Mixing Times of the Restricted Rook's Walk and a Generalized Curie-Weiss Model

Benjamin Savoie, Ana Wright, and Renjun Zhu
University of Michigan-Flint, Willamette University, and University of California, Berkeley

Faculty Advisor: Peter T. Otto, Willamette University

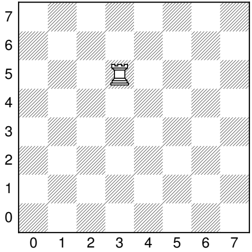
July 29, 2016

As part of the Willamette Valley REU

Outline

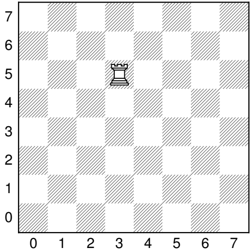
- Background
  - Markov Chains
  - Mixing Time
  - Coupling
  - Path Coupling
- Restricted Rook's Walk
  - Far / Near Restriction
  - Couplings
  - Mixing Time Bounds / Long Term Behavior
- Generalized Curie Weiss Model
  - Equilibrium Phase Strucure
  - Dynamic Phase Structure
  - β_M vs. β_C

Motivation



If the rook has made one move, what do we know about its starting position?
If the rook has made two moves, what do we know about its starting position?

Markov Chains



Markov chain: a sequence of random variables/vectors $X_1, X_2, \ldots$ such that

$$P[X_{t+1} = x_{t+1} | X_t = x_t, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0] = P[X_{t+1} = x_{t+1} | X_t = x_t]$$

Where $x_0, \ldots, x_t$ are states at time t, and Ω is the state space.

Markov chains

- Transition probability:

$$p(x, y) = \mathbb{P}(X_{t+1} = y | X_t = x)$$

- Transition matrix:

$$P = [p(x, y)]$$

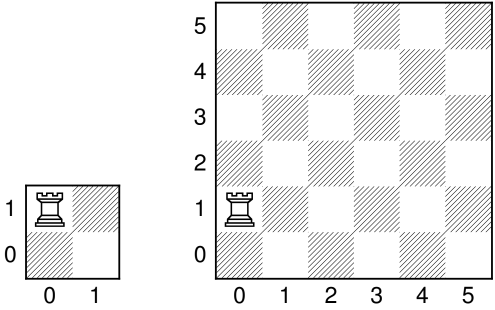
- Distribution at time t: $P^t(x, \cdot) = \mathbb{P}(X_t = \cdot | X_0 = x)$
  This is the x-th row of the transition matrix.

Convergence Theorem

For irreducible and aperiodic Markov chains

$$P^t(x, \cdot) \Longrightarrow \pi \qquad \text{as} \quad t \to \infty$$

where π is the unique stationary distribution of the chain, i.e. $\pi P = \pi$.



Mixing time

- Total variation distance

$$\|\mu - \nu\|_{\mathrm{TV}} := \max_{A \subset \Omega} |\mu(A) - \nu(A)| = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)|$$

- Distance from stationarity

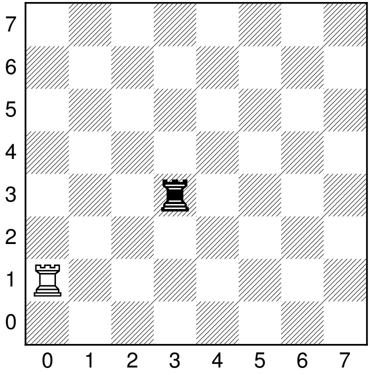
$$d(t) := \max_{x \in \Omega} \|P^t(x, \cdot) - \pi\|_{\mathrm{TV}}$$

- Mixing time: a measure of the convergence rate of the chain to its stationary distribution.

$$t_{\mathrm{mix}}(\varepsilon) := \min\{t : d(t) \leq \varepsilon\}$$

Bounding the Mixing Time

A **coupling** of two distributions μ and ν is a pair of random variables (X, Y) on a common probability space with marginals μ and ν.



Coupling Markov Chains

Coupling Inequality [LPW]:

$$\|\mu - \nu\|_{\text{TV}} = \|P^t(x,\cdot) - P^t(y,\cdot)\|_{\text{TV}} \le \mathbb{P}(X_t \neq Y_t)$$

Proof: For any $A \subset \Omega$,

$$\mu(A) - \nu(A) = \mathbb{P}(X_t \in A) - \mathbb{P}(Y_t \in A) \le \mathbb{P}(X_t \in A, Y_t \notin A) \le \mathbb{P}(X_t \neq Y_t)$$

So,

$$\begin{aligned} d(t) = \max_{x \in \Omega} \|P^t(x,\cdot) - \pi\|_{TV} &\le \max_{x,y} \|P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \\ &\le \max_{x,y} \mathbb{P}(X_t \neq Y_t) \end{aligned}$$

Coupling Markov Chains

Coupling Inequality [LPW]:

$$\|\mu - \nu\|_{\mathrm{TV}} = \|P^t(x,\cdot) - P^t(y,\cdot)\|_{\mathrm{TV}} \le \mathbb{P}(X_t \neq Y_t)$$

Proof: For any $A \subset \Omega$,

$$\mu(A) - \nu(A) = \mathbb{P}(X_t \in A) - \mathbb{P}(Y_t \in A) \le \mathbb{P}(X_t \in A, Y_t \notin A) \le \mathbb{P}(X_t \neq Y_t)$$

So,

$$d(t) = \max_{x \in \Omega} \|P^t(x,\cdot) - \pi\|_{TV} \le \max_{x,y} \|P^t(x,\cdot) - P^t(y,\cdot)\|_{TV}$$
$$\le \max_{x,y} \mathbb{P}(X_t \neq Y_t)$$

Coupling Markov Chains

Coupling Inequality [LPW]:

$$\|\mu - \nu\|_{\text{TV}} = \|P^t(x,\cdot) - P^t(y,\cdot)\|_{\text{TV}} \leq \mathbb{P}(X_t \neq Y_t)$$

Proof: For any $A \subset \Omega$,

$$\mu(A) - \nu(A) = \mathbb{P}(X_t \in A) - \mathbb{P}(Y_t \in A) \leq \mathbb{P}(X_t \in A, Y_t \notin A) \leq \mathbb{P}(X_t \neq Y_t)$$

So,

$$\begin{aligned} d(t) = \max_{x \in \Omega} \|P^t(x,\cdot) - \pi\|_{TV} &\leq \max_{x,y} \|P^t(x,\cdot) - P^t(y,\cdot)\|_{TV} \\ &\leq \max_{x,y} \mathbb{P}(X_t \neq Y_t) \end{aligned}$$

Coupling of Markov Chains

With a metric ρ on our state space in conjunction with Markov's inequality and results about $d(t)$ give us:

$$d(t) \leq \max_{x,y} \mathbb{P}(X_t \neq Y_t) = \max_{x,y} \mathbb{P}[\rho(X_t, Y_t) \geq 1] \leq \frac{\max_{x,y} E[\rho(X_t, Y_t)]}{1}$$

