Master Thesis

Ergodicity of the Dynamical XY-Model

David Geldbach

Supervisor: Prof. Markus Heydenreich
Ludwig-Maximilians-University Munich
August 9, 2021

Contents

Notation

Throughout this thesis we use some notational conventions which we explain here.

All spaces in this thesis are topological spaces: $\mathbb{R}^d$ and its subsets are equipped with the standard Euclidean topology. Similarly, we equip all manifolds, for example $\mathbb{S}^1$ the unit circle embedded in $\mathbb{R}^2$, with the standard topology. Countable sets carry the discrete topology. Consequently we always equip all topological spaces with their respective Borel sigma algebra. Hence we need not worry about measurability - as long as we consider continuous functions.

Given a measure space $(\Omega, \mathcal{F}, \mu)$ and a measurable function $f\colon \Omega \to \mathbb{R}$ we use different notations to denote the integral of f with respect to μ:

$$\int_\Omega f(x)\mu(dx) = \int f d\mu = \mu\big[f\big]$$

as long as the integral exists.

We often define probability measures with respect to a reference measure: assume we have given a measure space $(\Omega, \mathcal{F}, \mu)$, we define a new measure by

$$\nu(dx) = \rho(x)\mu(dx)$$

where ρ is a non-negative measurable function. This is to say that we define ν via its Radon-Nikodým derivative with respect to μ, i.e. $\frac{d\nu}{d\mu} = \rho$.

We often work with subsets of $\mathbb{Z}^d$, $B \Subset \mathbb{Z}^d$ denotes that $B \subset \mathbb{Z}^d$ is a finite subset.

Given a measurable space S, we define:

$$\begin{aligned} C(S) &= \big\{f\colon S \to \mathbb{R}, f \text{ continuous}\big\} \\ C_c(S) &= \big\{f \in C(S) : f \text{ is compactly supported}\big\} \end{aligned}$$

And if S is an open subset of $\mathbb{R}^d$, we further define:

$$C^k(S) = \big\{f \in C(S) : f \text{ is } k \text{ times continuously differentiable}\big\}$$

Given a measure space $(\Omega, \mathcal{F}, \mu)$ we define for $1 \le p < \infty$:

$$\begin{aligned} L^p(\mu) = L^p(\Omega, \mathcal{F}, \mu) &= \big\{f\colon \Omega \to \mathbb{R} \text{ measurable} : \|f\|_p = \mu\big[|f|^p\big]^{1/p} < \infty\big\} \\ L^\infty(\mu) = L^\infty(\Omega, \mathcal{F}, \mu) &= \big\{f\colon \Omega \to \mathbb{R} \text{ measurable} : \|f\|_\infty = ess\sup_{\omega\in\Omega} |f(\omega)| < \infty\big\} \end{aligned}$$

For two normed spaces $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ and a linear operator $A\colon X \to Y$, we define the corresponding operator norm

$$\|A\| = \|A\|_{X\to Y} = \sup_{f\colon X\to Y} \frac{\|Af\|_Y}{\|f\|_X}.$$

In the special case that $X = L^p(\mu)$ and $Y = L^q(\mu)$, we denote $\|A\|_{L^p\to L^q}$ by $\|A\|_{p,q}$.

1 Overview

In this thesis I will prove ergodicity of the XY-model with Glauber dynamics in one dimension ($d = 1$) for finite temperature and in higher dimensions ($d > 1$) for high temperature. This is achieved by showing a logarithmic Sobolev inequality for the Gibbs measure. In this short first chapter, I briefly introduce the dynamics of the XY-model, state the theorems that are being proven in the main part of this thesis, and explain how the thesis is structured.

Lattice models are a very important part of modern physics, in particular of statistical physics, where some of the most important exact solutions have been achieved for lattice models like the two-dimensional Ising-model. The Ising-model is a model for discrete spins on the $d-$dimensional lattice, with configurations taking values in $\{\pm 1\}^{\mathbb{Z}^d}$. A related but continuous spin model is the XY-model considered here, with configurations taking values in $(\mathbb{S}^1)^{\mathbb{Z}^d}$. Its probability distribution in a finite volume $\Lambda \subset \mathbb{Z}^d$ is given by

$$\mu_\Lambda(d\omega) = \frac{1}{Z} e^{-H_\Lambda(\omega)} \nu_\Lambda(d\omega),$$

where $H_\Lambda(\omega)$ is the energy of a configuration and ν_Λ the uniform distribution. The energy function is given by

$$H_\Lambda(\omega) = - \sum_{i,j \in \Lambda; |i-j|=1} \beta \langle \sigma_i, \sigma_j \rangle,$$

where σ_i is the spin at site i and $\beta > 0$ is a parameter corresponding to the inverse temperature, $\beta = 1/T$. This energy describes the tendency of neighbouring spins to align with each other. The probability measure which then describes the equilibrium in infinite volume is given by a Gibbs measure π. Similar to the Ising-model in $d = 2$, the XY-model undergoes a transition at a finite critical temperature, but due to its continuous character, this is not a conventional transition between a disordered and an ordered phase, but a change in correlations. In $d = 1$, both models do not have transitions at finite temperature.

We next add dynamics to the XY-model by defining a time-evolution operator P_t. Here we choose the Markov generator

$$\mathcal{L} = \sum_i \Delta_i - \nabla_i H \quad \cdot \nabla_i$$

which is also known as continuum Glauber dynamics and means that the spin at site i undergoes Brownian motion with diffusion and drift. The drift goes towards the local minimum of H, leading to local spin alignment.

The goal of this thesis is to show that Glauber dynamics of the XY-model converge to the Gibbs measure π in a uniform sense. One can only expect this to happen when π is unique, which automatically yields a necessary assumption: either we work on the one-dimensional lattice $\mathbb{Z}$, or we need to assume that $\beta < \beta_c(d)$ where $0 < \beta_c(d) \leq \infty$ is the uniqueness regime of the XY-model on $\mathbb{Z}^d$. The proof places an additional constraint on β, namely $\beta < \frac{1}{4d}$ which is why the result holds for $\beta < \tilde{\beta}_c = \min\{\beta_c, \frac{1}{4d}\}$.

We can now state the main theorems which are proven in the main part of this thesis.

Theorem: *For the XY model with unique Gibbs measure π, assume $d = 1$ or $\beta < \tilde{\beta}_c$. There are constants $\theta > 0, C$ such that for any differentiable f depending only on finitely many spins $\Lambda(f) \subset \mathbb{Z}^d$, we have*

$$\left\| P_t f - \pi[f] \right\|_\infty \leq C(\Lambda(f)) e^{-\theta t} \sum_{i \in \mathbb{Z}^d} \|\nabla_i f\|_\infty,$$

θ depends only on d, β and $C(\Lambda(f))$ depends on d, β and $\Lambda(f)$.

This means that P_t converges to π exponentially fast in a uniform sense and the associated process reaches equilibrium exponentially fast - irrespective of the starting configuration. The main tool needed to prove this theorem is a logarithmic Sobolev inequality for the unique Gibbs measure π:

Theorem: *For the XY model with unique Gibbs measure π, assume $d = 1$ or $\beta < \tilde{\beta}_c$. Then there is a constant $\alpha < \infty$ such that for any differentiable f depending only on finitely many spins we have*

$$Ent_\pi(f^2) = \pi\left[f^2 \log f^2\right] - \pi\left[f^2\right] \log \pi\left[f^2\right] \leq 2\alpha\pi\left[\left|\nabla f\right|^2\right].$$

This thesis starts by introducing the basics of the XY-model and of Markov processes in Chapter 2. In Chapters 3 and 4 I will present the spectral gap inequality and the logarithmic Sobolev inequality, respectively. In the final chapter 5, I will then prove the two theorems given here. The thesis finally closes with a short concluding chapter.

2 Introduction

2.1 Lattice models

The purpose of this section is to introduce the basics of lattice models from statistical mechanics, in particular the XY-model which we investigate here. We introduce Gibbs measures in finite and infinite volume. Here we need to prove that the Gibbs measure associated to the XY-model is unique in one dimensions for all temperatures, and in higher dimensions it is unique for sufficiently high temperatures. The main reference for this section is [9].

We consider spin configuration on the d-dimensional lattice $\mathbb{Z}^d$. This means that we are given a single-spin space Ω_0 which describes the possible values of a spin at a single site. Given $\Lambda \subset \mathbb{Z}^d$, Ω_Λ is the state of configuration of spins on Λ, i.e.:

$$\Omega_\Lambda = \Omega_0^\Lambda = \big\{(\omega_i)_{i\in\Lambda} : \omega_i \in \Omega_0 \ \forall i \in \Lambda\big\}$$

And if $\Lambda = \mathbb{Z}^d$, we abbreviate $\Omega = \Omega_{\mathbb{Z}^d}$. For two subsets $\Delta \subset \Lambda \subset \mathbb{Z}^d$ there are embeddings

$$\Omega_\Delta \hookrightarrow \Omega_\Lambda \hookrightarrow \Omega$$

by choosing the values of, say $\Lambda\backslash\Delta$ as some reference configuration. To make this a bit more precise, assume we have a reference configuration $\eta \in \Omega$ and $\omega \in \Omega_\Lambda$. We define a new configuration $\omega_\Lambda\eta_{\Lambda^c}$ as follows:

$$(\omega_\Lambda\eta_{\Lambda^c})_i = \begin{cases} \omega_i & \text{if } \ i \in \Lambda \\ \eta_i & \text{if } \ i \in \Lambda^c \end{cases}$$

Further we impose some technical assumptions on Ω_0: we want it be a Polish space. Henceforth, $(\Omega_\Lambda)_{\Lambda\subset\mathbb{Z}^d}$ are Polish as well when equipped with the product topology. We use the induced Borel σ-algebras, respectively.

More importantly, we impose a non-technical assumption on Ω_0 in this exposé: we require it to be compact. This is the case for the $XY-$model where we choose Ω_0 to be the unit circle in $\mathbb{R}^2$ and this is the case for the most famous model of this kind, the Ising model, where we have $\Omega_0 = \{\pm 1\}$.

