



Glauber-Dynamik für das Ising-Modell

Seminararbeit

“Moderne Anwendungen der Theorie der Markov-Ketten”

erstellt von
Xiaorui Wang

Betreuer: Prof. Dr. Wolfgang König
Priv.-Doz. Dr. Konstantin Fackeldey

Technische Universität Berlin
Fakultät II: Mathematik und Naturwissenschaften, Institut für Mathematik
Sommersemester 2016

Inhaltsverzeichnis

1	Einleitung	2
1.1	Motivation und Geschichte	2
1.2	Grundlagen des Ising-Modells	3
1.3	Glauber-Dynamik auf dem Ising-Modell	4
2	Pfadkopplung	6
2.1	Transportationsmetrik	7
2.2	Pfade und Pfadmetrik	8
2.3	Satz von Bublely und Dyer	8
3	Obere Grenzen für Mischzeiten im Ising-Modell	10
3.1	Schnelles Mischen bei hoher Temperatur	10
4	Die Mischzeit des Ising-Modells auf bestimmten Graphen	12
4.1	Vollständige Graphen	13
4.2	Ring-Graphen	13
4.3	b-äre Baum-Graphen	14

1 Einleitung

1.1 Motivation und Geschichte

Lange Zeit konnte Magnetismus nur phänomenologisch beschrieben werden ohne die physikalischen Mechanismen dahinter zu verstehen. Im späten 19. und frühen 20. Jahrhundert gab es zahlreiche Fortschritte, die neue Ansätze ermöglichten. Insbesondere ist dabei zu erwähnen das verbesserte Verständnis von Materie (Atommodelle von Rutherford [8] und Bohr [1], frühe Quantenmechanik, Spin) als auch komplexer Systeme aus vielen Teilchen (Beiträge zur statistischen Physik von Maxwell [6], Boltzmann, Gibbs [3]). Darauf aufbauend entwickelte Wilhelm Lenz 1920 ein Modell zur Beschreibung von Ferromagnetismus, d.h. des Magnetismus, der durch die Ausrichtungen der magnetischen Momente der Atome eines Festkörpers erzeugt wird. In diesem Modell werden die Atome eines Körpers als Knoten in einem Gitter dargestellt. Jedem Knoten wird dann ein Wert ± 1 zugewiesen, was der Orientierung des magnetischen Moments des Atoms in eine Richtung darstellt. Lenzs Student Ernst Ising, nachdem es heutzutage benannt ist, untersuchte dieses Modell in seiner Doktorarbeit und löste es im eindimensionalen Fall [4]. In mehr Dimensionen ist das Ising-Modell deutlich komplizierter. Lars Onsager konnte 1944 das zweidimensionale Gitter lösen [7] und in mehr als vier Dimensionen wird heutzutage die sogenannte "Mean-Field"-Theorie benutzt. Anstatt das Ising-Modell auf einem Gitter zu betrachten, kann man allerdings auch andere Arten von Graphen als Grundlage benutzen. Dies werden wir in dieser Arbeit für einige konkrete Beispiele machen. Dabei orientieren wir uns grob an den entsprechenden Abschnitten des Buchs [5].

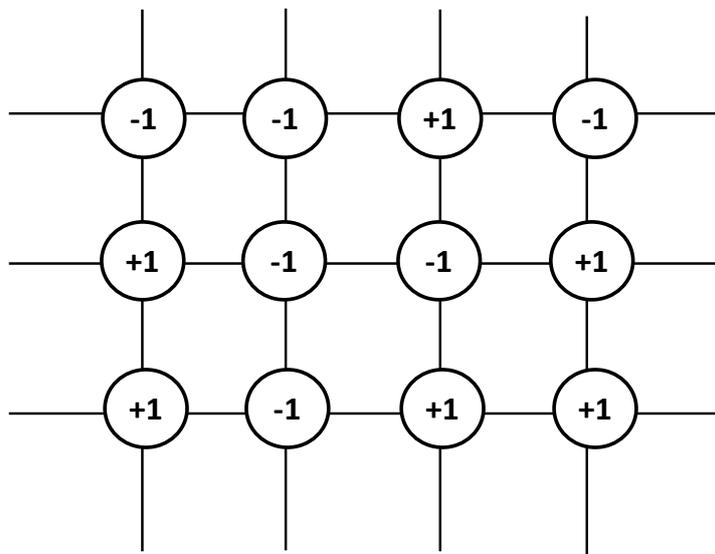


Abbildung 1: 2-dimensionales Ising-Gitter

1.2 Grundlagen des Ising-Modells

Definition 1.1. Sei $G = (V, E)$ ein Graph, V die Vertexmenge mit $N = |V| < \infty$ und E die Kantenmenge. Sei $\sigma \in \Omega = \{-1, +1\}^V$ eine **Konfiguration**. Der Wert $\sigma(v) \in \{-1, +1\}$ mit $v \in V$ heißt der **Spin** auf v .