Mean coupling distance: $\max_{x,y} E[\rho(X_t, Y_t)]$

Coupling of Markov Chains

With a metric ρ on our state space in conjunction with Markov's inequality and results about $d(t)$ give us:

$$d(t) \leq \max_{x,y} \mathbb{P}(X_t \neq Y_t) = \max_{x,y} \mathbb{P}[\rho(X_t, Y_t) \geq 1] \leq \frac{\max_{x,y} E[\rho(X_t, Y_t)]}{1}$$

Mean coupling distance: $\max_{x,y} E[\rho(X_t, Y_t)]$

Coupling of Markov Chains

With a metric ρ on our state space in conjunction with Markov's inequality and results about $d(t)$ give us:

$$d(t) \leq \max_{x,y} \mathbb{P}(X_t \neq Y_t) = \max_{x,y} \mathbb{P}[\rho(X_t, Y_t) \geq 1] \leq \frac{\max_{x,y} E[\rho(X_t, Y_t)]}{1}$$

Mean coupling distance: $\max_{x,y} E[\rho(X_t, Y_t)]$

Coupling of Markov Chains

Want to show contraction after one step:

$$E[\rho(X_t, Y_t)|x_{t-1}, y_{t-1}] = (1-\alpha)\rho(x_{t-1}, y_{t-1}) \leq e^{-\alpha}\rho(x_{t-1}, y_{t-1})$$

with $0 < \alpha < 1$

Iteration gives:

$$E[\rho(X_t, Y_t)] \leq e^{-\alpha t} E[\rho(X_0, Y_0)]$$

Coupling of Markov Chains

Since

$$d(t) \leq \max_{x,y} E[\rho(X_t, Y_t)]$$

and:

$$d(t_{mix}) \leq \varepsilon$$

then:

$$\max_{x,y} E[\rho(X_0, Y_0)]e^{-\alpha t} \leq \varepsilon$$

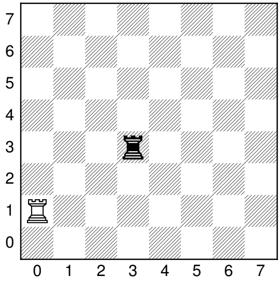
thus we have the Mixing Time Theorem:

$$t_{mix}(\varepsilon) \leq \frac{1}{\alpha} \log \left(\frac{\max_{x,y} E[\rho(X_0, Y_0)]}{\varepsilon} \right)$$

Path Coupling

If neighboring pairs contract, then a pair of chains will contract from any two states distance r apart.

$$X_t = x_0, x_1, ..., x_r = Y_t$$



$$E[\rho(X_t, Y_t)] \leq \sum_{i=1}^{r} E[\rho(X_{t,i}, X_{t,i-1})]$$

Coupling of Markov Chains

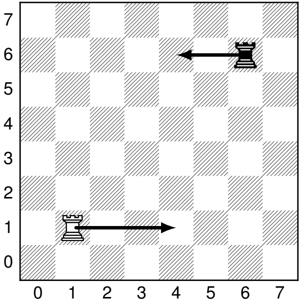
Overall goal: Find a ’good’ coupling; i.e a coupling that contracts with each time step. $(\alpha > 0)$

$$E[\rho(X_t, Y_t)|x_{t-1}, y_{t-1}] \leq e^{-\alpha}\rho(x_{t-1}, y_{t-1}) = e^{-\alpha}$$

"Good" Coupling

1. A 'good' coupling rule encourages two rooks to meet as fast as possible.
2. One approach we made was to require that after one step, $\rho(X_t, Y_t) = 0$ or 1, so that we could **guarantee** the rooks would contract.
3. It is possible to have bigger $\rho(X_t, Y_t)$ for coupling, and on average they might sill contract.

Literature Review: Unrestricted

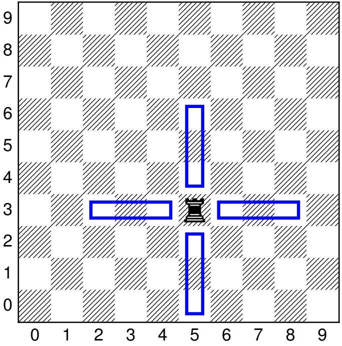


Theorem [MORS]

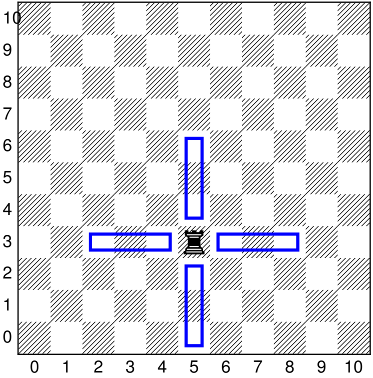
Mixing time bound: $t_{mix}(\varepsilon) \leq \left\lceil \frac{log(\frac{d}{\varepsilon})}{log \frac{d(n-1)}{(d-1)(n-1)+1}} \right\rceil \leq \left\lceil \frac{d(n-1)}{n-2} log(\frac{d}{\varepsilon}) \right\rceil$

Definition: Far Restricted Rook's Walk

Legal Moves: $K = \{1, 2, 3 = \lfloor \frac{n}{2} \rfloor - r\}$



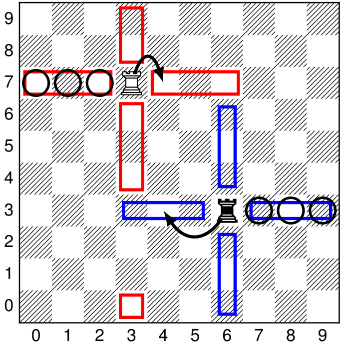
$n = 10$, $r = 2$,



$n = 11$, $r = 2$

Method: Far Restriction

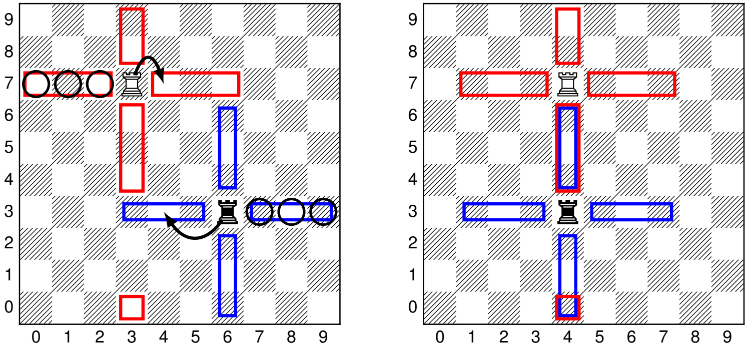
Coupling Rule 1: Move to Common Accessible set, so they will match in 1 dimension.