Before we can define any measure on $(\Omega_\Lambda)_{\Lambda\subset\mathbb{Z}^d}$, we need a notion of energy of a given configuration. We assume straight away that our energy function H_Λ, which is called Hamiltonian, is derived from a translation invariant potential [9, Def. 6.14]:

Definition 2.1: *Assume that for each $0 \in B \Subset \mathbb{Z}^d$ we are given a continuous $\Phi_B \colon \Omega^B \to \mathbb{R}$. For $\Lambda \Subset \mathbb{Z}^d$, the associated Hamiltonian is defined as*

$$H_\Lambda(\omega) = \sum_{i\in\Lambda} \sum_{0\in B\Subset\mathbb{Z}^d} \Phi_{i+B}(\omega) \quad \forall \omega \in \Omega.$$

We remark that one needs to impose a condition of Φ such that H is well defined: we assume $(\Phi_B)_{B\Subset\mathbb{Z}^d}$ to be absolutely summable in the sense that

$$\sum_{0\in B\Subset\mathbb{Z}^d} \|\Phi_B\|_\infty < \infty.$$

Lastly, we need to assume that we are given a reference measure ν on Ω_0, to avoid problems later, assume that ν is a probability measure, i.e. $\nu(\Omega_0) = 1$. Naturally, we equip Ω_Λ with the product measure $\nu_\Lambda = \nu^{\otimes\Lambda}$. This allows us to define Gibbs measures in finite volume:

Definition 2.2: *For fixed $\Lambda \Subset \mathbb{Z}^d$ and boundary condition $\eta \in \Omega$, and given Ω_0, H and ν as above, we define a probability measure on Ω_Λ*

$$\mu_\Lambda^\eta(d\omega) = \frac{1}{Z_\Lambda^\eta} e^{-H_\Lambda(\omega_\Lambda\eta_{\Lambda^c})} \nu_\Lambda(d\omega),$$

where the partition sum Z^η_Λ is defined as

$$Z^\eta_\Lambda = \int_{\Omega_\Lambda} e^{-H_\Lambda(\omega_\Lambda \eta_{\Lambda^c})} \nu_\Lambda(d\omega).$$

This probability measure is extended to Ω by choosing the reference configuration as η.

Due to our assumptions on H and ν, we have $0 < Z^\eta_\Lambda < \infty$ and therefore μ^η_Λ is well defined.

Definition 2.3: We can now introduce the model we are most interested in: the XY-model. Recall that we need to specify Ω_0, $(H_\Lambda)_{\Lambda \Subset \mathbb{Z}^d}$ and ν. We choose Ω_0 to be the unit circle embedded in $\mathbb{R}^2$

$$\Omega_0 = \big\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\big\},$$

equipped with the usual topology. Then we choose ν to be the normalised Lebesgue measure on Ω_0, which can be described as the push-forward of the Lebesgue measure on $[0,1[$ under the map $t \mapsto (\cos(2\pi t), \sin(2\pi t))$.

The Hamiltonian is given as a nearest neighbour interaction

$$H_\Lambda(\omega) = -\beta \sum_{\{i,j\}:\{i,j\}\cap\Lambda\neq\emptyset, |i-j|=1} \langle \omega_i, \omega_j \rangle,$$

where $\beta \geq 0$ is an order parameter which is called the *inverse temperature* and $\langle \cdot,\cdot \rangle$ denotes the standard inner product of $\mathbb{R}^2$. We always leave the dependence on β implicit.

Often we view the circle not as subset of $\mathbb{R}^2$ but as ${}^{[0,1]}/_{(0\sim 1)}$. This is done by using the homeomorphism

$$[0,1[\ni t \mapsto \big(\cos(t), \sin(t)\big) \in \big\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\big\}.$$

The Hamiltonian then becomes:

$$H^\eta_\Lambda(\omega) = -\beta \sum_{i,j\in\Lambda, |i-j|=1} \cos(2\pi(\omega_i - \omega_j)) - \beta \sum_{i\in\Lambda, j\in\Lambda^c, |i-j|=1} \cos(2\pi(\omega_i - \eta_j)),$$

for a configuration $\omega \in [0,1]^\Lambda$ with boundary condition η. We stress the dependence on η by writing H^η_Λ and we abuse notation and use the same letter for a configuration in $(\mathbb{S}^1)^{\mathbb{Z}^d}$ and the corresponding configuration in $[0,1]^{\mathbb{Z}^d}$.

Definition 2.4: *For the sake of completeness, we also introduce the Ising model. Here we have $\Omega_0 = \{\pm 1\}$ and ν is the uniform distribution on Ω_0. Just like in the XY-model, the Hamiltonian is a nearest neighbour interaction*

$$H_\Lambda(\omega) = -\beta \sum_{\{i,j\}:\{i,j\}\cap\Lambda\neq\emptyset, |i-j|=1} \omega_i \omega_j.$$

One should note that the family of measures $\{\mu^\eta_\Lambda : \Lambda \Subset \mathbb{Z}^d, \eta \in \Omega\}$ is actually a family of probability kernels: $\eta \mapsto \mu^\eta_\Lambda(A)$ for each measurable $A \subset \Omega_\Lambda$ is a measurable map with respect to the product σ-algebra of Ω_{Λ^c}. And for each $\eta \in \Omega$, $\mu^\eta_\Lambda(\cdot)$ is a probability measure on Ω by definition.

In light of this technical property, we also write $\mu^\eta_\Lambda(A) = \mu_\Lambda(A|\eta)$. Now given another probability measure π on Ω, we can define a composition of π with μ_Λ

$$\pi\mu_\Lambda(A) = \int_\Omega \mu_\Lambda(A|\eta)\pi(d\eta), \quad A \subset \Omega \text{ measurable.}$$

Probabilistically, this corresponds to sampling a boundary condition η using π first, and the using μ^η_Λ to sample the values of the spins in Λ. A natural choice for π arises when we have $\Delta \subset \Lambda \Subset \mathbb{Z}^d$ and a fixed $\eta \in \Omega$. Similar to $\pi\mu_\Lambda$ we define $\mu_\Lambda\mu_\Delta$:

$$\mu_\Lambda\mu_\Delta(A|\eta) = \int_{\Omega_\Lambda} \mu_\Delta(A|\omega)\mu_\Lambda(d\omega|\eta), \quad A \subset \Omega \text{ measurable.}$$

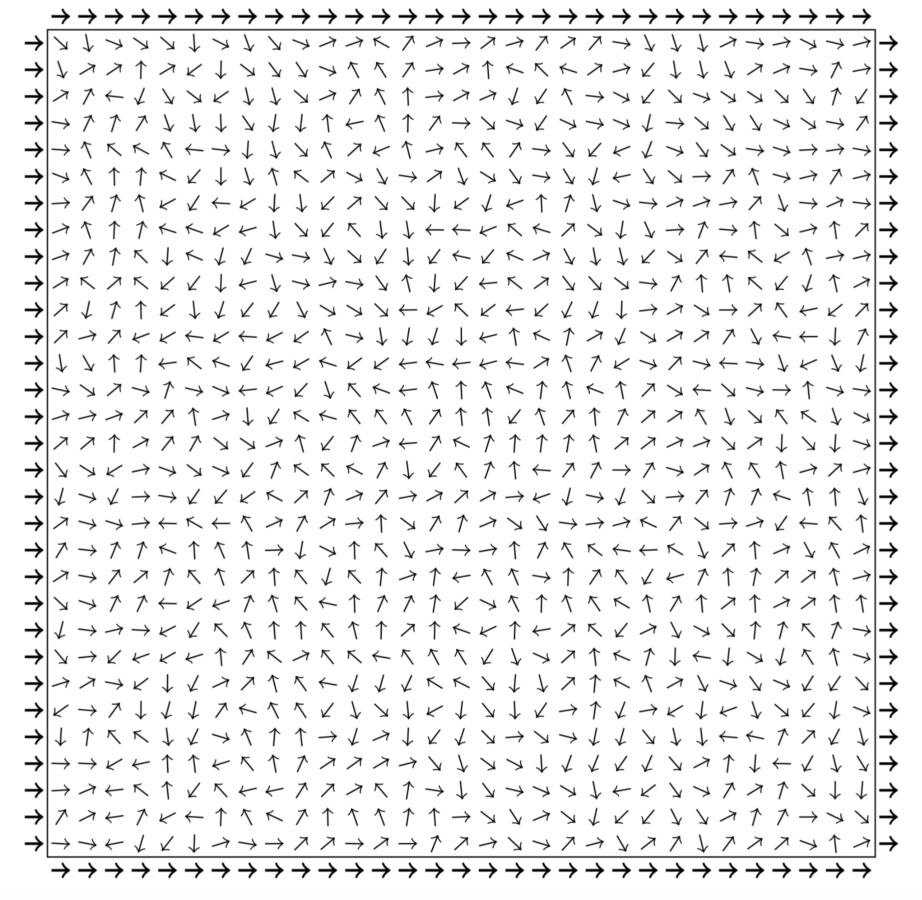


Figure 1: A typical configuration of the XY-model for β small, taken from [9]. The thick arrows represent the boundary condition η.

This corresponds to sampling the configuration in Λ in two steps, first in $\Lambda\backslash\Delta$ and then in Δ. Computing this expression yields the consistency of the finite volume Gibbs in the following sense:[1]

Lemma 2.5: *For* $\{\mu_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ *defined as above,* $\Delta \subset \Lambda \Subset \mathbb{Z}^d$ *and any* $\eta \in \Omega$ *we have*

$$\mu_\Lambda\mu_\Delta(\cdot|\eta) = \mu_\Lambda(\cdot|\eta).$$

Proof. In the case of $\Omega_0 = \{\pm 1\}$ an analogous lemma can be found in [9, Lemma 6.15]. We adapt the proof to arbitrary Ω_0. Let $f\colon \Omega_\Lambda \to \mathbb{R}$ be measurable and bounded. We compute $\mu_\Lambda\mu_\Delta(f|\eta)$:

[1]This renders $\{\mu_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ a specification, more precisely a Gibbsian specification [9, Def. 6.11]

$$\begin{aligned}
\mu_\Lambda\mu_\Delta(f|\eta) &= \int_{\Omega_\Lambda}\int_{\Omega_\Delta} f(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})\mu_\Delta(d\sigma|\omega)\mu_\Lambda(d\omega|\eta)\\
&= \frac{1}{Z^\eta_\Lambda}\int_{\Omega_\Lambda}\frac{1}{Z^\omega_\Delta}\int_{\Omega_\Delta} f(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})e^{-H_\Delta(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}e^{-H_\Lambda(\omega_\Lambda\eta_{\Lambda^c})}\nu_\Delta(d\sigma)\nu_\Lambda(d\omega)\\
&= \frac{1}{Z^\eta_\Lambda}\int_{\Omega_{\Lambda\setminus\Delta}}\int_{\Omega_\Delta}\int_{\Omega_\Delta}\frac{1}{Z^\omega_\Delta} f(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})\\
&\qquad\qquad \cdot e^{-H_\Delta(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}e^{-H_\Lambda(\omega_\Lambda\eta_{\Lambda^c})}\nu_\Delta(d\sigma)\nu_\Delta(d\omega_\Delta)\nu_{\Lambda\setminus\Delta}(d\omega_{\Lambda\setminus\Delta})\\
&= \frac{1}{Z^\eta_\Lambda}\int_{\Omega_{\Lambda\setminus\Delta}}\int_{\Omega_\Delta} f(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})e^{-H_\Lambda(\sigma_\Delta\omega_{\Lambda\setminus\Delta}\eta_{\Lambda^c})}\nu_\Delta(d\sigma)\nu_{\Lambda\setminus\Delta}(d\omega_{\Lambda\setminus\Delta})\\
&= \mu_\Lambda(f|\eta)
\end{aligned}$$

where we used the definition of Z^ω_Δ. Because the above identity is true for any measurable and bounded f, we obtain the lemma. □

This leads us to a reasonable definition of Gibbs measures in infinite volume: the so-called DLR-condition which is named after Dobrushin, Lanford and Ruelle.