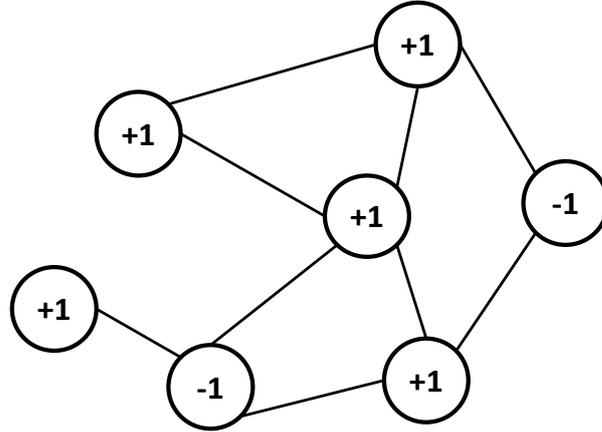


Abbildung 2: Allgemeinerer Ising-Graph. In dieser Konfiguration hat er die Energie $H(\sigma) = -(1 + 1 + 1 - 1 - 1 + 1 - 1 - 1 - 1) = 1$.

Das nächster Nachbar **Ising-Modell** ist eine Wahrscheinlichkeitsverteilung auf $\Omega := \{-1, +1\}^V$ gegeben durch eine sogenannte **Gibbs-Verteilung**

$$\mu(\sigma) := \frac{e^{-\beta H(\sigma)}}{Z(\beta)}. \quad (1)$$

Dabei ist $\sigma \in \Omega = \{-1, +1\}^V$ eine Konfiguration,

$$H(\sigma) := - \sum_{\substack{v, w \in V \\ v \sim w}} \sigma(v)\sigma(w) \quad (2)$$

die Energie der Konfiguration σ , und die Funktion $Z(\beta)$ ein Normierungsfaktor für die Wahrscheinlichkeitsverteilung μ , gegeben durch

$$Z(\beta) := \sum_{\sigma \in \Omega} e^{-\beta H(\sigma)}. \quad (3)$$

Der Parameter $\beta \geq 0$ ist die Gewichtung der Energiefunktion. In der physikalischen Interpretation ist β das Inverse der Temperatur. Verschiedene Werte dieses Parameters verändern das Verhalten des betrachteten Modells stark.

Es gibt drei qualitativ unterschiedliche Fälle:

$\beta = 0$, **unendlich hohe Temperatur**. In diesem Fall ist $e^{-\beta H(\sigma)} = 1$. Daraus folgt $Z(\beta) = |\Omega|$, so dass $\mu(\sigma) = \frac{1}{Z(\beta)} = \frac{1}{|\Omega|}$ gleichverteilt auf Ω ist. Alle Konfigurationen sind also gleichwahrscheinlich. Die Spins $+1$ und -1 sind gut gemischt.

$\beta \approx \beta_k$, **kritische Temperatur**. bei endlicher Temperatur sind Konfigurationen mit geringerer Energie in der Gibbs-Verteilung stärker gewichtet. Es bilden sich Gebiete in denen benachbarte Vertizes den selben Spin tragen, da solche Konfigurationen im Allgemeinen eine niedrigere Energie haben. Die kritische Temperatur ist der Wert von β , bei dem sich das Verhalten des Systems qualitativ ändert (siehe z.B. Schranken für Mischzeiten in Abschnitt 4).

$\beta \rightarrow \infty$, **sehr niedrige Temperatur**. Bei extrem niedriger Temperatur sind fast alle Spins gleich orientiert, was der geringsten möglichen Energie entspricht.

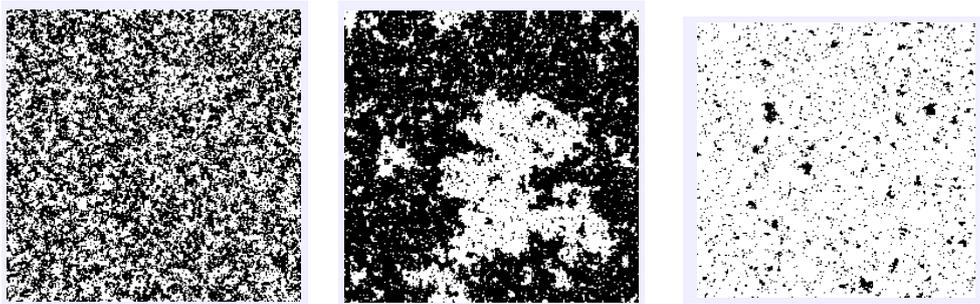


Abbildung 3: links: β klein (hohe Temperatur) – mitte: $\beta \approx \beta_k$ (kritische Temperatur) – rechts: β groß (niedrige Temperatur). Bilder entnommen aus [5].

1.3 Glauber-Dynamik auf dem Ising-Modell

Im vorherigen Abschnitt haben wir davon gesprochen, dass sich bei bestimmten Temperaturen Gebiete gleichen Spins bilden oder Spins gemischt werden. Wir müssen nun überlegen wie wir eine solche zeitliche Änderung (“Dynamik”) eines Systems mathematisch modellieren. Für das Ising-Modell bietet sich die sogenannte Glauber-Dynamik an.

Eine **Glauber-Dynamik** für eine Wahrscheinlichkeitsverteilung μ ist eine Markovkette $(X_t)_{t \in \mathbb{N}}$ mit Zustandsraum $\Omega = \{-1, +1\}^V$ so dass μ ihre stationäre Verteilung ist. Wir konstruieren nun die Übergangswahrscheinlichkeit von einer bestimmten Konfiguration σ zu einer anderen Konfiguration σ' in der Glauber-Dynamik.