$n = 10, r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

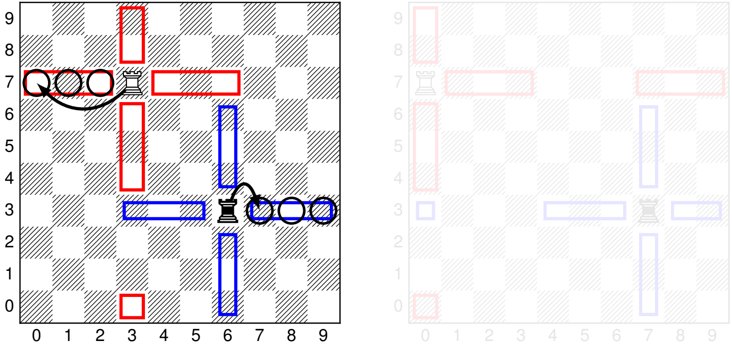
Coupling Rule 1: Move to Common Accessible set, so they will match in 1 dimension.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

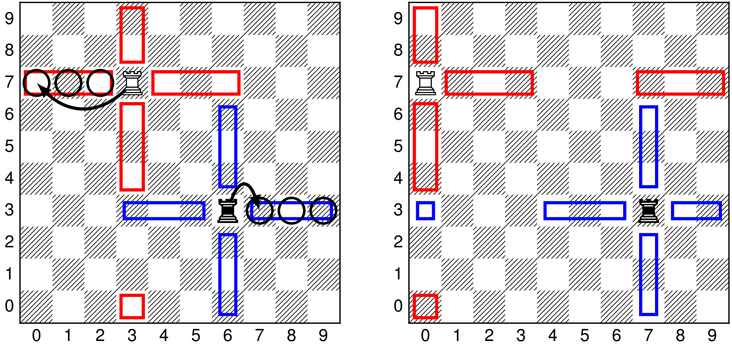
Coupling Rule 2: Move to circle set (accessible for one, not the other). ”First Match First”, etc.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

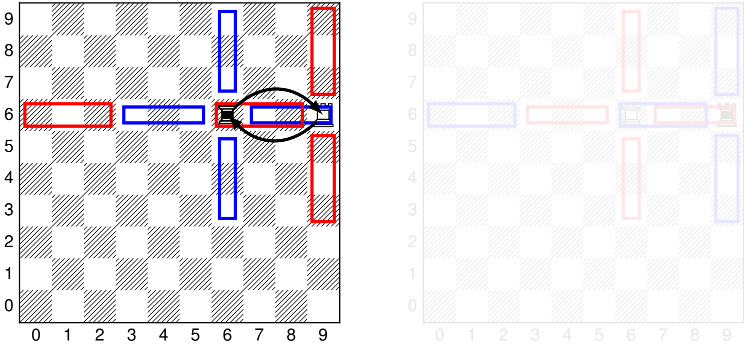
Coupling Rule 2: Move to circle set (accessible for one, not the other). "First Match First", etc.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

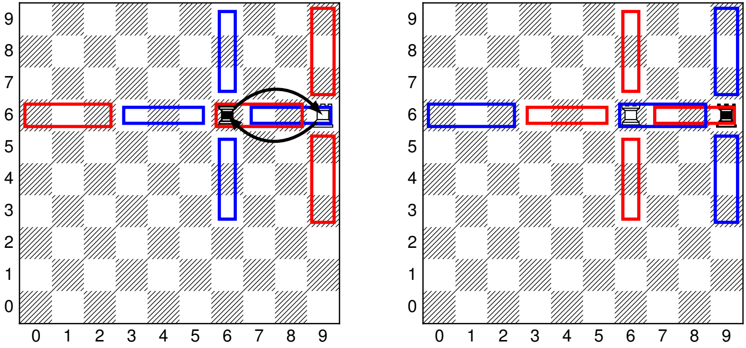
Coupling Rule 3: Swap.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

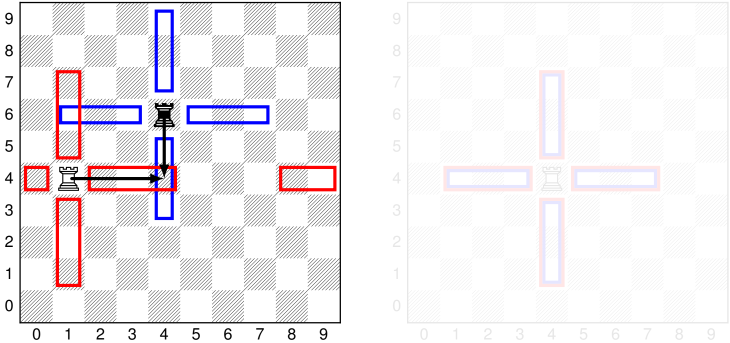
Coupling Rule 3: Swap.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

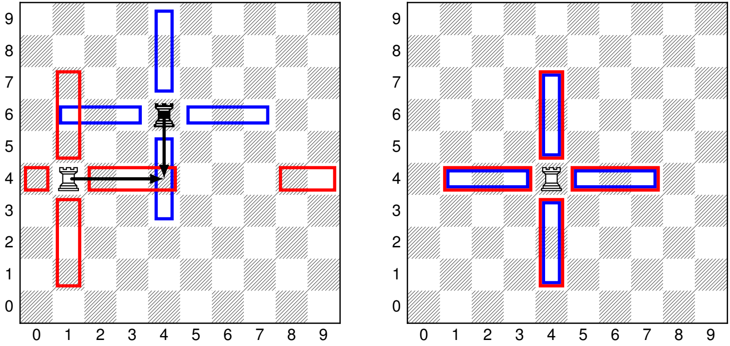
Coupling Rule 4: Move to Common Accessible square, so they will match.



$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Method: Far Restriction

Coupling Rule 4: Move to Common Accessible square, so they will match.

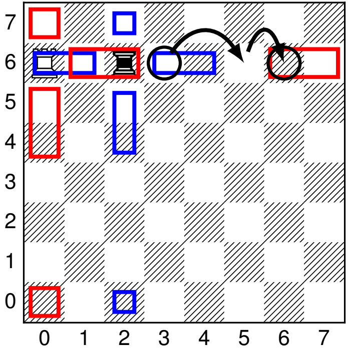


$n = 10$, $r = 2$, so $K = \{1, 2, 3\}$

Far Restriction: NOT Contract

What if $n = 8$ and $r = 2$???
$\Longrightarrow \rho(X_t, Y_t)$ increases

$$E[\rho(X_t, Y_t) | \rho(x_{t-1}, y_{t-1}) = 1] = \frac{2}{8} \cdot 2 + \frac{4+1}{8} \cdot 1 + \frac{1}{8} \cdot 0 = \frac{9}{8} > 1$$



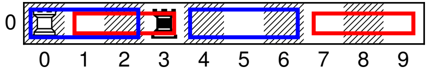
Here, we have $\rho(X_t, Y_t) = 2$ in 2 cases with $K = \{1, 2\}$.

Far Restriction: Condition

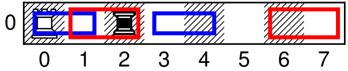
Claim: For even n, if $n \geq 6r - 2$, then $\rho(X_t, Y_t) = 1, 0$ for all neighboring pairs.

Idea: Prevent increase in $\rho(X_t, Y_t)$. The number of inaccessible squares is $n - 2k - 1 = 2r - 1$. This needs to be reached by other rook. So:

$$n - 2k - 1 \leq \frac{n}{2} - r$$
$$n \geq 6r - 2$$



For $r = 2$ Example, $8 = n < 6 \cdot 2 - 2 = 10$, condition **fails**!!