Definition 2.6: *A measure π on Ω is called infinite volume Gibbs measure (or simply Gibbs measure) if it satisfies*

$$\pi\mu_\Lambda = \pi$$

for any $\Lambda \Subset \mathbb{Z}^d$. We denote the set of all Gibbs measures by $\mathscr{G}(\Phi)$, where Φ indicates the dependence on the potential involved in the definition of $\{\mu_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$.

The immediate question is that of existence. If Ω_0 is compact, then an infinite volume Gibbs measure always exists. This question becomes more intricate when this is note the case, see [9].

Theorem 2.7: *For $\{\mu_\Lambda\}_{\Lambda\Subset\mathbb{Z}^d}$ defined as above and Ω_0 compact, we have $\mathscr{G}(\Phi) \neq \emptyset$.*

Proof. This is [9, Thm. 6.26] in the case of $\Omega_0 = \{\pm 1\}$, we adapt the proof to compact Ω_0.

The main tool need is Prokhorov's theorem: because Ω_0 is compact and Polish, so is Ω and therefore the space of probability measures on Ω is sequentially compact with respect to weak convergence.

Choose and arbitrary $x_0 \in \Omega_0$ and let $\eta = (\eta_i)_{i\in\mathbb{Z}^d}$ with $\eta_i = x_0$ for every $i \in \mathbb{Z}^d$. Further let $\Lambda_n = \{-n, ..., n\}^d$. Define $\mu_n = \mu^\eta_{\Lambda_n}$. By Prokhorov's theorem $(\mu_n)_{n\geq 1}$ is sequentially compact, hence we can pick a subsequence $(n_k)_{k\geq 1}$ so that there exists a probability measure π with $\mu_{n_k} \rightharpoonup \pi$ weakly as $k \to \infty$. We show that π satisfies the DLR-condition in Definition 2.6.

For this fix an arbitrary continuous and bounded $f\colon \Omega \to \mathbb{R}$ and fix $\Lambda \Subset \mathbb{Z}^d$. By our assumptions on the potential Φ, we have that $\omega \mapsto \mu^\omega_\Lambda[f]$ is also continuous and bounded. Now choose k_0 big enough such that $\Lambda \subset \Lambda_{n_k}$. Then we have:

$$\pi\mu_\Lambda[f] = \lim_{\substack{k\to\infty\\ k>k_0}} \underbrace{\mu_{\Lambda_{n_k}}\mu_\Lambda[f|\eta]}_{=\mu_{\Lambda_{n_k}}[f|\eta]} = \lim_{\substack{k\to\infty\\ k>k_0}} \mu_{n_k}[f] = \pi[f]$$

where the first and last equality follow from the convergence of the μ_{n_k} and the middle equality follows from Lemma 2.5 which can be lifted to expectations of bounded functions by measure theoretic induction.

Because the above equality is true for any f, we have $\pi\mu_\Lambda = \pi$. And because this is true for any Λ, π satisfies the DLR-condition and therefore $\pi \in \mathscr{G}(\Phi)$. □

Having settled the question of existence, the next natural question is uniqueness of Gibbs measures. In fact, this is a very hard question and we only answer it partially and only for the XY-model: if $d = 1$ or if β is very small, then the Gibbs measure is unique.

Theorem 2.8: *For the XY−model - as defined in Definition 2.3 - we have two sufficient conditions for uniqueness:*

(a) $d = 1$ *and* $\beta > 0$ *arbitrary.*

(b) $d \geq 2$ *and* $0 < \beta < \beta_c(d)$.

for some $0 < \beta_c(d) \leq \infty$.

A significant part of the theory of lattice models deals with proving uniqueness or non-uniqueness of Gibbs measures. In particular, one could show that

$$\beta_c = \sup_{\beta>0}\{\text{the Gibbs measure at inverse temperature } \beta \text{ is unique}\}.$$

Furthermore, in $d \geq 3$ we indeed have that $\beta_c < \infty$, see [9].

We now turn to the proof of Theorem 2.8. As one would expect, if the finite dimensional measures μ_Λ^η depend only weakly on the boundary condition η, then the Gibbs measure in infinite volume is unique. This is illustrated by the following lemma:

Lemma 2.9: *We have* $\mathscr{G}(\Phi) = \{\pi\}$ *if*

$$\sup_{\eta,\omega\in\Omega} \left|\mu^\eta_{\Lambda_n}[f] - \mu^\omega_{\Lambda_n}[f]\right| \longrightarrow 0,$$

for all local, measurable and bounded f *and some sequence* $\Lambda_n \uparrow \mathbb{Z}^d$ *such that* $\Lambda(f) \subset \Lambda_n$ *for all* n.

Proof. This lemma is similar to [9, Lemma 6.30] and consists of applying the DLR-condition. Assume we have $\pi_1, \pi_2 \in \mathscr{G}(\Phi)$. By the definition of Gibbs measures we have for any f like in the lemma:

$$\begin{aligned}\left|\pi_1[f] - \pi_2[f]\right| = \left|\ \pi_1\mu_{\Lambda_n}[f] - \pi_2\mu_{\Lambda_n}[f]\right| &= \left|\iint \mu^\eta_{\Lambda_n}[f] - \mu^\omega_{\Lambda_n}[f]\pi_1(d\eta)\pi_2(d\omega)\right| \\ &\leq \iint \left|\mu^\eta_{\Lambda_n}[f] - \mu^\omega_{\Lambda_n}[f]\right|\pi_1(d\eta)\pi_2(d\omega) \xrightarrow{n\to\infty} 0\end{aligned}$$

by use of the dominated convergence theorem. Hence $\pi_1[f] = \pi_2[f]$ for all appropriate f and thus $\pi_1 = \pi_2$. This means that the Gibbs measure is unique. □

For the XY−model on $\mathbb{Z}$ we can directly estimate $\sup_{\eta,\omega\in\Omega}\left|\mu^\eta_\Lambda[f] - \mu^\omega_\Lambda[f]\right|$ and we can even show that this quantity decays exponentially:

Lemma 2.10: *For the one-dimensional XY-model and any* $\beta > 0$ *we have: there exists* $\gamma > 0$ *such that for any continuous function* $f\colon (\mathbb{S}^1)^\mathbb{Z} \to \mathbb{Z}$ *with* $\Lambda(f) = \{-N, ..., N\}$ *for some* N *we have:*

$$\sup_{x_1,y_1,x_2,y_2\in\mathbb{S}^1}\left|\mu^{\{x_1,x_2\}}_{\{-n,...,n\}}[f] - \mu^{\{y_1,y_2\}}_{\{-n,...,n\}}[f]\right| \leq C(f)e^{-\gamma(n-N)},$$

for all $n \geq N$ *and for some* $C(f)$ *which depends only on* f*. Here,* $\mu^{\{x_1,x_2\}}_{\{-n,...,n\}}$*, denotes the finite volume Gibbs measure with boundary condition* x_1 *for the spin at* $-(n+1)$ *and* x_2 *for the spin at* $n+1$.

Proof. We postpone the proof to Section 5.1 and revisit this lemma as Lemma 5.3 in a slightly different formulation. □

A similar result holds for the XY-model on $\mathbb{Z}^d$ for β small:

Proposition 2.11: *For the XY-model on* $\mathbb{Z}^d$*: there exists* $0 < \beta_c(d) < \infty$ *such that for any* $\beta < \beta_c(d)$ *the limit*

$$\lim_{\Lambda \uparrow \mathbb{Z}^d} \mu_\Lambda^\eta[f] = \ell(f)$$

independent of $\eta \in \Omega$ *and for any local, continuous* $f \colon \Omega \to \mathbb{R}$*. Furthermore, there exists* $\gamma > 0$ *such that for any such* f *we have:*

$$\left|\mu_\Lambda^\eta[f] - \ell(f)\right| \leq C(f) e^{-\gamma dist(\Lambda, \Lambda(f))},$$

where $C(f)$ *is a constant that depends only on* f*.*

Proof of Theorem 2.8. The uniqueness of the Gibbs measure for the one-dimensional XY-model now follows directly from Lemma 2.9 and Lemma 2.10. Respectively, the uniqueness of Gibbs measure for the XY-model on $\mathbb{Z}^d, d \geq 2$ follows from Lemma 2.9 and the previous proposition. □

Sketch of the proof of Proposition 2.11. This proposition is similar to [9, Prop. 6.39] which concerns itself with models on $\{\pm 1\}^{\mathbb{Z}^d}$. The proof uses the so-called cluster expansion and we sketch this approach.

Fix f, η as required and let $\Lambda \Subset \mathbb{Z}^d$ such that $\Lambda(f) \subset \Lambda$. To rewrite $\mu_\Lambda^\eta[f]$, we denote $Z_\Lambda^\eta = \int_{(\mathbb{S}^1)^\Lambda} \exp(-H_\Lambda(\sigma_\Lambda \eta_{\Lambda^c})) \nu_\Lambda(d\sigma)$ and let $\Delta = \Lambda \backslash \Lambda(f)$:

$$\begin{aligned}
\mu_\Lambda^\eta[f] &= \frac{1}{Z_\Lambda^\eta} \int_{(\mathbb{S}^1)^\Lambda} f(\sigma) e^{-H_\Lambda(\sigma_\Lambda \eta_{\Lambda^c})} \nu_\Lambda(d\sigma) \\
&= \frac{1}{Z_\Lambda^\eta} \int_{(\mathbb{S}^1)^{\Lambda(f)}} \underbrace{f(\sigma) e^{\beta \sum_{i,j \in \Lambda(f): |i-j|=1} \langle \sigma_i, \sigma_j \rangle}}_{:= F(\sigma_{\Lambda(f)})} \underbrace{\int_{(\mathbb{S}^1)^\Delta} e^{-H_\Delta(\sigma_{\Lambda(f)} \sigma_\Delta \eta_{\Lambda^c})} \nu_\Delta(d\sigma_\Delta)}_{= Z_\Delta^{\sigma_{\Lambda(f)} \eta_{\Lambda^c}}} \nu_{\Lambda(f)}(d\sigma_{\Lambda(f)}) \\
&= \int_{(\mathbb{S}^1)^{\Lambda(f)}} F(\sigma_{\Lambda(f)}) \frac{Z_\Delta^{\sigma_{\Lambda(f)} \eta_{\Lambda^c}}}{Z_\Lambda^\eta} \nu_{\Lambda(f)}(d\sigma_{\Lambda(f)}) \qquad (2.1)
\end{aligned}$$