Bedingte Gibbs-Verteilung. Sei ein Vertex $w \in V$ gleichverteilt zufällig gewählt. Dann wähle σ' gemäß einer bedingten Wahrscheinlichkeitsverteilung, nämlich der Gibbs-Verteilung μ eingeschränkt auf Konfigurationen, die mit σ in allen Vertizes außer w übereinstimmen. Sei für alle $\sigma \in \Omega$ und $w \in V$

$$\Omega(\sigma, w) = \{\sigma' \in \Omega \mid \sigma'(v) = \sigma(v) \forall v \neq w\}. \quad (4)$$

Damit schreiben für diese bedingte Wahrscheinlichkeitsverteilung

$$\mu(\sigma' \mid \Omega(\sigma, w)) = \begin{cases} \frac{\mu(\sigma')}{\mu(\Omega(\sigma, w))} = \frac{\mu(\sigma')}{\mu(\sigma) + \mu(\sigma')} = \frac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} & \sigma' \in \Omega(\sigma, w), \\ 0 & \sigma' \notin \Omega(\sigma, w). \end{cases} \quad (5)$$

Dieser Ausdruck kann weiter umgeschrieben werden, so dass man sieht, dass $\mu(\sigma' \mid \Omega(\sigma, w))$ nur von den Spinwerten auf w und seinen Nachbarvertizes abhängt. Man kann die Energie H aufteilen um den relevanten Teil zu isolieren. Es gilt

$$\begin{aligned} H(\sigma) &= - \sum_{\substack{u, v \in V \\ u \sim v}} \sigma(u)\sigma(v) \\ &= - \underbrace{\sigma(w) \sum_{\substack{u \in V \\ u \sim w}} \sigma(u)}_{S(\sigma, w)} - \underbrace{\sum_{\substack{v \in V \\ v \neq w}} \sum_{\substack{u \in V \\ u \sim v \\ u \neq w}} \sigma(u)\sigma(v)}_{S'(\sigma, w)} \end{aligned} \quad (6)$$

Damit folgt $S(\sigma, w) = -S(\sigma', w)$ und $S'(\sigma, w) = S'(\sigma', w)$. Für die obige Wahrscheinlichkeitsverteilung hat man dann

$$\begin{aligned} \mu(\sigma' \mid \Omega(\sigma, w)) &= \frac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} \mathbb{1}_{\sigma' \in \Omega(\sigma, w)} \\ &= \frac{e^{\beta(S(\sigma', w) + S'(\sigma', w))}}{e^{\beta(S(\sigma, w) + S'(\sigma, w))} + e^{\beta(S(\sigma', w) + S'(\sigma', w))}} \mathbb{1}_{\sigma' \in \Omega(\sigma, w)} \\ &= \frac{e^{\beta S(\sigma', w)} e^{S'(\sigma', w)}}{e^{-\beta S(\sigma', w)} e^{S'(\sigma', w)} + e^{\beta S(\sigma', w)} e^{S'(\sigma', w)}} \mathbb{1}_{\sigma' \in \Omega(\sigma, w)} \\ &= \frac{e^{\beta S(\sigma', w)}}{e^{\beta S(\sigma', w)} + e^{-\beta S(\sigma', w)}} \mathbb{1}_{\sigma' \in \Omega(\sigma, w)}. \end{aligned} \quad (7)$$

Übergangswahrscheinlichkeit. Die Übergangswahrscheinlichkeit hat bei diesem Vorgehen allgemein die Form

$$\begin{aligned} p_{\sigma, \sigma'} &= \mathbb{P}(X_{t+1} = \sigma' \mid X_t = \sigma) \\ &= \sum_{w \in V} \underbrace{\mathbb{P}(w)}_{=1/|V|} \underbrace{\mathbb{P}(X_{t+1} = \sigma' \mid X_t = \sigma, w)}_{=\mu(\sigma' \mid \Omega(\sigma, w))} \end{aligned} \quad (8)$$

woraus sich durch einsetzen ergibt

$$\begin{aligned}
p_{\sigma,\sigma'} &= \frac{1}{|V|} \sum_{w \in V} \frac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} \mathbb{1}_{\sigma' \in \Omega(\sigma,w)} \\
&= \frac{1}{|V|} \sum_{w \in V} \frac{e^{\beta S(\sigma',w)}}{e^{\beta S(\sigma,w)} + e^{-\beta S(\sigma',w)}} \mathbb{1}_{\sigma' \in \Omega(\sigma,w)}. \tag{9}
\end{aligned}$$

Es bleibt nun noch zu überprüfen, dass die Gibbs-Verteilung μ tatsächlich die stationäre Verteilung der Markovkette $(X_t)_{t \in \mathbb{N}}$ mit dieser Übergangswahrscheinlichkeit ist. Wir machen dies mit den *detailed balance equations*, d.h. wir zeigen, dass $\mu(\sigma)p_{\sigma,\sigma'} = \mu(\sigma')p_{\sigma',\sigma}$. Nach [5, Abschnitt 1.6, Prop. 1.19] ist dann μ stationär für $p_{\sigma,\sigma'}$.