Result: Far Restriction Condition

Theorem [OSWZ'16]

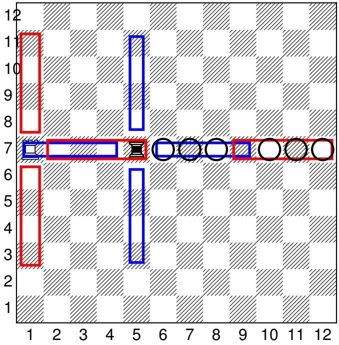
For any far restricted rook's walk on an n^d board with legal moves $K = \{1, 2, ..., \lfloor \frac{n}{2} \rfloor - r\}$, if $n \geq 6r - 2$ for even n, and $n \geq 6r + 1$ for odd n, then

1. $diam(\Omega) = 2d$;
2. $\rho(X_t, Y_t) = 1, 0$ for all neighboring pairs $\rho(X_{t-1}, Y_{t-1})$.

Analysis: Mean Coupling Distance

Condition: $n \geq 6r - 2$.

$$E_{even}[\rho(X_{t+1}, Y_{t+1}) \mid \rho(x_t, y_t) = 1] = (\frac{d-1}{d}) \cdot 1 + \frac{(2r-1)+1}{2d(\frac{n}{2}-r)} \cdot 1 + \left[1 - \frac{d-1}{d} - \frac{(2r-1)+1}{2d(\frac{n}{2}-r)}\right] \cdot 0 = 1 - \left[\frac{1}{d} - \frac{r}{d(\frac{n}{2}-r)}\right] = 1 - \left[\frac{\frac{n}{2}-2r}{d(\frac{n}{2}-r)}\right] \leq e^{-\alpha}$$



Result: Mixing Bound (Far Restriction)

Contraction bound: $\begin{cases} \mathrm{E}[\rho(X_t+1, Y_t+1) | \rho(x_t, y_t) = 1] \leq e^{-\alpha}. \\ \mathrm{diam}(\Omega) = 2d \end{cases}$

Mixing bound for Restricted Rook's Walk:

Theorem [OSWZ'16]

Mixing Time Bound: $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r)}{n-4r} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$, for even n.

Result: Mixing Bound (Far Restriction)

Contraction bound: $\begin{cases} \mathrm{E}[\rho(X_t+1, Y_t+1)|\rho(x_t, y_t)=1] \leq e^{-\alpha}. \\ \mathrm{diam}(\Omega) = 2d \end{cases}$

Mixing bound for Restricted Rook's Walk:

Theorem [OSWZ'16]

Mixing Time Bound: $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r)}{n-4r} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$, for even n.

Result: Odd n (Far Restriction)

Condition: $n \geq 6r + 1$

$$E_{odd}[\rho(X_{t+1}, Y_{t+1}) \mid \rho(x_t, y_t) = 1] =$$
$$(\frac{d-1}{d}) \cdot 1 + \frac{(2r)+1}{2d(\frac{n-1}{2}-r)} \cdot 1 + (1 - \frac{d-1}{d} - \frac{2r+1}{2d(\frac{n-1}{2})-r}) \cdot 0 =$$
$$1 - \left[\frac{1}{d} - \frac{2r+1}{d(n-2r-1)}\right] = 1 - \left[\frac{n-4r-2}{d(n-2r-1)}\right] \leq e^{-\alpha}$$

Theorem [OSWZ'16]

Mixing Time Bound: $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r-1)}{n-4r-2} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$, for odd n.

Far Restriction: Mixing Time Bound Behavior

$$\lim_{n\to\infty} E_{even/odd}[\rho(X_t, Y_t)|\rho(x_{t-1}, y_{t-1}) = 1] = \frac{d-1}{d} = 1 - \frac{1}{d}$$

Corollary

For both even and odd n,

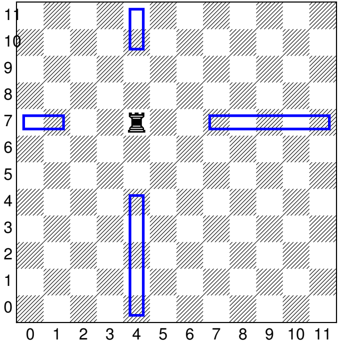
$$\lim_{n\to\infty} t_{mix}(\varepsilon) \leq \left\lceil d \log\left(\frac{2d}{\varepsilon}\right)\right\rceil$$

Corollary

Mixing Time Bound $t_{mix}(\varepsilon)$ strictly increases with r.

Definition: Near Restriction

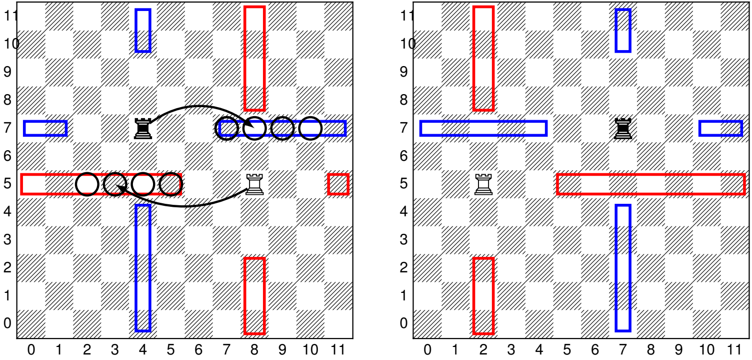
For the **near restriction**, $K = \{r+1, r+2, ..., \lfloor \frac{n}{2} \rfloor\}$, where r is the restriction.



Example: $n = 12$, $r = 2$, so $K = \{3, 4, 5, 6\}$

Method: Near Restriction

Coupling Rule: "First to First, No Swap!!!" Circle set maintains the same distance apart, $\rho(X_t, Y_t) = 1$.

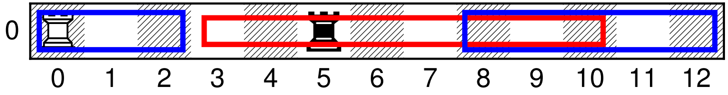


$$n = 12, r = 2, \text{ so } K = \{3, 4, 5, 6\},$$

Result: $\rho(X_t, Y_t)$ Theorem (Near Restriction)

Theorem [OSWZ'16]

For any neighboring pair of near restricted rook's walk on an n^d board with legal moves $K = \{r+1, r+2, ..., \lfloor \frac{n}{2} \rfloor\}, \rho(X_t, Y_t) = 0, 1.$



$n = 13$, $r = 2$, $K = \{3, 4, 5, 6\}$
Common Accessible set: $\rho(X_t, Y_t) = 0$
Rest: $\rho(X_t, Y_t) = 1$, i.e. $3 \to 11, 4 \to 12$, etc. 5 apart.