The only quantity that depends on Λ and η is the ratio in (2.1). Hence we want to estimate this ratio, we present an analysis of Z_Λ^η. From now on we want to emphasise the dependence on β more, let $V_{i,j}(\sigma) = -\langle \sigma_i, \sigma_j \rangle \mathbb{1}_{|i-j|=1}$ the interaction of the spins at i and j. We then have:

$$e^{-H_\Lambda} = \prod_{\{i,j\} \cap \Lambda \neq \emptyset} e^{-\beta V_{i,j}} = \prod_{\{i,j\} \cap \Lambda \neq \emptyset} \left(e^{-\beta V_{i,j}} - 1 + 1\right) = \sum_{B \in \mathscr{B}} \prod_{\{i,j\} \in B} \left(e^{-\beta V_{i,j}} - 1\right), \qquad (2.2)$$

where $\mathscr{B}$ is the powerset of $G = \left\{\{i,j\} : \{i,j\} \cap \Lambda \neq \emptyset, |i-j| = 1\right\}$. We endow G with a graph structure: the elements of G are vertices and $\{i_1, j_1\}, \{i_2, j_2\}$ are connected if and only if $\{i_1, j_1\} \cap \{i_2, j_2\} \neq \emptyset$. Let $\mathscr{S}$ be the set of subgraphs of G and $\mathcal{C}(S), S \in \mathscr{S}$ be the set of connected components of S. (2.2) then becomes

$$e^{-H_\Lambda} = \sum_{S \in \mathscr{S}} \prod_{C \in \mathcal{C}(S)} \prod_{\{i,j\} \in C} \left(e^{-\beta V_{i,j}} - 1\right).$$

The advantage of this decomposition is that if $C \cap C' = \emptyset$ then $\prod_{\{i,j\} \in C} \left(e^{-\beta V_{i,j}} - 1\right)$ and $\prod_{\{i,j\} \in C'} \left(e^{-\beta V_{i,j}} - 1\right)$ do not depend on the same spins which causes the integral in Z_Λ^η to factorise, let $\overline{C} = \bigcup_{\{i,j\} \in C} \{i,j\}$:

$$\begin{aligned} Z_\Lambda^\eta &= \int_{(\mathbb{S}^1)^\Lambda} e^{-H_\Lambda(\sigma_\Lambda \eta_{\Lambda^c})} \nu_\Lambda(d\sigma) \\ &= 2^{|\Lambda|} \sum_{S \in \mathscr{S}} \prod_{C \in \mathcal{C}(S)} \underbrace{2^{-|\overline{C} \cap \Lambda|} \int_{(\mathbb{S}^1)^{\overline{C}}} \prod_{\{i,j\} \in C} \big(e^{-\beta V_{i,j}(\sigma_\Lambda \eta_{\Lambda^c})} - 1\big) \nu_{\overline{C}}(d\sigma)}_{=w(C)} \\ &= 2^{|\Lambda|} \sum_{S \in \mathscr{S}} \prod_{C \in \mathcal{C}(S)} w(C) \end{aligned} \tag{2.3}$$

Implicitly, the weight of C, $w(C)$, depends on Λ and η.

We have now arrived at the point where the formalism of cluster expansion is used. This formalism allows us to rewrite the sum of (2.3) as $\exp\big(\sum_{n\geq 0}[...]\big)$ and can be found in [9, Chapter 5]. In the end we are given an expansion of the form

$$\log\Big(\sum_{S \in \mathscr{S}} \prod_{C \in \mathcal{C}(S)} w(C)\Big) = 1 + \sum_{n \geq 1} \sum_{C_1,...,C_n \in \chi_\Lambda} \Psi_{\Lambda,\eta}(C_1, ..., C_n), \tag{2.4}$$

where the C_i are connected subgraphs like above, possibly containing duplications, χ_Λ indicates the dependence of the C_i on Λ, and Ψ is of the following form:

$$\Psi_{\Lambda,\eta}(C_1, ..., C_n) = \Big(\prod_{i=1}^n w(C_i)\Big) \cdot \big(\text{combinatorial factors}\big).$$

For the series in (2.4) to converge absolutely, we need the weights $w(C)$ to be small. This can be achieved by noting that

$$|w(C)| \leq \prod_{\{i,j\} \in C} \|e^{-V_{i,j}} - 1\|_\infty = \prod_{\{i,j\} \in C} (e^\beta - 1) \leq (e^\beta - 1)^{|\overline{C}|}$$

can be made arbitrarily small if β is small, the last inequality holds if $e^\beta - 1 \leq 1$.

Once the expansion (2.4) is established, we can consider the ratio in (2.1):

$$\begin{aligned} \frac{Z_\Delta^{\sigma_{\Lambda(f)} \eta_{\Lambda^c}}}{Z_\Lambda^\eta} &= 2^{-|\Lambda(f)|} \frac{\exp\big(\sum_{n \geq 1, C_1,...,C_n \in \chi_\Delta} \Psi_{\Delta, \sigma_{\Lambda(f)} \eta_{\Lambda^c}}(C_1, ..., C_n)\big)}{\exp\big(\sum_{n \geq 1, C_1,...,C_n \in \chi_\Lambda} \Psi_{\Lambda,\eta}(C_1, ..., C_n)\big)} \\ &= 2^{-|\Lambda(f)|} \frac{\exp\big(\sum_{n \geq 1, C_1,...,C_n \in \chi_\Delta; \bigcup \overline{C_i} \cap \Lambda(f) \neq \emptyset} \Psi_{\Delta, \sigma_{\Lambda(f)} \eta_{\Lambda^c}}(C_1, ..., C_n)\big)}{\exp\big(\sum_{n \geq 1, C_1,...,C_n \in \chi_\Lambda; \bigcup \overline{C_i} \cap \Lambda(f) \neq \emptyset} \Psi_{\Lambda,\eta}(C_1, ..., C_n)\big)} \end{aligned} \tag{2.5}$$

The cancellation in the second identity comes from the fact that all the terms that do not intersect $\Lambda(f)$ are present both in the numerator and denominator.

Recall that we do not want to keep Λ and $\Delta = \Lambda \backslash \Lambda(f)$ fixed but we want to consider $\Lambda_n \uparrow \mathbb{Z}^d$. Due to the condition that $\bigcup_{i=1}^n \overline{C_i} \cap \Lambda(f) \neq \emptyset$, the contribution of the additional terms in series becomes negligible and the fraction $Z_{\Lambda_n \backslash \Lambda(f)}^{\sigma_{\Lambda(f)} \eta_{\Lambda^c}} / Z_{\Lambda_n}^\eta$ converges. Furthermore, in (2.5) η influences only the terms that satisfy both $\bigcup_{i=1}^n \overline{C_i} \cap \Lambda(f) \neq \emptyset$ and $\bigcup_{i=1}^n \overline{C_i} \cap \Lambda^c \neq \emptyset$. These terms have weights of order smaller than $(e^\beta - 1)^{dist(\Lambda_n, \Lambda(f))}$ which is why the limit is independent of η. Further, this hints at the exponential rate of convergence in the proposition. This completes our sketch of the proof. □

2.2 Markov Processes

In this section we recall the basic notions of Markov processes in particular the relation of the generator with the reversible measures of the process. Further, we introduce the framework which we need later to introduce spectral gap and logarithmic Sobolev inequalities.

We do recall the most basic definitions and refer to [12] for that, in particular the third chapter. We also refer to [15] for more definitions, especially more properties of the carré du champ.

We assume that our underlying space S is a locally compact and separable topological space equipped with its Borel $\sigma-$algebra $\mathcal{S}$. Mainly we deal with $\mathbb{R}, \mathbb{S}^1$ or $[0,1]$ and products thereof. We use the correspondence of $S-$valued stochastic processes $(X_t)_{t\geq 0}$ that satisfy the strong Markov property, the semi-groups of operators $(P_t)_{t\geq 0}$ acting on some Banach space $\mathcal{B}$ of functions on S, equipped with the $\|\cdot\|_\infty$-norm, and probability generators $\mathcal{L}$ defined on some subset of $\mathcal{B}$. A priori, we assume $\mathcal{B}$ to be a subspace of $C(S)$ and that $\mathcal{B}$ contains constant functions. Recall what it means for $(P_t)_{t\geq 0}$ to be a probability semi-group:

Definition 2.12: *A family of operators $(P_t)_{t\geq 0}$ acting on $\mathcal{B}$ is called a probability semi-group if the following conditions are satisfied for all $f \in \mathcal{B}$:*

1. $P_0 f = f$.
2. $\lim_{t\to 0} P_t f = f$.
3. $P_t P_s f = P_{t+s}$ *for all* $t, s \geq 0$.
4. $P_t f \geq 0$ *for all* $f \geq 0$ *and* $t \geq 0$.
5. $P_t 1 = 1$ *for the constant function* 1 *and any* $t \geq 0$.

We then define $\mathcal{L}f = \lim_{t\to 0} \frac{1}{t}(P_t f - f)$ whenever the limit exists. The set of function for which $\mathcal{L}$ is well defined is the domain of $\mathcal{L}$, denoted by $D(\mathcal{L})$. As a matter of fact, $\mathcal{L}$ defines $(P_t)_t$ uniquely. Further recall that $D \subset D(\mathcal{L})$ is called a *core* for $\mathcal{L}$ if the closure of $\mathcal{L}\big|_D$ is $\mathcal{L}$, i.e. if $\mathcal{L}$ is uniquely determined by its values on D [12, Def. 3.31].

Throughout this thesis, we are interested in the invariant measures of the process. A measure μ is called invariant (or stationary) with respect to the semi-group $(P_t)_t$ if $\int P_t f \; d\mu = \int f \; d\mu$ for all $t \geq 0$ and $f \in \mathcal{B}$. A convenient criterion to check invariance is the following, for the proof see [12, Thm. 3.37]:

Theorem 2.13: *A measure μ is invariant if and only if $\int \mathcal{L}f \; d\mu = 0$ for all $f \in D$ where D is a core of $\mathcal{L}$.*

Note that we do not require μ to be a finite measure here. Of course, this criterion is only of any use if we are able to guess a good candidate for μ. And even if we should succeed in that regard, this leaves two questions: uniqueness of the invariant measure and convergence to it. These questions will be discussed over the course of the next chapters.

If we are given an invariant measure μ, we implicitly use that the operators (P_t) can be extended to $L^p(\mu)$ for any $p \geq 1$ which is defined in the usual way [15, Property 1.14] using Hahn-Banach's theorem. In particular this allows us to define:

Definition 2.14: *A measure μ is called reversible if for all $f, g \in L^2(\mu)$:*

$$\int f(P_t g) \; d\mu = \int (P_t f) g \; d\mu.$$

By choosing $g = 1$ we see that reversible probability measure are also invariant. Further, this renders $(P_t)_t$ and $\mathcal{L}$ symmetric operators on $L^2(\mu)$ and $L^2(\mu) \cap D(\mathcal{L})$ respectively. The resulting formula

$$\int f(\mathcal{L}g) \; d\mu = \int (\mathcal{L}f) g \; d\mu$$

is also called the integration by parts formula for $\mathcal{L}$ which is equivalent to μ being reversible.