$$\begin{aligned}
\mu(\sigma)p_{\sigma,\sigma'} &= \frac{e^{-\beta H(\sigma)}}{Z(\beta)} \frac{1}{|V|} \sum_{w \in V} \frac{e^{-\beta H(\sigma')}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} \mathbb{1}_{\sigma' \in \Omega(\sigma,w)} \\
&= \frac{e^{-\beta H(\sigma')}}{Z(\beta)} \frac{1}{|V|} \sum_{w \in V} \frac{e^{-\beta H(\sigma)}}{e^{-\beta H(\sigma)} + e^{-\beta H(\sigma')}} \mathbb{1}_{\sigma \in \Omega(\sigma',w)} \\
&= \mu(\sigma')p_{\sigma',\sigma}. \tag{10}
\end{aligned}$$

Außerdem ist die so konstruierte Markovkette reversibel. Sie ist allerdings nicht eindeutig, d.h. andere Algorithmen können zu anderen Markovketten führen, die jedoch auch die stationäre Verteilung μ haben. Siehe dazu z.B. die Seminarvorträge zu den Themen “Metropolis-Algorithmus” und “Gibbs-Sampling” vom 22.06.2016.

Offene Fragen. Wir haben jetzt eine Übergangswahrscheinlichkeit, die uns die Dynamik des Ising-Modells bestimmt. Damit können wir nun die Eigenschaften des Modells untersuchen. Insbesondere interessiert uns dabei die Mischzeit t_ε , bzw. ob wir obere und untere Schranken für sie finden können. Um dies beantworten zu können, müssen wir erst eine neue Technik einführen - die sogenannte *Pfadkopplung*.

2 Pfadkopplung

Die Kopplung von Markovketten ist eine sehr nützliche Technik, mit der unter anderem Schranken für Mischzeiten von Markovketten berechnet werden können. Hier benötigen wir die Pfadkopplung, um dann mit dem Satz von Bubley und Dyer (Satz 2.3) eine gewisse Metrik zu beschränken. Dies wird dann im folgenden Kapitel für Berechnungen im Ising-Modell benutzt. Zunächst müssen wir jedoch einige neue Begriffe einführen.

2.1 Transportationsmetrik

Abstand von Konfigurationen. Sei P eine Übergangsmatrix auf einem metrischem Raum (Ω, ρ) , wobei Ω ein Zustandsraum ist und der Abstand ρ erfüllt $\rho(x, y) \geq \mathbb{1}_{\{x \neq y\}}$, mit $x, y \in \Omega$.

Zur Anwendung auf das Ising-Modell wählen wir als Abstand zweier Konfigurationen die Anzahl der unterschiedlichen Vertizes. Man sieht dann, dass für $x, y, z \in \Omega$ gilt:

- $\rho(x, y) \geq 0$ und $\rho(x, y) = 0 \Leftrightarrow x = y$
- $\rho(x, y) = \rho(y, x)$
- $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$

Transportationsmetrik. Wir betrachten die Kopplung (X, Y) von zwei Markovketten mit den Startverteilungen μ und ν . Die *Transportationsmetrik* ρ_K zwischen μ und ν definieren wir als folgendes:

$$\rho_K(\mu, \nu) := \inf\{\mathbb{E}(\rho(X, Y)) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \quad (11)$$

Für den Fall, dass $\rho(x, y) = \mathbb{1}_{\{x \neq y\}}$ kann man sehen, dass dann folgt $\rho_K(\mu, \nu) = \|\mu - \nu\|_{TV}$, die Totalvariationsmetrik:

$$\begin{aligned} \rho_K(\mu, \nu) &= \inf\{\mathbb{E}(\rho(X, Y)) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \inf\{\mathbb{E}[\mathbb{1}_{\{X \neq Y\}}] : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \inf\{\mathbb{P}(X \neq Y) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \|\mu - \nu\|_{TV} \end{aligned} \quad (12)$$

Wenn wir die Kopplung (X, Y) als die Wahrscheinlichkeitsverteilung auf dem Produktraum $\Omega \times \Omega$ statt als Zufallsvariablen schreiben, dann ist

- $\mu = q(\cdot \times \Omega) = \sum_{y \in \Omega} q(\cdot, y)$
- $\nu = q(\Omega \times \cdot) = \sum_{y \in \Omega} q(y, \cdot)$

Daraus folgt

$$\begin{aligned} \rho_K(\mu, \nu) &= \inf\{\mathbb{E}(\rho(X, Y)) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \inf\left\{ \sum_{(x, y) \in \Omega \times \Omega} \rho(x, y) q(x, y) \mid \mu = q(\cdot \times \Omega), \nu = q(\Omega \times \cdot) \right\}. \end{aligned} \quad (13)$$

Aus einfachen topologische Überlegungen (vgl. [5], S. 190) kann man schließen, dass immer eine *optimale Verteilung* $q^*(x, y)$ existiert, für die das Infimum erreicht wird. Dies nutzen wir um den Satz von Bublely und Dyer zu beweisen.

2.2 Pfade und Pfadmetrik

Sei $(X_t)_{t \in \mathbb{N}}$ eine Markovkette und $G_0 = (V_0, E_0)$ ein Graph, dessen Vertexmenge $V_0 = \Omega = \{-1, +1\}^V$ der Zustandsraum des Ising-Modells ist. Außerdem wird jeder Kante $\{x, y\} \in E_0$ eine Länge $\ell(x, y)$ zugewiesen.