Result: Mixing Time Bound (Near Restriction)

Condition $n \geq 4r + 4$ for even n, and $n \geq 4r + 3$ for odd n;

then:

$$E_{even/odd}[\rho(X_t, Y_t) \mid \rho(x_{t-1}, y_{t-1}) = 1] =$$

$$(\frac{d-1}{d}) \cdot 1 + (\frac{2r+1}{d(n-2r-1)}) \cdot 1 + (1 - \frac{d-1}{d} - \frac{2r+1}{d(n-2r-1)}) \cdot 0 =$$

$$1 - \left[\frac{1}{d} - \frac{2r+1}{d(n-2r-1)}\right] = 1 - \frac{n-4r-2}{d(n-2r-1)} \leq e^{-\alpha}$$

Theorem [OSWZ'16]

Mixing Time Bound: $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r-1)}{n-4r-2} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$

Near Restriction: Mixing Time Bound Behavior

$$\lim_{n\to\infty} E_{even/odd}[\rho(X_t, Y_t)|\rho(x_{t-1}, y_{t-1} = 1] = \frac{d-1}{d} = 1 - \frac{1}{d}$$

Corollary

$$\lim_{n\to\infty} t_{mix}(\varepsilon) \leq \left\lceil d \log\left(\frac{2d}{\varepsilon}\right)\right\rceil$$

Corollary

$t_{mix}(\varepsilon)$ *strictly increases with* r

Different Restriction

Corollary

For both the far and near restricted rook's walk, if $r = \frac{n}{f}$*, then* $\lim_{n\to\infty} t_{mix}(\varepsilon) \leq \frac{d(f-2)}{f-4} \log\left(\frac{2d}{\varepsilon}\right)$

Corollary

For a near or far restricted rook's walk, if $r = o(n)$*, i.e.* $r = n^{\frac{1}{p}}$*, then* $\lim_{n\to\infty} t_{mix}(\varepsilon) \leq \left\lceil d \cdot log\left(\frac{2d}{\varepsilon}\right)\right\rceil$

Summary: Unrestricted Rook's Walk

1. **No Restriction**: $r = 0$
   $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-1)}{n-2} log(\frac{2d}{\varepsilon}) \right\rceil$
2. **Far Restriction**: $r < \frac{n}{6}$
   $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r)}{n-4r} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$, for even n.
   $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r-1)}{n-4r-2} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$, for odd n.
3. **Near Restriction**: $r < \frac{n}{4}$
   $t_{mix}(\varepsilon) \leq \left\lceil \frac{d(n-2r-1)}{n-4r-2} \log\left(\frac{2d}{\varepsilon}\right) \right\rceil$

Remark: As $n \to \infty$, all mixting time bound converge to the bound of $\left\lceil d \cdot log(\frac{2d}{\varepsilon}) \right\rceil$.

Statistical Mechanics

- ”In statistical mechanics, one derives macroscopic properties of a substance from a probability distribution that describes the complicated interactions among the individual constituent particles.” [1]

Curie Weiss Model

- In the *Curie Weiss model*, there are n particles, each with spin $+1$ or -1. A state is a complete configuration $\omega \in \{-1, 1\}^n$ that describes the spin at each particle. $S_n(\omega)$ is the sum of all the spins of ω.
- Microscopic quantity: Spin at each particle.
- Macroscopic quantity: The mean spin / magnetization $\frac{S_n(\omega)}{n}$.

Curie Weiss Model

- In the *Curie Weiss model*, there are n particles, each with spin $+1$ or -1. A state is a complete configuration $\omega \in \{-1, 1\}^n$ that describes the spin at each particle. $S_n(\omega)$ is the sum of all the spins of ω.

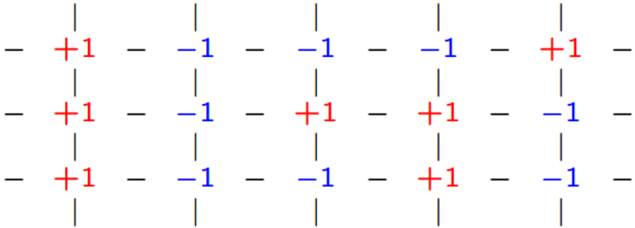
- Microscopic quantity: Spin at each particle.

- Macroscopic quantity: The mean spin / magnetization $\frac{S_n(\omega)}{n}$.

Curie Weiss Model

- In the *Curie Weiss model*, there are n particles, each with spin $+1$ or -1. A state is a complete configuration $\omega \in \{-1, 1\}^n$ that describes the spin at each particle. $S_n(\omega)$ is the sum of all the spins of ω.
- Microscopic quantity: Spin at each particle.
- Macroscopic quantity: The mean spin / magnetization $\frac{S_n(\omega)}{n}$.

Curie Weiss Model



A particular configuration ω where $\frac{S_n(\omega)}{n} = -\frac{1}{15}$.

Gibbs Ensemble

- The stationary distribution is given by the *Gibbs Ensemble*:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$$

- The partition function $Z_n(\beta) = \sum_{\omega \in \Omega} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$ normalizes the probabilities and $\beta = \frac{1}{T}$.

- $g(\frac{S_n(\omega)}{n})$ defines the interaction between particles.

Gibbs Ensemble

- The stationary distribution is given by the *Gibbs Ensemble*:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$$

- The partition function $Z_n(\beta) = \sum_{\omega \in \Omega} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$ normalizes the probabilities and $\beta = \frac{1}{T}$.

- $g(\frac{S_n(\omega)}{n})$ defines the interaction between particles.

Gibbs Ensemble

- The stationary distribution is given by the *Gibbs Ensemble*:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$$

- The partition function $Z_n(\beta) = \sum_{\omega \in \Omega} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$ normalizes the probabilities and $\beta = \frac{1}{T}$.

- $g(\frac{S_n(\omega)}{n})$ defines the interaction between particles.

Gibbs Ensemble

- The stationary distribution is given by the *Gibbs Ensemble*:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$$

- The partition function $Z_n(\beta) = \sum_{\omega\in\Omega} e^{n\beta g(\frac{S_n(\omega)}{n})} \prod_{i=1}^{n} \rho(\omega_i)$ normalizes the probabilities and $\beta = \frac{1}{T}$.

- $g(\frac{S_n(\omega)}{n})$ defines the interaction between particles.

Our Generalized Curie-Weiss Model

- In the classical CW model, $g(x) = x^2$.
- We instead have:

$$g(x) = \frac{\alpha_1}{4!}x^4 + \frac{\alpha_2}{2!}x^2$$

- $\alpha_1 > 0$ and $\alpha_2 > 0$ represent interaction strengths.
- The Gibbs Ensemble then becomes:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta \frac{\alpha_1}{4!}\left(\frac{S_n(\omega)}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{S_n(\omega)}{n}\right)^2} \prod_{i=1}^{n} \rho(\omega_i)$$

Our Generalized Curie-Weiss Model

- In the classical CW model, $g(x) = x^2$.
- We instead have:

$$g(x) = \frac{\alpha_1}{4!}x^4 + \frac{\alpha_2}{2!}x^2$$

- $\alpha_1 > 0$ and $\alpha_2 > 0$ represent interaction strengths.
- The Gibbs Ensemble then becomes:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta \frac{\alpha_1}{4!}\left(\frac{S_n(\omega)}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{S_n(\omega)}{n}\right)^2} \prod_{i=1}^{n} \rho(\omega_i)$$