In this chapter and the following ones we always need an assumption which justifies that our calculations are well defined: the existence of a core which contains only well behaved functions, compare to [1, Def. 2.4.2] or [2].

Hypothesis 2.15: *For a generator $\mathcal{L}$ with invariant probability measure μ, we assume the existence of an algebra $\mathcal{A} \subset D(\mathcal{L})$ which is a core for $\mathcal{L}$. $\mathcal{A}$ is assumed to be dense in $L^p(\mu), 1 \leq p < \infty$, stable under multiplication, addition and composition with smooth functions. Further we assume that if $f \in \mathcal{A}$ then $\mathcal{L}f \in \mathcal{A}$ and $P_t f \in \mathcal{A}$ for all $t \geq 0$. Lastly, we assume that $\mathcal{A}$ contains all constant functions.*

In specific examples we can check that this hypothesis is actually satisfied. Note that when we define $\mathcal{L}$ only on $\mathcal{A}$ it is technically not a generator on the bigger space $\mathcal{B}$ but rather a pre-generator. The unique closure of $\mathcal{L}$ then is a generator. We always abuse notation and call both objects $\mathcal{L}$, for more details check [12].

We need one more operator related to $\mathcal{L}$:

Definition 2.16: *To a generator $\mathcal{L}$, we associate its bilinear carré du champ operator Γ_1*

$$\Gamma_1(f,g) = \tfrac{1}{2}\big[\mathcal{L}(fg) - f(\mathcal{L}g) - g(\mathcal{L}f)\big],$$

for all $f, g \in \mathcal{A}$.

Given a reversible measure μ, observe that by using Theorem 2.13 and the integration by parts formula for $\mathcal{L}$ we obtain the following:

$$\int \Gamma_1(f,g) = \tfrac{1}{2}\int \mathcal{L}(fg)\ d\mu + \int f(-\mathcal{L}g)\ d\mu = \int f(-\mathcal{L}g)\ d\mu$$

In particular, using that $-\mathcal{L}$ is a positive operator, we have that $\mu\big[\Gamma_1(f,f)\big] \geq 0$. This can also be seen by the following characterisation of Γ_1, $f \in \mathcal{A}$:

$$\Gamma_1(f,f) = \frac{d}{dt}\frac{1}{2}\Big(P_t(f^2) - (P_t f)^2\Big)\Big|_{t=0} = \lim_{t\to 0}\frac{1}{2t}\big(P_t(f^2) - (P_t f)^2\big) \geq 0$$

Further, this already hints at $\mu\big[\Gamma_1(f,f)\big]$ being an interesting quantity:

Definition 2.17: *For a generator $\mathcal{L}$ and reversible measure μ we define for the associated Dirichlet form, which is sometimes also called the energy:*

$$\mathcal{E}(f,f) = \int f(-\mathcal{L}f)\ d\mu = \int \Gamma_1(f,f)\ d\mu,$$

for all $f \in \mathcal{A}$.

The use of these definitions only becomes fully apparent in later sections. Nevertheless, we discuss some examples now.

Example 2.18: Using $\mathbb{R}$ as state space, we define the Ornstein-Uhlenbeck generator to be

$$\mathcal{L}f(x) = f''(x) - xf'(x).$$

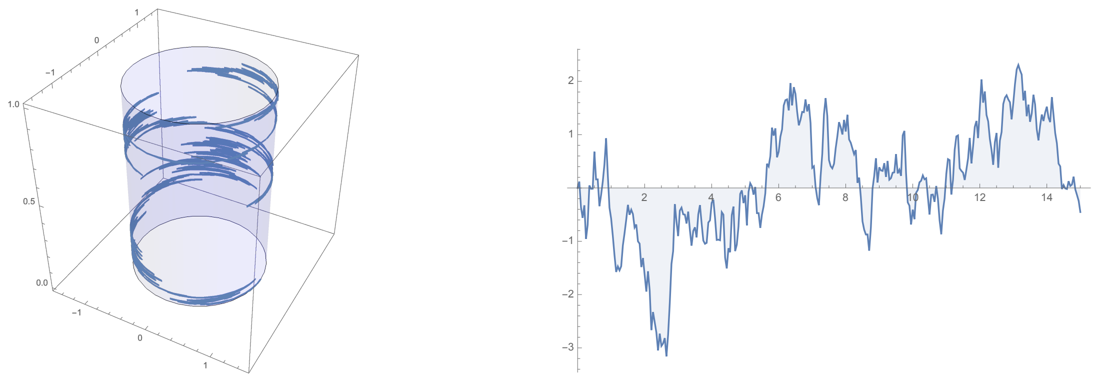
It is defined for $f \in \mathcal{A} = \{f \in C^\infty(\mathbb{R}) : \exists \text{P polynomial, such that: } |f| \leq |P|\}$, the set of smooth functions which grow slower than polynomials. Physically, this process corresponds to a Brownian particle drifting in an harmonic field. For this generator, the standard Gaussian measure γ

$$\gamma(dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx$$

is reversible. Indeed, let $f, g \in \mathcal{A}$. Partial integration and usage of the fact that

$$\lim_{|x|\to\infty} h(x) e^{-\frac{x^2}{2}} = 0$$

for any $h \in \mathcal{A}$, we obtain:



(1) (2)

Figure 2: (1) Brownian motion on the circle; (2) Ornstein-Uhlenbeck process.

$$\begin{aligned}\int_{\mathbb{R}} f(x)\mathcal{L}g(x)\gamma(dx) &= \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g''(x)f(x)e^{-\frac{x^2}{2}}\,dx - \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g'(x)f(x)xe^{-\frac{x^2}{2}}\,dx \\ &= \left[g'(x)f(x)\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right]_{-\infty}^{\infty} - \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g'(x)f(x)xe^{-\frac{x^2}{2}}\,dx \\ &\quad + \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g'(x)f(x)xe^{-\frac{x^2}{2}}\,dx - \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g'(x)f(x)e^{-\frac{x^2}{2}}\,dx \\ &= -\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g'(x)f(x)xe^{-\frac{x^2}{2}}\,dx \\ &= \int_{\mathbb{R}} g(x)\mathcal{L}f(x))\gamma(dx)\end{aligned}$$

Therefore γ is reversible with respect to the Ornstein-Uhlenbeck generator. We also compute Γ_1, let $f, g \in \mathcal{A}$:

$$\begin{aligned}2\Gamma_1(f,g)(x) &= \mathcal{L}(fg) - f(\mathcal{L}g) - g(\mathcal{L}f) \\ &= \partial_x^2(f(x)g(x)) - x\partial_x(f(x)g(x)) \\ &\quad - g''(x)f(x) + xg'(x)f(x) - f''(x)g(x) + xf'(x)g(x) \\ &= 2f'(x)g'(x) - x\big[\partial_x(f(x)g(x)) - g'(x)f(x) - f'(x)g(x)\big] \\ &= 2f'(x)g'(x)\end{aligned}$$

Example 2.19: Using the circle defined as $\mathbb{S}^1 = {}^{[0,1]}/_{(0\sim 1)}$ as state space, we define a (rescaled) Brownian motion by its generator $\mathcal{L}f = f''$ for all $f \in \mathcal{A} = C^\infty(\mathbb{S}^1)$. We keep in mind that there is a bijection between $C^\infty(\mathbb{S}^1)$ and the periodic functions contained in $C^\infty(\mathbb{R})$:

$$C^\infty_{per}(\mathbb{R}) = \left\{f \in C^\infty(\mathbb{R}) : f^{(k)}(x) = f^{(k)}(x+n)\ \forall n \in \mathbb{Z},\ \forall k \in \mathbb{N}_0\right\} \tag{2.6}$$

We use this bijection implicitly and we use it to define f'' on the circle.

Here the reversible measure is the Lebesgue measure. To check this, let $f, g \in \mathcal{A}$:

$$\begin{aligned}\int_{\mathbb{S}^1} f(\mathcal{L}g)\ dx &= \int_0^1 f(x)g''(x)\ dx \\ &= \big[f(x)g'(x)\big]_0^1 - \big[f'(x)g(x)\big]_0^1 + \int_0^1 f''(x)g(x)\ dx = \int_{\mathbb{S}^1}(\mathcal{L}f)g\ dx\end{aligned}$$

where the boundary terms of the partial integration disappear precisely because we work on the circle. And again, we have that $\Gamma_1(f,g) = f'g'$:

$$2\Gamma_1(f,g) = \partial_x^2(f(x)g(x)) - f(x)g''(x) - g(x)f''(x) = 2f'(x)g'(x).$$

To conclude this section, we present different modes of convergence to equilibrium.

Definition 2.20: *A Markov semi-group $(P_t)_{t\geq 0}$ with invariant probability measure μ is called*

- *weakly ergodic, if for all $f \in L^1(\mu)$:*

$$P_t f \to \mu[f] \quad \mu - a.s. \tag{2.7}$$

- *$L^2(\mu)$-ergodic, if for all $f \in L^2(\mu)$:*

$$\int (P_t f - \mu[f])^2 d\mu \to 0. \tag{2.8}$$

- *uniformly ergodic if for all $f \in L^1(\mu)$:*

$$\|P_t f - \mu[f]\|_\infty \to 0. \tag{2.9}$$

These different types of ergodictiy are linked to different functional inequalities which we will explore in the subsequent chapters.

3 Spectral Gap

In this chapter we introduce the spectral gap inequality. We explore some of its properties and in particular its connection to L^2-ergodicity and its stability under perturbation and tensoration. Afterwards we discuss some examples including the one-dimensional XY-model in finite volume.

3.1 General properties

We remind ourselves of Hypothesis 2.15 which assumes the existence of a class of functions $\mathcal{A}$ for which all relevant calculations are well defined.

The first step is to define the inequality:

Definition 3.1: *Let $\mathcal{L}$ be a probability generator with carré du champ Γ_1 and a reversible probability measure μ. μ is said to satisfy a spectral gap inequality with constant $\lambda > 0$ if*

$$Var_\mu(f) \leq \frac{1}{\lambda}\mu\big[\Gamma_1(f,f)\big] \tag{3.1}$$

for any $f \in \mathcal{A}$. $Var_\mu(f) = \mu\big[\big(f - \mu[f]\big)^2\big]$ denotes the variance of f under μ. The optimal constant, i.e. the largest λ that satisfies the above inequality for all $f \in \mathcal{A}$, is denoted by λ_{gap}.

The main reason why we are interested in this inequality is its equivalence to (exponential) L^2-ergodicity which we introduced earlier in (2.8).

Theorem 3.2: *$\mathcal{L}$ and reversible μ satisfy a spectral gap inequality with constant $\lambda > 0$ if and only if*

$$\big\|P_t f - \mu[f]\big\|_2^2 \leq Var_\mu(f)e^{-2\lambda t} \tag{3.2}$$

for all $t \geq 0$ and $f \in \mathcal{A}$.