Definition 2.1. Sei für alle Kanten $\ell(x, y) \geq 1$. Ein **Pfad** (path) in Ω von x nach y ist definiert als eine Folge von Zuständen $\xi = (x_0, x_1, \dots, x_r)$, sodass $x = x_0, y = x_r$ und $\{x_{i-1}, x_i\} \in E_0$ eine Kante ist, für alle $i = 1, \dots, r - 1$. Die **Pfadlänge** von x nach y ist definiert als Summe der Kantenlängen der Kanten, die x und y verbinden, also $\sum_{i=1}^r \ell(x_{i-1}, x_i)$. Die **Pfadmetrik** auf Ω ist definiert als das Minimum aller möglichen Pfadlängen, d.h.

$$\rho(x, y) = \min\{\text{Länge von } \xi : \xi \text{ ein Pfad von } x \text{ nach } y\} \quad (14)$$

Satz 2.2. Wir nutzen die selbe Annahme wie oben, nämlich dass eine Kante zwischen x und y existiert und jede Kantenlänge $\ell(x', y') \geq 1$ ist, für beliebige $x', y' \in \Omega$. Dann gilt es:

$$\rho_{TV}(\mu, \nu) \leq \rho_K(\mu, \nu) \quad (15)$$

Beweis. Aus der Kantenlänge $\ell(x, y) \geq 1$, folgt

$$\rho(x, y) \geq \ell(x, y) \geq 1 \geq \mathbf{1}_{\{x \neq y\}} \quad (16)$$

Darauf folgt

$$\begin{aligned} \rho_{TV}(\mu, \nu) &= \inf\{\mathbb{E}(\mathbf{1}_{\{X \neq Y\}}) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \inf\left\{\sum_{(x,y) \in \Omega \times \Omega} \mathbf{1}_{\{x \neq y\}} q(x, y) \mid \mu = q(\cdot \times \Omega), \nu = q(\Omega \times \cdot)\right\} \\ &\leq \inf\left\{\sum_{(x,y) \in \Omega \times \Omega} \rho(x, y) q(x, y) \mid \mu = q(\cdot \times \Omega), \nu = q(\Omega \times \cdot)\right\} \\ &= \inf\{\mathbb{E}(\rho(X, Y)) : (X, Y) \text{ ist eine Kopplung von } \mu \text{ und } \nu\} \\ &= \rho_K(\mu, \nu) \end{aligned} \quad (17)$$

□

2.3 Satz von Bublely und Dyer

Satz 2.3. (Bublely und Dyer [2, 1997])

Angenommen für jede Kante $\{x, y\}$ existiert eine Kopplung (X_1, Y_1) zwischen den Verteilungen $P(x, \cdot)$ und $P(y, \cdot)$, sodass

$$\mathbb{E}_{x,y}(\rho(X_1, Y_1)) \leq \rho(x, y)e^{-\alpha} = \ell(x, y)e^{-\alpha} \quad (18)$$

Dann gilt für jede zwei Wahrscheinlichkeitsverteilungen μ und ν auf Ω

$$\rho_K(\mu P, \nu P) \leq e^{-\alpha} \rho_K(\mu, \nu) \quad (19)$$

Beweis. Seien $x, y \in \Omega$ zwei feste aber beliebige Konfigurationen. Wir suchen zunächst eine Beziehung zwischen ρ und ρ_K für diese. Sei (x_0, x_1, \dots, x_r) der minimale Pfad von x nach y ($x_0 = x, x_r = y$), d. h. $\rho(x, y) = \sum_{k=1}^r \ell(x_{k-1}, x_k)$. Dann können wir mit der Dreiecksungleichung und der Voraussetzung abschätzen:

$$\begin{aligned} \rho_K(P(x, \cdot), P(y, \cdot)) &\leq \sum_{k=1}^r \rho_K(P(x_{k-1}, \cdot), P(x_k, \cdot)) \\ &\leq \sum_{k=1}^r e^{-\alpha} \ell(x_{k-1}, x_k) \\ &= e^{-\alpha} \rho(x, y) \end{aligned} \quad (20)$$

Sei nun θ eine Kopplung von μP und νP mit

$$\rho_K(\mu P, \nu P) \leq \sum_{u, w \in \Omega} \rho(u, w) \theta(u, w) \quad (21)$$

und η eine optimale Kopplung (die wie in Abschnitt 2.1 erwähnt immer existiert) von μ und ν , so dass

$$\rho_K(\mu, \nu) = \sum_{x, y \in \Omega} \rho(x, y) \eta(x, y). \quad (22)$$

Dann können wir schreiben

$$\begin{aligned} \rho_K(\mu P, \nu P) &\leq \sum_{u, w \in \Omega} \rho(u, w) \theta(u, w) \\ &= \sum_{u, w \in \Omega} \rho(u, w) \sum_{x, y \in \Omega} \mathbb{P}(X_1 = u, Y_1 = w \mid X_0 = x, Y_0 = y) \eta(x, y) \\ &= \sum_{x, y \in \Omega} \eta(x, y) \underbrace{\sum_{u, w \in \Omega} \rho(u, w) \mathbb{P}(X_1 = u, Y_1 = w \mid X_0 = x, Y_0 = y)}_{\mathbb{E}_{x, y}(\rho(X_1, Y_1))} \end{aligned} \quad (23)$$