Our Generalized Curie-Weiss Model

- In the classical CW model, $g(x) = x^2$.
- We instead have:

$$g(x) = \frac{\alpha_1}{4!}x^4 + \frac{\alpha_2}{2!}x^2$$

- $\alpha_1 > 0$ and $\alpha_2 > 0$ represent interaction strengths.
- The Gibbs Ensemble then becomes:

$$P_{n,\beta}(\omega) = \frac{1}{Z_n(\beta)} e^{n\beta \frac{\alpha_1}{4!}\left(\frac{S_n(\omega)}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{S_n(\omega)}{n}\right)^2} \prod_{i=1}^{n} \rho(\omega_i)$$

Phase Transition Structure

- In the Curie Weiss model, we see how changing the temperature $(1/\beta)$ affects the equilibrium structure and dynamic structure.
- **Equilibrium Structure:** How many global minimizers of the free energy function $G_\beta(z)$?
- **Dynamic Structure:** Does the Glauber dynamics Markov chain mix rapidly or slowly? $f_\beta(z)$

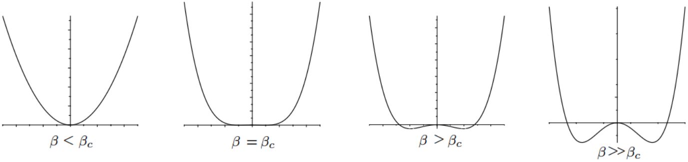
Phase Transition Structure

- In the Curie Weiss model, we see how changing the temperature ($1/\beta$) affects the equilibrium structure and dynamic structure.
- **Equilibrium Structure:** How many global minimizers of the free energy function $G_\beta(z)$?
- **Dynamic Structure:** Does the Glauber dynamics Markov chain mix rapidly or slowly? $f_\beta(z)$

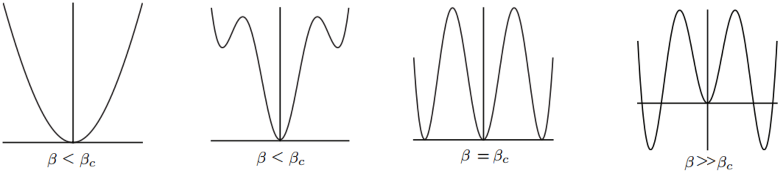
Phase Transition Structure

- In the Curie Weiss model, we see how changing the temperature $(1/\beta)$ affects the equilibrium structure and dynamic structure.
- **Equilibrium Structure:** How many global minimizers of the free energy function $G_\beta(z)$?
- **Dynamic Structure:** Does the Glauber dynamics Markov chain mix rapidly or slowly? $f_\beta(z)$

Equilibrium Phase Transition Structure



$G_\beta(S_n/n)$ for second order continuous phase transition.

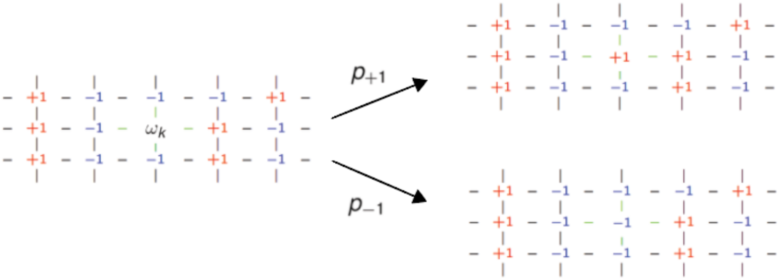


$G_\beta(S_n/n)$ for first order discontinuous phase transition.

Glauber Dynamics

Glauber dynamics defines a Markov chain that is guaranteed to converge to a given stationary distribution (Gibbs Ensemble).

Glauber Dynamics



Glauber dynamics have **local update probabilities.**

$$p_{\pm 1}(\omega, k) = \frac{e^{n\beta\left(\frac{\alpha_1}{4!}\left(\frac{\widetilde{S}(\omega,k)\pm 1}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{\widetilde{S}(\omega,k)\pm 1}{n}\right)^2\right)}}{e^{n\beta\left(\frac{\alpha_1}{4!}\left(\frac{\widetilde{S}(\omega,k)+1}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{\widetilde{S}(\omega,k)+1}{n}\right)^2\right)} + e^{n\beta\left(\frac{\alpha_1}{4!}\left(\frac{\widetilde{S}(\omega,k)-1}{n}\right)^4 + \frac{\alpha_2}{2!}\left(\frac{\widetilde{S}(\omega,k)-1}{n}\right)^2\right)}}$$

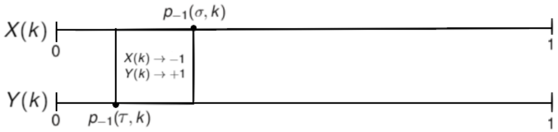
Greedy Coupling

- Let $X_t = \sigma$ and $Y_t = \tau$ where σ and τ are any two different configurations on $\{-1, 1\}^n$. With a common source of randomness, $U \in (0, 1)$, define the **greedy coupling**:

$$X(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\sigma, k) \\ +1 & \text{if } p_{-1}(\sigma, k) \leq U \leq 1 \end{cases}$$

$$Y(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\tau, k) \\ +1 & \text{if } p_{-1}(\tau, k) \leq U \leq 1 \end{cases}$$

- Define the metric $\rho(s, t) = \frac{1}{2}\Sigma_{i=1}^{n}|s_i - t_i|$.



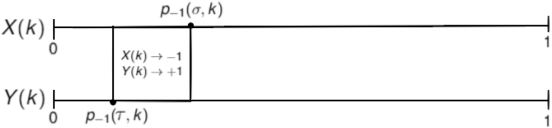
Greedy Coupling

- Let $X_t = \sigma$ and $Y_t = \tau$ where σ and τ are any two different configurations on $\{-1, 1\}^n$. With a common source of randomness, $U \in (0, 1)$, define the **greedy coupling**:

$$X(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\sigma, k) \\ +1 & \text{if } p_{-1}(\sigma, k) \leq U \leq 1 \end{cases}$$

$$Y(k) = \begin{cases} -1 & \text{if } 0 \leq U \leq p_{-1}(\tau, k) \\ +1 & \text{if } p_{-1}(\tau, k) \leq U \leq 1 \end{cases}$$

- Define the metric $\rho(s, t) = \frac{1}{2}\Sigma_{i=1}^{n}|s_i - t_i|$.



Aggregate Path Coupling

- With standard path coupling, contraction is required between all neighboring configurations.
- With aggregate path coupling, contraction is only required between a configuration close to the equilibrium and any other configuration.

Aggregate Path Coupling

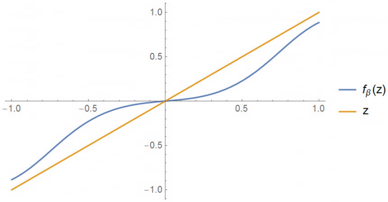
- With standard path coupling, contraction is required between all neighboring configurations.
- With aggregate path coupling, contraction is only required between a configuration close to the equilibrium and any other configuration.