Proof. This proof is standard, see for example [15, Property 2.4] or [17, Lemma 2.1.4].
First, assume that (3.1) is satisfied. Assume $\mu[f] = 0$ which implies $\mu[P_t f] = 0$ by invariance of μ. Define $u(t) = \mathrm{Var}_\mu(P_t f)$ and observe:

$$\partial_t u(t) = \partial_t \mu\big[(P_t f)^2\big] = 2\mu\big[(P_t f)(\mathcal{L}P_t f)\big] = -2\mu\big[\Gamma_1(P_t f, P_t f)\big]$$

where we used that P_t and $\mathcal{L}$ commute. Now, using the spectral gap inequality we obtain that $u(t)$ satisfies the differential inequality

$$\partial_t u(t) \leq -2\lambda u(t).$$

And by Grönwall's lemma: $u(t) \leq u(0)e^{-2\lambda t}$ for all $t \geq 0$. Using $u(0) = \mathrm{Var}_\mu(f)$ yields (3.2).

For the converse statement, assume (3.2) to be satisfied. Assume again $\mu[f] = 0$. (3.2) then reads as

$$\mu\big[(P_t f)^2\big] \leq e^{-2\lambda t}\mu\big[f^2\big] \implies \frac{\mu\big[(P_t f)^2\big] - \mu\big[f^2\big]}{t} \leq \frac{e^{-2\lambda t}\mu\big[f^2\big]\big] - \mu\big[f^2\big]}{t}$$

for all $t > 0$. Taking the limit $t \to 0$ yields:

$$\frac{d}{dt}\mu\big[(P_t f)^2\big]\bigg|_{t=0} \leq \frac{d}{dt}e^{-2\lambda t}\mu\big[f^2\big]\bigg|_{t=0} \implies 2\mu\big[f(\mathcal{L}f)\big] \leq -2\lambda\mu\big[f^2\big]$$

Diving this by -2λ yields the desired spectral gap inequality. □

Before we discuss examples in the next section, we want to present two useful stability properties of the spectral gap inequality: it is stable under tensoration (i.e. $\mu_1 \otimes \mu_2$) and perturbation.

Theorem 3.3 (Perturbation Property)**:** *Assume that ν satisfies the spectral gap inequality with respect to $\mathcal{L}$ and with constant $\lambda > 0$ on the state space E. Let H be a bounded measurable function and define a new probability measure μ^H:*

$$\mu^H(dx) = \frac{1}{Z_H} e^{-H(x)} \nu(dx), \tag{3.3}$$

where $Z_H = \int e^{-H} d\nu$. Then μ^H satisfies the spectral gap inequality with respect to the same carré du champ Γ_1 and with constant λ_H where $\lambda_H = \lambda e^{-2osc(H)}$ for $osc(H) = \sup_x H(x) - \inf_x H(x)$.

Theorem 3.4 (Tensoration Property)**:** *Assume that μ_1 and μ_2 both satisfy the spectral gap inequality with constant $\lambda > 0$ - respectively with respect to $\mathcal{L}_1$ and $\mathcal{L}_2$ on the state spaces E_1 and E_2. Then the product measure $\mu_1 \otimes \mu_2$ satisfies the spectral gap with respect to the same constant λ and the generator $\mathcal{L}_1 \oplus \mathbb{1} + \mathbb{1} \oplus \mathcal{L}_2$.*

Before we start with the proof of the perturbation property, let us state a small lemma:

Lemma 3.5: *For μ^H defined as in* (3.3) *and any $f \geq 0$ measurable we have*

$$e^{-osc(H)} \nu[f] \leq \mu^H[f] \leq e^{osc(H)} \nu[f].$$

Proof of lemma. By the definition of μ^H, we have $Z_H = \nu[e^{-H}]$. And therefore

$$\mu[f] = \frac{\int f e^{-H} d\nu}{\int e^{-H} d\nu} \leq \frac{\left(\sup_x e^{-H(x)}\right) \int f d\nu}{\left(\inf_x e^{-H(x)}\right) \int 1 d\nu} = e^{osc(H)} \nu[f].$$

The lower bound is analogous. □

Proof of the perturbation property. The proof is standard, see for example [1, Thm. 3.4.1] or [15, Property 2.6]. It uses the following characterisation of the variance:

$$\mathrm{Var}_{\mu^H}(f) = \inf_{a \in \mathbb{R}^2} \mu^H\left[(f-a)^2\right], \quad f \in \mathcal{A}$$

Using this, the spectral gap inequality for ν and the previous lemma twice, we obtain for $f \in \mathcal{A}$:

$$\mathrm{Var}_{\mu^H} \leq \mu^H\left[(f - \nu[f])^2\right] \leq e^{osc(H)} \mathrm{Var}_\nu(f) \leq \frac{e^{osc(H)}}{\lambda} \nu\left[\Gamma_1(f,f)\right] \leq \frac{e^{2osc(H)}}{\lambda} \nu\left[\Gamma_1(f,f)\right]$$

This is the desired spectral gap inequality. □

Proof of the tensoration property. Again, this result is standard, see [1, Thm 3.2.1] or [15, Thm. 2.5]. We combine the two references.

Fix $f\colon E_1 \times E_2 \to \mathbb{R}$ measurable such that $\mu_1 \otimes \mu_2\left[f(-\mathcal{L}_i)f\right]$ is well defined and finite for $i = 1, 2$. We observe:

$$\begin{aligned}
\mathrm{Var}_{\mu_1 \otimes \mu_2}(f) &= \left(\mu_1 \otimes \mu_2\left[f^2\right] - \mu_1\left[\mu_2[f]^2\right]\right) - \left(\mu_1 \otimes \mu_2[f]^2 - \mu_1\left[\mu_2[f]^2\right]\right) \\
&= \mu_1\left[\mathrm{Var}_{\mu_2}(f)\right] + \mathrm{Var}_{\mu_1}\left(\mu_2[f]\right)
\end{aligned}$$

And by the convexity of $\mathrm{Var}_{\mu_1}(\cdot)$ and Jensen's inequality we have:

$$\mathrm{Var}_{\mu_1 \otimes \mu_2}(f) \leq \mu_1 \otimes \mu_2\left[\mathrm{Var}_{\mu_1}(f) + \mathrm{Var}_{\mu_2}(f)\right]$$

Here we can apply the spectral gap inequalities for μ_1, μ_2:

$$\begin{aligned}
\mathrm{Var}_{\mu_1 \otimes \mu_2}(f) &\leq \frac{1}{\lambda} \mu_1 \otimes \mu_2 \Big[\mu_1[f(-\mathcal{L}_1)f] + \mu_2[f(-\mathcal{L}_2)f]\Big] \\
&= \frac{1}{\lambda} \mu_1 \otimes \mu_2 \Big[f(-\mathcal{L}_1)f + f(-\mathcal{L}_2)f\Big]
\end{aligned}$$

This is the desired spectral gap inequality. □

3.2 Examples

We now turn to some examples which satisfy the spectral gap inequality. These extend the discussions found in Section 2.2.

Brownian motion on the circle

We continue Example 2.19. Recall that we considered $\mathcal{L} = \frac{d^2}{dx^2}$ defined on $C^\infty([0,1]/(0 \sim 1))$ and we computed $\Gamma_1(f,f)(x) = (f'(x))^2$. Using this, we can show the spectral gap inequality:

Proposition 3.6: *The Lebesgue measure μ on $[0,1]$ satisfies the spectral gap inequality as follows:*

$$Var_\mu(f) \leq \frac{1}{2\pi} \int \left(f'(x)\right)^2 \, dx \tag{3.4}$$

where $f \in C^\infty([0,1]/(0 \sim 1))$. This inequality is sharp, hence $\lambda_{gap} = 2\pi$.

Proof. This inequality can most easily be shown by Fourier expansion. To this end, let

$$\forall k \in \mathbb{Z}: \ \phi_k(x) = e^{2\pi i k x}, \quad \phi_k'(x) = 2\pi i k \phi_k(x)$$

We also need to extend the inner product of $L^2(\mu)$ to complex functions by

$$\langle f, g \rangle_\mu = \int_0^1 \overline{f(x)} g(x) dx.$$

Then we have $\langle \phi_k, \phi_l \rangle_\mu = \delta_{kl}$. We expand f in terms of the orthonormal basis $(\phi_k)_{k\in\mathbb{Z}}$:

$$f(x) = \sum_{k\in\mathbb{Z}} a_k \phi_k(x) \quad \text{and} \quad f'(x) = \sum_{k\in\mathbb{Z}} 2\pi i k a_k \phi_k(x); \quad \forall k \in \mathbb{Z} : a_k \in \mathbb{C}$$

Again, assume that $\int_0^1 f(x)dx = 0$ which translates to $a_0 = 0$. This yields the estimate:

$$\begin{aligned} \mathrm{Var}_\mu(f) = \langle f, f \rangle_\mu &= \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} \overline{a_k} a_l \langle \phi_k, \phi_l \rangle_\mu = \sum_{k\in\mathbb{Z}} |a_k|^2 \overset{a_0=0}{\leq} \sum_{k\in\mathbb{Z}} k^2 |a_k|^2 \\ &= \frac{1}{2\pi^2} \frac{1}{2} \sum_{k\in\mathbb{Z}} 4\pi^2 k^2 |a_k|^2 = \frac{1}{2\pi^2} \sum_{k\in\mathbb{Z}} \sum_{l\in\mathbb{Z}} 4\pi^2 k^2 \overline{a_k} a_l \langle \phi_k, \phi_l \rangle_\mu = \frac{1}{2\pi^2} \frac{1}{2} \int_0^1 (f'(x))^2 dx \end{aligned}$$

Thus we have shown a spectral gap inequality with constant $\lambda = 2\pi^2$. This estimate is sharp, take $f(x) = \sin(2\pi x)$:

$$\mathrm{Var}_\mu(f) = \int_0^1 \sin^2(2\pi x) dx = \frac{1}{2} = \frac{1}{2\pi^2} \frac{1}{2} \int_0^1 4\pi^2 \cos^2(2\pi x) dx = \frac{1}{2\pi^2} \frac{1}{2} (f'(x))^2 dx$$

□

Gaussian measures

We continue Example 2.18. Recall that we considered $\mathcal{L}f(x) = f''(x) - xf'(x)$ defined on $\mathcal{A}$, the smooth real functions which grow slower than polynomials. The associated Ornstein-Uhlenbeck process admits the standard Gaussian measure γ as reversible measure and we computed $\Gamma_1(f,f)(x) = \left(f'(x)\right)^2$. Using this, we can show a spectral gap inequality:

Proposition 3.7: *Let γ be the standard Gaussian measure on $\mathbb{R}$, it satisfies the following spectral gap inequality:*

$$Var_\gamma(f) \leq \int_{\mathbb{R}} \left(f'(x)\right)^2 \gamma(dx) \tag{3.5}$$

for any $f \in \mathcal{A}$. The inequality is sharp if and only if $f(x) = ax + b$ for some $a, b \in \mathbb{R}$.