Hier können wir die Annahme $\mathbb{E}_{x, y}(\rho(X_1, Y_1)) \leq \rho(x, y) e^{-\alpha}$ nutzen, so dass

$$\rho_K(\mu P, \nu P) \leq e^{-\alpha} \sum_{x, y \in \Omega} \eta(x, y) \rho(x, y) = e^{-\alpha} \rho_K(\mu, \nu). \quad (24)$$

□

Korollar 2.4. *Mit den selben Annahmen wie zuvor können wir aus dem Theorem von Buble und Dyer außerdem schließen, dass*

$$d(t) \leq e^{-\alpha t} \text{diam}(\Omega) \quad (25)$$

und darauß folgt¹

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{\log(\text{diam}(\Omega)) + \log(1/\varepsilon)}{\alpha} \right\rceil. \quad (26)$$

Beweis. Wir iterieren das Ergebnis $\rho_K(\mu P, \nu P) \leq e^{-\alpha} \rho_K(\mu, \nu)$. Es folgt

$$\rho_K(\mu P^t, \nu P^t) \leq e^{-\alpha t} \rho_K(\mu, \nu) \leq e^{-\alpha t} \max_{x,y} \rho(x, y) \leq e^{-\alpha t} \text{diam}(\Omega) \quad (27)$$

Wir setzen $\mu = \delta_x$, $x \in \Omega$ beliebig und $\nu = \pi$. Wegen $\rho_{TV}(\mu, \nu) \leq \rho_K(\mu, \nu)$ zeigt es, dass

$$\begin{aligned} \|P^t(x, \cdot) - \pi\|_{TV} &= \rho(P^t(x, \cdot), \pi) \\ &\leq \rho_K(P^t(x, \cdot), \pi) \\ &\leq e^{-\alpha t} \rho_K(P(x, \cdot), \pi) \\ &\leq e^{-\alpha t} \text{diam}(\Omega) \end{aligned} \quad (28)$$

was die Behauptung beweist. \square

3 Obere Grenzen für Mischzeiten im Ising-Modell

3.1 Schnelles Mischen bei hoher Temperatur

Wir nutzen die Pfadkopplung um zu beweisen, dass für kleine β (d.h. bei hoher Temperatur), die Glauber-Dynamik für das Ising-Modell sehr schnell mischt.

Wir konstruieren die Glauber-Dynamik auf dem Ising-Modell auf einem Graph mit n Vertizes und maximalem Grad Δ , d.h. jeder Vertex hat höchstens Δ Nachbarn.

Proposition 3.1. *Definiere die Funktionen*

$$c(\beta) := 1 - \Delta \tanh(\beta) \quad c_g(\beta) := 1 - \frac{\Delta}{2} \tanh(2\beta) \quad (29)$$

¹ Zur Erinnerung:

$$\begin{aligned} d(t) &= \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{TV} \\ \text{diam}(\Omega) &= \max_{x \in \Omega} \rho(x, y) \\ t_{mix}(\varepsilon) &= \min\{t : d(t) \leq \varepsilon\} \end{aligned}$$

(i) Falls $c(\beta) > 0$, dann

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{n(\log n + \log(1/\varepsilon))}{c(\beta)} \right\rceil. \quad (30)$$

Insbesondere gilt diese Schranke für hohe Temperaturen $\beta < \Delta^{-1}$.

(ii) Angenommen jeder Vertex hat eine gerade Anzahl von Nachbarn. Falls dann $c_g(\beta) > 0$, dann

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{n(\log n + \log(1/\varepsilon))}{c_g(\beta)} \right\rceil \quad (31)$$

Beweis. Wir beweisen Aussage (i). Aussage (ii) lässt sich analog beweisen.

Wir definieren den Abstand ρ zwischen zwei Konfiguration σ, τ durch²

$$\rho(\sigma, \tau) := \frac{1}{2} \sum_{u \in V} |\sigma(u) - \tau(u)| \quad (32)$$

mit $\sigma(u), \tau(u) \in \{+1, -1\}$. Wenn $\sigma(u), \tau(u)$ benachbart sind, sie sich also nur in einem Vertex v unterscheiden, dann gilt $\rho(\sigma, \tau) = 1$. Wir nehmen o.B.d.A. an, dass $\sigma(v) = +1$, und $\tau(v) = -1$. Außerdem definieren wir die Menge $\mathcal{N}(v) := \{w \in V, w \sim v\}$ als die Vertexmenge, die alle Nachbarvertizes von v enthält. Wir betrachten nun eine Kopplung (X, Y) , die bei Konfigurationen σ, τ startet, d.h. $(X_0 = \sigma, Y_0 = \tau)$ und wollen dann untersuchen, wie die Zustände X_1 und Y_1 nach einem Schritt in der Glauberdynamik aussehen.

Wie üblich wählen wir einen Vertex $w \in V$ zufällig gleichverteilt, der nach Glauberdynamik aktualisiert werden soll (vgl. Abschnitt 1.3). Wir betrachten den Abstand zwischen beiden Konfigurationen nachdem das passiert ist. Die Wahrscheinlichkeit $\mathbb{P}(X_1(w) = +1)_{X_1, w}$, dass w nach einem Schritt den Wert $+1$ oder -1 annimmt ist nur durch die Anzahl von Nachbarvertizes von w mit Spins $+1$ und -1 bestimmt (Prop. ??). Wir müssen drei Fälle diskutieren, wie sich der Abstand $\rho(X_1, Y_1)$ ändern kann. In den ersten beiden sind die Nachbarvertizes von w in beiden Konfigurationen gleich und beide können gemeinsam aktualisiert werden. Im dritten Fall, wenn w ein Nachbar von v ist, dann muss man genauer unterscheiden, was passiert.