Mean Coupling Distance

Let σ be a configuration close to the origin and τ be any other configuration that is neighboring σ. Then the mean coupling distance is:

$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \le 1 - \left(\frac{1}{n} - \frac{1}{2}\right)\left[f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)\right] + \mathcal{O}\left(\frac{1}{n^2}\right)$$

$$\frac{f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)}{\left(\frac{S_n(\sigma)}{n}\right) - \left(\frac{S_n(\tau)}{n}\right)} \le \alpha'$$



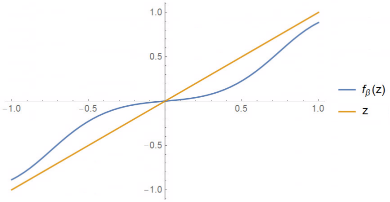
$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \le 1 + \frac{\alpha' - 1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

Mean Coupling Distance

Let σ be a configuration close to the origin and τ be any other configuration that is neighboring σ. Then the mean coupling distance is:

$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \leq 1 - \left(\frac{1}{n} - \frac{1}{2}\right)\left[f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)\right] + \mathcal{O}\left(\frac{1}{n^2}\right)$$

$$\frac{f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)}{\left(\frac{S_n(\sigma)}{n}\right) - \left(\frac{S_n(\tau)}{n}\right)} \leq \alpha'$$



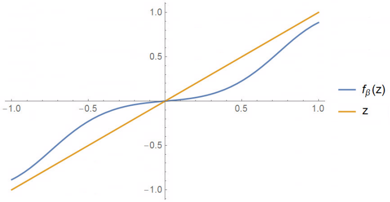
$$\mathbb{E}_{\sigma,\tau}[\rho(X,Y)] \leq 1 + \frac{\alpha' - 1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

Mean Coupling Distance

Let σ be a configuration close to the origin and τ be any other configuration that is neighboring σ. Then the mean coupling distance is:

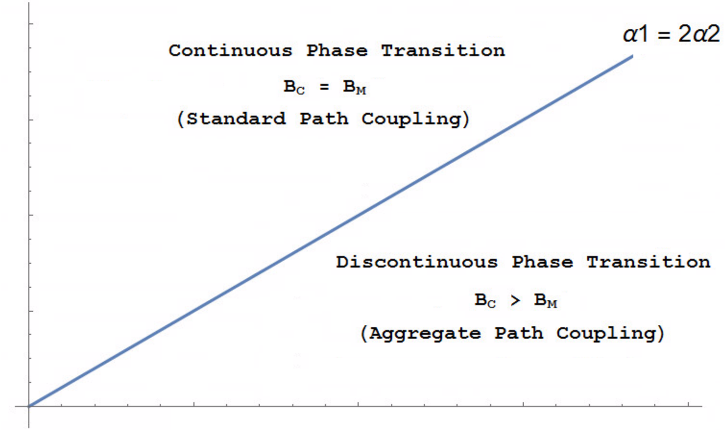
$$\mathbb{E}_{\sigma,\tau}[\rho(X, Y)] \leq 1 - \left(\frac{1}{n} - \frac{1}{2}\right)\left[f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)\right] + \mathcal{O}\left(\frac{1}{n^2}\right)$$

$$\frac{f_\beta\left(\frac{S_n(\sigma)}{n}\right) - f_\beta\left(\frac{S_n(\tau)}{n}\right)}{\left(\frac{S_n(\sigma)}{n}\right) - \left(\frac{S_n(\tau)}{n}\right)} \leq \alpha'$$

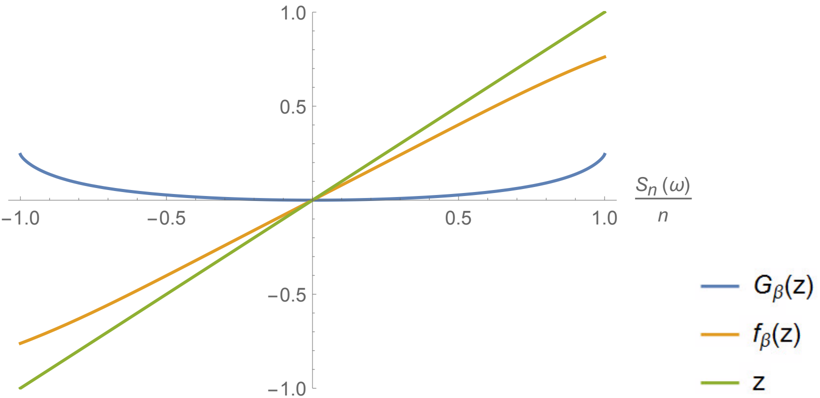


$$\mathbb{E}_{\sigma,\tau}[\rho(X, Y)] \leq 1 + \frac{\alpha' - 1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right)$$

Phase Transition Structure

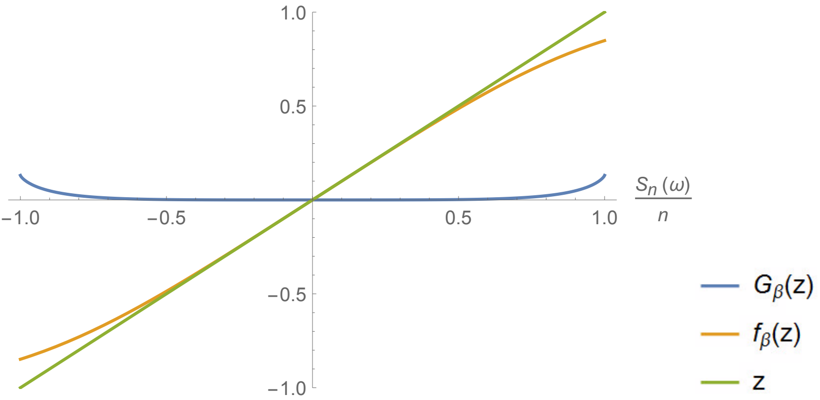


Continuous Phase Transition ($\alpha_1 < 2\alpha_2$)



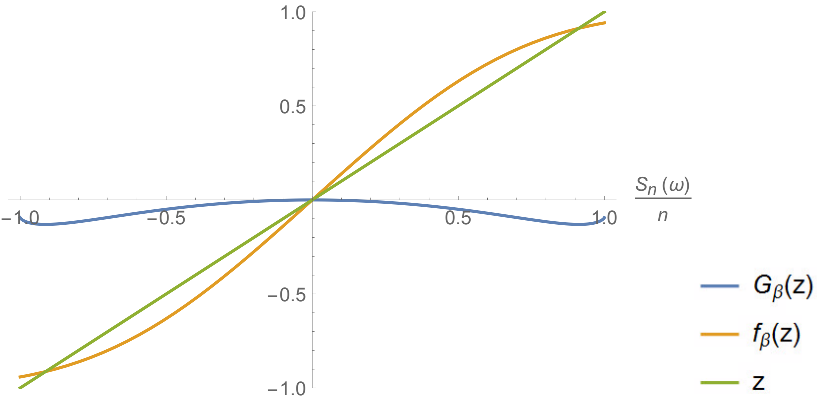
$0 < \beta < \beta_C$

Continuous Phase Transition ($\alpha_1 < 2\alpha_2$)