Proof. This result is due to [8] but we adapt the exposition of [4].

Similar to the proof of Proposition 3.6, we need to choose the correct basis for $L^2(\gamma)$. Here the choice are the Hermitian polynomials, $(H_k(x))_{k\geq 0}$. We define them via their generating series

$$G_s(x) = e^{sx-\frac{s^2}{2}} = \sum_{k=0}^{\infty} \frac{s^k}{\sqrt{k!}} H_k(x). \tag{3.6}$$

Hence, $H_k(x) = \frac{1}{\sqrt{k!}} \frac{d^k}{ds^k} G_s(x)\big|_{s=0}$. First, we use this to check that $(H_k)_k$ is indeed an orthonormal sequence in $L^2(\gamma)$:

$$\begin{aligned}
\int_{\mathbb{R}} H_k(x)H_\ell(x)\gamma(dx) &= \frac{1}{\sqrt{k!\ell!}} \frac{d^k}{ds^k} \frac{d^\ell}{dt^\ell} e^{-\frac{s^2}{2}} e^{-\frac{t^2}{2}} \int_{\mathbb{R}} e^{sx} e^{tx} \gamma(dx) \bigg|_{s=t=0} \\
&= \frac{1}{\sqrt{k!\ell!}} \frac{d^k}{ds^k} \frac{d^\ell}{dt^\ell} e^{-\frac{s^2}{2}} e^{-\frac{t^2}{2}} e^{\frac{(s+t)^2}{2}} \bigg|_{s=t=0} \\
&= \frac{1}{\sqrt{k!\ell!}} \frac{d^k}{ds^k} \frac{d^\ell}{dt^\ell} e^{st} \bigg|_{s=t=0} \\
&= \sum_{n=0}^{\infty} \frac{1}{\sqrt{k!\ell!}} \frac{1}{n!} \frac{d^k}{ds^k} \frac{d^\ell}{dt^\ell} (st)^n \bigg|_{s=t=0} \\
&= \delta_{k,\ell}
\end{aligned}$$

Secondly, we can use this expansion to check that $(H_k)_k$ diagonalises $\mathcal{L}$. For this we introduce an additional parameter: $G_{s,\theta}(x) = G_{\theta s}(x)$, hence $G_s(x) = G_{s,1}(x)$. Next, we notice that $\mathcal{L}G_{s,\theta}$ is still well-defined for any s, θ. Observe:

$$\mathcal{L}G_s(x) = (s^2 - sx)e^{sx-\frac{s^2}{2}} = -\frac{d}{d\theta} G_{s,\theta}(x)\bigg|_{\theta=1}$$

And by applying (3.6) to the equation above

$$\sum_{k=0}^{\infty} \frac{s^k}{\sqrt{k!}} \mathcal{L}H_k(x) = -\sum_{k=0}^{\infty} \frac{d}{d\theta} \frac{\theta^k s^k}{\sqrt{k!}} H_k(x)\bigg|_{\theta=1} = -\sum_{k=0}^{\infty} \frac{s^k}{\sqrt{k!}} kH_k(x).$$

Thus, by comparing coefficients, we obtain $\mathcal{L}H_k = -kH_k$ for all $k \geq 0$.

The desired spectral gap inequality follows immediately from the properties above. Let $f \in \mathcal{A}$ with $f = \sum_{k=0}^{\infty} a_k H_k$ for some sequence $a_k \in \mathbb{R}$ for all $k \geq 0$. We then have:

$$\mathrm{Var}_\gamma(f) = \sum_{k=1}^{\infty} a_k^2 \leq \sum_{k=1}^{\infty} k a_k^2 = \sum_{k,\ell=0}^{\infty} \langle H_\ell, -\mathcal{L}H_k \rangle_{L^2(\gamma)} = \int_{\mathbb{R}} f(-\mathcal{L}f) d\gamma = \int_{\mathbb{R}} \big(f'(x)\big)^2 \gamma(dx)$$

Furthermore, the above inequality is sharp if and only if $f(x) = aH_1(x) + bH_0(x) = ax + b$ for some $a, b \in \mathbb{R}$. □

Using this, we can show that any Gaussian measure satisfies a similar inequality:

Corollary 3.8: *The standard Gaussian measure on $\mathbb{R}^d$, $\gamma^{\otimes d}$, satisfies a spectral gap inequality as follows:*

$$Var_{\gamma^{\otimes d}}(f) \leq \int_{\mathbb{R}^d} |\nabla f|^2 \gamma^{\otimes d}(dx)$$

for any $f \in \mathcal{A}$.[2] *Furthermore, the inequality is sharp.*

Proof. This follows directly from the previous proposition and the tensoration property, Theorem 3.4. □

[2]Here we have $\mathcal{A} = \mathcal{A}_0^{\otimes d}$ where $\mathcal{A}_0$ is the standard algebra associated to the one-dimensional case.

3.3 One dimensional XY-model

In this section we turn to the spectral gap of the $XY-$model in the one-dimensional case. The one-dimensional structure allows us to improve Theorem 3.4.

Recall the definition of the XY-model from Definition 2.3. In this section we perceive the circle as $[0,1]/0 \sim 1$.

Further we restrict ourselves to a one-dimensional lattice so that we can consider $\Lambda = \{1, ..., L\}$ without loss of generality. The measure of interest then is

$$\mu_\Lambda^\eta(d\omega) = \frac{1}{Z_\Lambda^\eta} e^{-H_\Lambda^\eta(\omega)} \nu_\Lambda(d\omega),$$

where ν_Λ is the normalised Lebesgue measure on $[0,1]^\Lambda$.

First observe that the tensoration property, Theorem 3.4, and Proposition 3.6 imply that our reference measure ν_Λ satisfies a spectral gap inequality - uniformly in Λ and with constant 2π. Theorem 3.3 actually already tells us that μ_Λ^η satisfies the spectral gap inequality with constant bounded by $\inf_\eta 2\pi e^{-2osc(H_\Lambda^\eta)} = 2\pi e^{-4\beta(L+1)}$ uniformly in η. This constant decays exponentially fast in $L = |\Lambda|$ which can be improved:

Proposition 3.9: *For the setting described as above, the measure μ_Λ^η with $\Lambda = \{1, ..., L\}$ satisfies the spectral gap inequality with $\lambda \geq \frac{1}{L} 4\pi e^{-24\beta}$. This means that the family of measures $\{\mu_{\{1,...,L\}}^\eta\}_{\eta \in \Omega}$ satisfies the spectral gap inequality uniformly.*

Proof. This is a special case of [15, Property 2.7] and we adapt the proof accordingly.

The idea of the proof is to exchange one spin at the time with a uniform spin to apply the spectral gap inequality of the reference ν.

Fix η and fix $f \in L^2(\mu_\Lambda^\eta)$, observe:

$$\mathrm{Var}_{\mu_\Lambda^\eta}(f) = \mu_\Lambda^\eta\left[f^2\right] - \mu_\Lambda^\eta\left[f\right]^2 = \frac{1}{2} \iint \left(f(\omega) - f(\tilde{\omega})\right)^2 \mu_\Lambda^\eta(\omega) \mu_\Lambda^\eta(\tilde{\omega})$$

Let us write $f(\omega) - f(\tilde{\omega})$ as a telescoping sum, exchanging one variable at the time:

$$f(\omega) - f(\tilde{\omega}) = \sum_{k=1}^{L} \underbrace{f(..., \omega_k, \tilde{\omega}_{k+1}, ...) - f(..., \omega_{k-1}, \tilde{\omega}_k, ...)}_{:= \Delta_k f(\omega, \tilde{\omega})} = \sum_{k=1}^{L} \Delta_k f(\omega, \tilde{\omega})$$

And by use of the representation of the variance above:

$$\mathrm{Var}_{\mu_\Lambda^\eta}(f) = \frac{1}{2} \left(\mu_\Lambda^\eta \otimes \mu_\Lambda^\eta\right) \left[\left(\sum_{k=1}^{L} \Delta_k f \right)^2 \right] \leq \frac{L}{2} \left(\mu_\Lambda^\eta \otimes \mu_\Lambda^\eta\right) \left[\sum_{k=1}^{L} \left(\Delta_k f\right)^2 \right] \tag{3.7}$$

The inequality comes from the elementary inequality $\left(\sum_{k=1}^n a_k\right)^2 \leq n \sum_{k=1}^n a_k^2$ which is applied to $a_k = \Delta_k f(\omega, \tilde{\omega})$ for fixed $\omega, \tilde{\omega}$.

Now we need to bound $\left(\mu_\Lambda^\eta \otimes \mu_\Lambda^\eta\right)\left[\left(\Delta_k f\right)^2\right]$. We do this by replacing the spin at site k with a uniform spin, i.e. it is distributed according to our reference measure ν. Formally, we fix $k \in \{1, ..., L\}$ and construct a new measure. First, we decompose the Hamiltonian:

$$H_k^{\text{left}}(\omega) = -\beta \sum_{\substack{0 \geq i,j < k \\ |i-j|=1}} \cos(2\pi(\omega_i - \omega_j)) \qquad H_k^{\text{right}}(\omega) = -\beta \sum_{\substack{k < i,j \leq L \\ |i-j|=1}} \cos(2\pi(\omega_i - \omega_j))$$

Observe that $\left\|H^\eta - (H_k^{\text{left}} + H_k^{\text{right}})\right\|_\infty \leq 4\beta$ because we leave out (up to) 4 interaction terms: two of the k-th spin and the two boundary terms. Now we define two measure on $\{1, ..., k-1\}$ and $\{k+1, ..., L\}$:

$$\mu_k^{\text{left}}(d\omega) = \frac{1}{Z^{\text{left}}} e^{-H_k^{\text{left}}(\omega)} d\nu_{\{1,...,k-1\}}(d\omega) \qquad \mu_k^{\text{right}}(d\omega) = \frac{1}{Z^{\text{right}}} e^{-H_k^{\text{right}}(\omega)} d\nu_{\{k+1,...,L\}}(d\omega)$$

where $Z^{\text{left}}, Z^{\text{right}}$ are normalising constants. By applying Lemma 3.5 twice, we have for any measurable $g \geq 0$:

$$e^{-16\beta}\big(\mu_k^{\text{left}} \otimes \nu \otimes \mu_k^{\text{right}}\big)^{\otimes 2}\big[g\big] \leq \big(\mu_\Lambda^\eta \otimes \mu_\Lambda^\eta\big)\big[g\big] \leq e^{16\beta}\big(\mu_k^{\text{left}} \otimes \nu \otimes \mu_k^{\text{right}}\big)^{\otimes 2}\big[g\big] \tag{3.8}$$