- Falls $w = v$, sind die Wert des Nachbarn von w in beiden Konfiguration gleich: $\mathbb{P}(X_1(w) = +1)_{X_1, w} = \mathbb{P}(Y_1(w) = +1)_{Y_1, w}$. Es folgt $\rho(X_1, Y_1) = 0$. Diese Situation hat die Wahrscheinlichkeit

$$\mathbb{P}(w = v) = \frac{1}{n} \quad (33)$$

- Falls $w \in V \setminus (\mathcal{N}(v) \cup \{v\})$, dann sind die Spinwerte von w und seinen Nachbarvertizes in beiden Konfiguration gleich. Es folgt $\rho(X_1, Y_1) = 1 =$

²Dies entspricht wieder der ursprünglichen Definition “Anzahl der unterschiedlichen Vertizes”.

$\rho(X_0, Y_0)$. Nach einem Schritt ist der Abstand zwischen Konfigurationen also immer noch 1. Die Wahrscheinlichkeit, dass dieser Fall auftritt ist

$$\mathbb{P}(w \notin \mathcal{N}(v) \cup v) = \frac{n - (|\mathcal{N}(v)| + 1)}{n} \leq 1 - \frac{2}{n} \quad (34)$$

- Falls $w \in \mathcal{N}(v)$, also w ein Nachbar von v ist, dann ist die Wahrscheinlichkeit $\mathbb{P}(X_1(w) = +1)_{X_1, w} \neq \mathbb{P}(Y_1(w) = +1)_{Y_1, w}$. Zunächst stellen wir fest, dass die Wahrscheinlichkeit für diesen Fall

$$\mathbb{P}(w \in \mathcal{N}(v)) = \frac{|\mathcal{N}(v)|}{n} \leq \frac{\Delta}{n} \quad (35)$$

ist. Es gibt eine Änderung zu $\rho(X_1, Y_1) = 2$, falls sich die beiden Konfigurationen unterschiedlich ändern. Die Wahrscheinlichkeit dafür zu berechnen ist etwas aufwändiger, da man nicht beide Konfigurationen gemeinsam aktualisieren kann.³ Wir zitieren hier nur das Ergebnis und verweisen für Details auf [5, Kap. 15.1].

Der Erwartungswert des Abstandes ist dann

$$\mathbb{E}_{X_0, Y_0}[\rho(X_1, Y_1)] \leq 1 - \frac{1}{n} + \frac{\Delta}{n} \tanh(\beta) \leq e^{-\frac{1-\Delta \tanh(\beta)}{n}} \cdot \rho(X_0, Y_0) \quad (36)$$

Nach Voraussetzung ist $\frac{1-\Delta \tanh(\beta)}{n} > 0$. Das heißt dieser Erwartungswert erfüllt gerade die Voraussetzung des Satzes von Bubley und Dyer (18) mit $\alpha = \frac{c(\beta)}{n}$. Wir können ihn und den Korollar 2.4 anwenden um

$$t_{mix}(\varepsilon) \leq \left\lceil \frac{n(\log n + \log(1/\varepsilon))}{c(\beta)} \right\rceil. \quad (37)$$

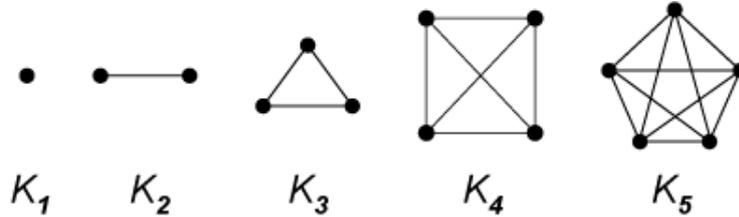
zu erhalten. Dabei haben wir genutzt $\text{diam}(\Omega) \leq n$. □

4 Die Mischzeit des Ising-Modells auf bestimmten Graphen

Nun haben wir genug theoretische Vorarbeit geleistet um für einige Spezialfälle konkrete Mischzeiten auszurechnen. Wir beschränken uns auf einfache Graphentypen, bei denen wir die Struktur ausnutzen können um kombinatorische Eigenschaften zu finden. Insbesondere ist es hilfreich die Anzahl der Nachbarn beliebiger Vertices zu kennen. Genauer gesagt betrachten wir *vollständige Graphen* bei denen jeder Vertex Nachbar jedes anderen Vertex ist, *Ringe*, in denen jeder Vertex genau zwei Nachbarn hat und regelmäßige *b*-äre Bäume in denen (mit Ausnahme der Blätter) jeder Vertex $b + 1$ Nachbarn hat.

³Man verwendet einen Trick um beide mit Hilfe nur einer Zufallsvariable zu aktualisieren. Dann beschränkt man die damit berechnete Wahrscheinlichkeit indem man eine gewisse Hilfsfunktion einführt und deren Supremum untersucht.