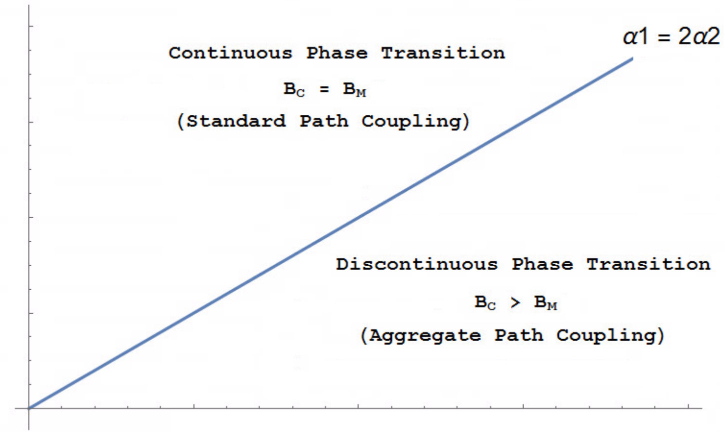
$\beta = \beta_C$

Continuous Phase Transition ($\alpha_1 < 2\alpha_2$)

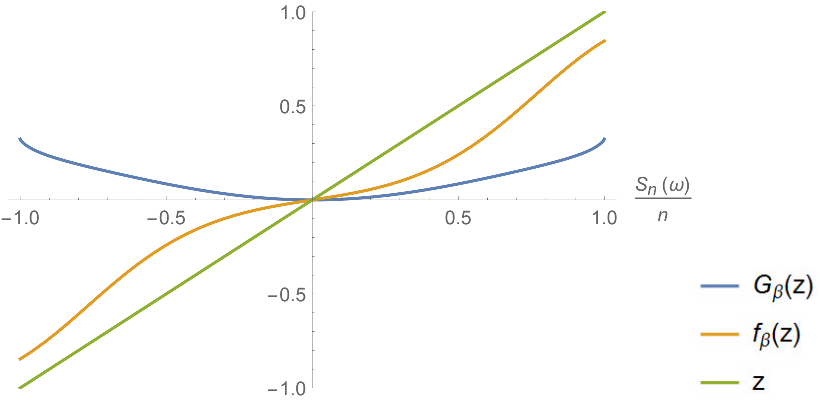


$\beta > \beta_C$

Phase Transition Structure

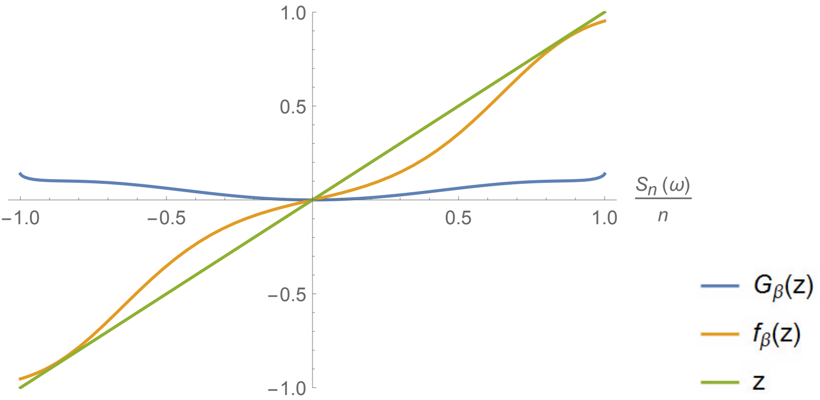


Disontinuous Phase Transition ($\alpha_1 > 2\alpha_2$)



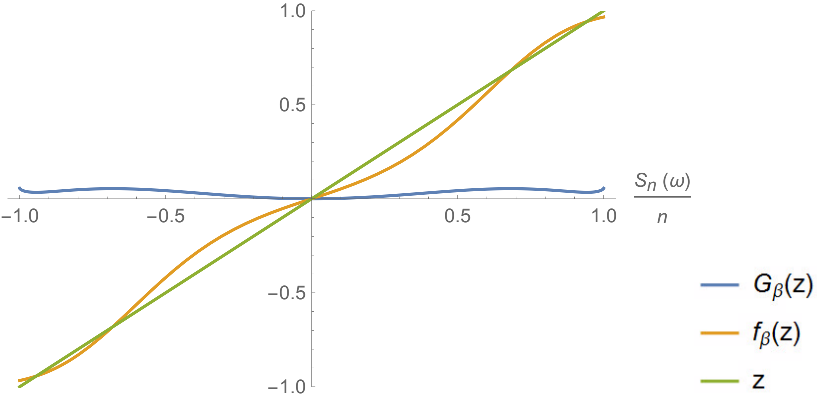
$$0 < \beta < \beta_M$$

Disontinuous Phase Transition ($\alpha_1 > 2\alpha_2$)



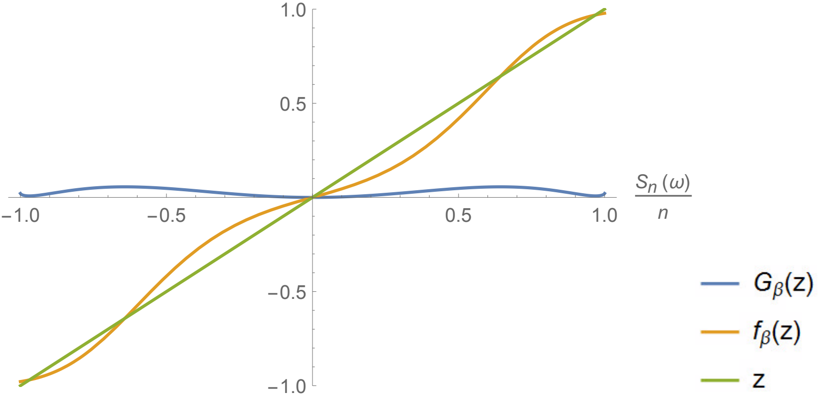
$$\beta = \beta_M$$

Disontinuous Phase Transition ($\alpha_1 > 2\alpha_2$)



$$\beta_M < \beta < \beta_C$$

Disontinuous Phase Transition ($\alpha_1 > 2\alpha_2$)



$\beta = \beta_C$

Curie-Weiss Conclusion

Main Results:

- $\beta_M = \sup_{\beta>0}\{\alpha'_\beta < 1\}$
- $\beta_C = \sup_{\beta>0}\{G_\beta(0) = G'_\beta(0) = 0, G''_\beta(0) \geq 0\}$
- For the second order continuous phase transition, $\beta_M = \beta_C$.
- For the first order discontinuous phase transition, $\beta_M < \beta_C$.

Acknowledgments

- Willamette Valley REU Consortium for Mathematics Research
- National Science Foundation for funding the REU
- Dr. Peter Otto, faculty mentor

References

- S. Kim,
  Mixing time of a Rook's Walk.
  (2012)
- D. Levin, Y. Peres, E. Wilmer,
  Markov Chains and Mixing Times.
  American Mathematical Society, USA (2009)
- C. Mcleman, P. Otto, J. Rahmani , M. Sutter
  Mixing Times For The Rook's Walk Via Path Coupling
  (2014)
- R. S. Ellis,
  Entropy, Large Deviations and Statistical Mechanics
  2006