGNIBEDA, T. und WOESS, W. (2002). Random walks on trees with finitely many cone types. *J. Theoret. Probab.* **15** 399-438.
- [18] PAPADIMITRIOU, C. H. (1994). Computational Complexity. *Addison-Wesley*
- [19] SIPSER, M. (1996) Introduction to the theory of computation. *PWS Publishing Company*
- [20] SPITZER, F. (1976) Principles of Random Walk. *Springer-Verlag*
- [21] TOLMAN, S. und MEAKIN, P. (1989) Off-lattice and hypercubic-lattice models for diffusion-limited aggregation in dimensionalities 2-8. *Phys. Rev. A* **40** 428.
- [22] WITTEN, T. A. und SANDER, L. M. (1981). Diffusion-limited aggregation, a kinetic critical phenomenon. *Phys. Rev. Lett.* **47** 1400.
- [23] WOESS, W. (2000). Random walks on Infinite Graphs and Groups. *Cambridge Univ. Press, Cambridge*