4.1 Vollständige Graphen



Sei $G = (V, E)$ ein vollständiger Graph, d.h. für jede zwei beliebige Vertices $v, w \in V$ existiert eine Kante $\{v, w\} \in E$.

Für einen Vollständigen Graph G mit n Vertices ist $\Delta = n - 1$. Es folgt dann für die Mischzeiten

(i) Falls $\beta < 1/n$, dann

$$t_{mix}(\varepsilon) \leq \frac{n(\log n + \log(1/\varepsilon))}{1 - n\beta} \quad (38)$$

(ii) Falls $\beta > 1/n$, dann gibt es ein positive Funktion $r(n\beta)$, so dass

$$t_{mix} \geq O(\exp[nr(n\beta)]) \quad (39)$$

Beweis von (i). Man hat

$$c(\beta) = 1 - (n - 1) \tanh(\beta) \geq 1 - (n - 1)\beta > 1 - n\beta > 0. \quad (40)$$

Anwenden von Korollar 2.4 und einsetzen liefert das Ergebnis. □

Beweis von (ii). Für die untere Schranke verwendet man die sogenannte *bottleneck ratio*⁴. Man nutzt sozusagen “Engstellen” im Zustandsraum um zu schätzen, wie lange der Mischvorgang mindestens dauern muss. □

Man kann hier sehr deutlich den sogenannten Phasenübergang beobachten. Die kritische Temperatur ist $\beta_c = 1/n$ und das Verhalten des Systems für $\beta > \beta_c$ und $\beta < \beta_c$ ist extrem unterschiedlich.

4.2 Ring-Graphen

Sei $G = (V, E)$ ein Ring-Graph, d.h. ein Graph in dem jeder Vertex genau 2 Nachbarn hat. Es ist also $\Delta = 2$ und insbesondere hat jeder Vertex eine gerade Anzahl von Nachbarn, so dass wir die Funktion $c_g(\beta)$ aus Fall (ii) der Proposition 3.1.

Für jedes $\beta > 0$ und ein festes $\varepsilon > 0$ hat die Mischzeit in der Glauber-Dynamik

⁴Siehe Vortrag und Seminararbeit *Untere Schranken für Mischzeiten* von Christopher Tili.

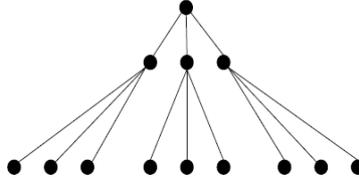
des Ising-Modells auf einem Ring mit n Vertices die Schranken

$$\frac{1 + o(1)}{2c_g(\beta)} n \log n \leq t_{mix}(\varepsilon) \leq \frac{1 + o(1)}{c_g(\beta)} n \log n \quad (41)$$

Beweis. Die obere Schranke folgt wieder direkt aus der Anwendung von Proposition 3.1 und Korollar 2.4. Für die untere Schranke braucht man *Wilson's Methode*. Hier verweisen wir wieder nur auf [5, Kap. 13.2 und 15.3]. \square

Wir sehen, dass die obere und untere Schranke für die Mischzeiten recht nah beieinander liegen. Es kann also keine große qualitative Änderung im Verhalten des Systems geben. Dies ist konsistent mit dem Ergebnis von Ising [4], dass das 1-dimensionale Gitter (d.h. eine Kette von Vertices, also ein Ring mit einer Kante entfernt) keinen Phasenübergang hat, denn dies ist dem Ring sehr ähnlich.

4.3 b -äre Baum-Graphen



Sei T ein b -ärer Baum (d.h. jeder Vertex hat b Kinder) mit Tiefe k . Man kann dann zeigen [5, Kap. 15.4], dass für jedes $\theta = \tanh(\beta) < 1/2\sqrt{b}$ (d.h. für hinreichend große Temperaturen) existiert eine Konstante $0 < c_\theta < 1$, sodass

$$\mathbb{E}_{\sigma, \tau}[\rho(X_1, Y_1)] \leq (1 - \frac{c_\theta}{n})\rho(\sigma, \tau) \quad (42)$$

und damit

$$t_{mix}(\varepsilon) \leq \frac{n}{c_\theta} \left[\frac{3}{2} \log n + \log(1/\varepsilon) \right]. \quad (43)$$

Man sieht also wieder, dass bei hohen Temperaturen ein System in der Regel schnell gemischt wird.

Literatur

- [1] Niels Bohr. On the Constitution of Atoms and Molecules, Part I-III. *Philosophical Magazine*, 26:1–25, 476–502, 857–875, 1913.
- [2] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in markov chains. In *FOCS '97: Proceedings of the 38th Annual Symposium on Foundations of Computer Science (FOCS)*, 1997.
- [3] Josiah Willard Gibbs. *Elementary Principles in Statistical Mechanics*. 1902.
- [4] Ernst Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.
- [5] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.
- [6] James Clerk Maxwell. Illustrations of the dynamical theory of gases. Part I+II. *Philosophical Magazine*, 19 und 20:19–32, 21–37, 1860.
- [7] Lars Onsager. Crystal statistics. I. a two-dimensional model with an order-disorder transition. *Physical Review*, 65:117–149, 1944.
- [8] Ernest Rutherford. The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine*, 21:669–688, 1911.