Before we apply this to $\Delta_k f$, notice that $\nu^{\otimes 2}\big[\Delta_k\big] = 0$ where $\nu^{\otimes 2}$ acts on the k-th variable of ω and $\tilde{\omega}$ respectively. The spectral gap inequality for ν thus yields:

$$\nu^{\otimes 2}\big[\big(\Delta_k f\big)^2\big] \leq \frac{1}{2\pi}\nu\big[\big(\tfrac{d}{d\omega_k} f\big)^2\big]$$

Combining this with (3.8) we get:

$$\begin{aligned}
\big(\mu_\Lambda^\eta \otimes \mu_\Lambda^\eta\big)\big[\big(\Delta_k f\big)^2\big] &\leq e^{16\beta}\big(\mu_k^{\text{left}} \otimes \nu \otimes \mu_k^{\text{right}}\big)^{\otimes 2}\big[\big(\Delta_k f\big)^2\big] \\
&\leq \frac{e^{16\beta}}{2\pi}\big(\mu_k^{\text{left}} \otimes \nu \otimes \mu_k^{\text{right}}\big)\big[\big(\tfrac{d}{d\omega_k} f\big)^2\big] \\
&\leq \frac{e^{24\beta}}{2\pi}\mu_\Lambda^\eta\big[\big(\tfrac{d}{d\omega_k} f\big)^2\big]
\end{aligned}$$

We conclude by combining this estimate and (3.7):

$$\mathrm{Var}_{\mu_\Lambda^\eta}(f) \leq L\frac{e^{24\beta}}{4\pi}\mu_\Lambda^\eta\big[|\nabla f|^2\big].$$

□

4 Logarithmic Sobolev inequality

Having introduced the spectral gap inequality in the previous chapter, we now turn to the so-called logarithmic Sobolev inequality. It is an improvement over the spectral gap inequality: it still satisfies the tensoration property, but instead of L^2-ergodicity, it is connected to uniform ergodicity. Further, it implies the spectral gap inequality. The only downside is that it is more technical and much harder to prove in explicit examples - even in the Gaussian case.

In this chapter we introduce the logarithmic Sobolev inequality and prove its aforementioned properties before we discuss some examples. We also discuss the Bakry-Emery criterion which can be conveniently used to show the logarithmic Sobolev inequality.

4.1 General properties

We remind ourselves of Hypothesis 2.15 which assumes the existence of a class of functions $\mathcal{A}$ for which all relevant calculations are well defined.

Before we can define the logarithmic Sobolev inequality, we need to define the entropy of a function. Physically, entropy is used as a measure for disorder. In the following, let $\mathcal{L}$ be a Markov generator with carré du champ Γ_1 and μ a reversible probability measure.

Definition 4.1: *For measurable $f \geq 0$ define its entropy with respect to μ:*

$$Ent_\mu(f) = \mu[f \log f] - \mu[f] \log \mu[f],$$

provided the integrals exist. We also define the (relative) entropy of another probability measure ν with respect to μ by:

$$Ent(\nu|\mu) = \begin{cases} Ent_\mu\left(\frac{d\nu}{d\mu}\right) & \text{if} \quad \nu \ll \mu \\ \infty & \text{else} \end{cases}$$

Remark that we make use of the convention $0 \log 0 = 0$. Here are some very basic facts about the entropy:

Lemma 4.2: *The following hold true of Ent_μ:*

1. *Positivity: $Ent_\mu(f) \geq 0$ with equality if and only if f is constant.*
2. *Homogeneity: $Ent_\mu(\gamma f) = \gamma Ent_\mu(f)$ for $\gamma \geq 0$.*
3. *$Ent(\nu|\mu) = \mu\left[\frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu}\right] = \nu\left[\log \frac{d\nu}{d\mu}\right]$ if $\nu \ll \mu$.*

Proof. 2 and 3 follow directly from the definition of Ent_μ. For 1, let $\phi(x) = x \log x$. $\phi''(x) = \frac{1}{x} > 0$, hence ϕ is strictly convex and Jensen's inequality reads

$$\mu[f \log f] = \mu[\phi(f)] \geq \phi(\mu[f]) = \mu[f] \log[f]$$

with equality if and only if f is constant. Rearranging yields the desired statement. □

Having defined the entropy, we can introduce the logarithmic Sobolev inequality.

Definition 4.3: *μ satisfies a logarithmic Sobolev inequality (abbreviated LSI) with constant $0 < \alpha < \infty$ with respect to $\mathcal{L}$ if*

$$Ent_\mu[f^2] \leq 2\alpha\mu[\Gamma_1(f, f)] \tag{4.1}$$

for any $f \in \mathcal{A}$. The optimal constant α_{LS}, i.e. the smallest α which satisfies the above inequality for all $f \in \mathcal{A}$, is called the logarithmic Sobolev constant.

It is not obvious why there should be any measure satisfying such an inequality. Unfortunately, even the simple examples - Markov chains on $\{0,1\}$ and the standard Gaussian measure - have rather complicated proofs. Therefore we adjourn the examples to the next section and discuss some general properties of logarithmic Sobolev inequalities instead.

An important property of LSI is its stability under tensoration and perturbation. The proofs are very similar to the corresponding proofs for the spectral gap inequality, Theorems 3.3 and 3.4, albeit more technical.

Theorem 4.4 (Perturbation property)**:** *Assume that ν satisfies a logarithmic Sobolev inequality with respect to $\mathcal{L}$ and with constant $\alpha < \infty$. Let H be a bounded measurable function and define a new probability measure μ^H:*

$$\mu^H(dx) = \frac{1}{Z_H} e^{-H(x)} \nu(dx),$$

where $Z_H = \int e^{-H} d\nu$. Then μ^H satisfies a logarithmic Sobolev inequality with respect to the same carré du champ Γ_1 and constant bounded by $\alpha e^{2osc(H)}$. Recall $osc(H) = \sup_x H(x) - \inf_x H(x)$.

Theorem 4.5 (Tensoration property)**:** *Assume that μ_1 and μ_2 satisfy a logarithmic Sobolev inequality with constant $\alpha < \infty$ - with respect to two generators $\mathcal{L}_1$ and $\mathcal{L}_2$ on the state spaces E_1 and E_2 respectively. Then the product measure $\mu_1 \otimes \mu_2$ satisfies a logarithmic Sobolev inequality with the same constant α and with respect to $\mathcal{L}_1 \oplus \mathbb{1} + \mathbb{1} \oplus \mathcal{L}_2$.*

The proofs of these two theorems are based on the following variational formulas for the entropy:

Lemma 4.6: *(Variational formulas for the entropy) For a probability measure μ and $f \geq 0$ such that $Ent_\mu(f)$ is well defined, we have the following variational formulas:*

$$\begin{cases} Ent_\mu(f) = \sup\left\{\mu[fg] : g \text{ measurable with } \mu[e^g] = 1\right\} \\ Ent_\mu(f) = \inf\left\{\mu\left[f \log \frac{f}{t} - f + t\right] : t > 0\right\} \end{cases}$$

Proof of the lemma. This formulas are mentioned in [1, Chapter 1.2] and in the proof of [1, Thm. 3.4.3] and we prove them here. Fix f and without loss of generality let $\mu[f] = 1$. For the purpose of the proof denote

$$\begin{cases} \mathrm{Ent}^*_\mu(f) = \sup\left\{\mu[fg] : g \text{ measurable with } \mu[e^g] = 1\right\} \\ \mathrm{Ent}^{**}_\mu(f) = \inf\left\{\mu\left[f \log \frac{f}{t} - f + t\right] : t > 0\right\} \end{cases}$$

Clearly we have $\mathrm{Ent}_\mu(f) \leq \mathrm{Ent}^*_\mu(f)$ by choosing $g = \log f$. For the converse inequality we need the inequality $uv \leq u \log u - u + e^v$ which is valid for all $v \in \mathbb{R}$ and $u \geq 0$. This inequality can be shown by observing that the function $\varphi_u(v) = e^v - u + u \log u - uv$ has a global minimum at $\log u$ and $\varphi_u(\log u) = 0$. With this inequality we get, assume $\mu[e^g] = 1$:

$$\mu[fg] \leq \mu[f \log f] - \mu[f] + \mu[e^g] = \mathrm{Ent}_\mu(f).$$

This completes the proof of $\mathrm{Ent}_\mu(f) = \mathrm{Ent}^*_\mu(f)$.

Regarding $\mathrm{Ent}^{**}_\mu(f)$, we consider the function $\psi_f(t) = \mu\left[f \log \frac{f}{t} - f + t\right]$ and observe that it achieves its global minimum at $t = \mu[f]$. By using $\mu[f] = 1$ we get:

$$\mathrm{Ent}^{**}_\mu(f) = \psi_f(\mu[f]) = \mu[f \log f] = \mathrm{Ent}_\mu(f)$$

This completes the proof of the lemma. □

Proof of Theorem 4.4. This theorem is standard and can be found as [1, Thm. 3.4.3] or [15, Property 4.6]. Fix $f \in \mathcal{A}$ with $f \geq 0$. We make use of the second variational formula of Lemma 4.6

$$\mathrm{Ent}_\mu(f) = \inf_{t>0} \left\{\mu\left[f \log \frac{f}{t} - f + t\right]\right\}.$$

The advantage of this formula is that for any fixed $t > 0$ we have $g_t(x) = x \log \frac{x}{t} - x + t \geq 0$ for any $x \geq 0$. This is true because g_t achieves its global minimum at $x = t$ and $g_t(t) = 0$. This allows us to apply Lemma 3.5 to $\mu^H\big[g_t(f)\big]$, note that it would not be possible to apply the same lemma to $\mu\big[f \log \frac{f}{\mu[f]}\big]$ because $x \log x \not\geq 0$. Using the lemma twice and the logarithmic Sobolev inequality for ν once, we obtain:

$$\begin{aligned}\mathrm{Ent}_{\mu^H}(f) &= \inf_{t>0} \mu^H\big[g_t(f)\big] \leq e^{osc(H)} \inf_{t>0} \nu\big[g_t(f)\big] = e^{osc(H)} \mathrm{Ent}_\nu(f) \\ &\leq 2\alpha e^{osc(H)} \nu\big[\Gamma_1(f^{1/2}, f^{1/2})\big] \leq 2\alpha e^{2osc(H)} \mu^H\big[\Gamma_1(f^{1/2}, f^{1/2})\big].\end{aligned}$$

By replacing f with f^2 we obtain the desired inequality. □

Proof of the tensoration property. Again, this theorem can be found as [1, Thm. 3.2.2] or [15, Thm. 4.4]. We follow the former reference.

Fix $f \colon E_1 \times E_2 \to \mathbb{R}$ measurable, $f \geq 0$ such that $f(-\mathcal{L}_i)f, i = 1, 2$ is well defined. Now let $g \colon E_1 \times E_2 \to \mathbb{R}$ such that $\mu_1 \otimes \mu_2[e^g] = 1$. Define

$$g_1 = g - \log \int e^g d\mu_1; \quad g_2 = \log \int e^g d\mu_1.$$

We then have $g = g_1 + g_2$ and $\mu_1[e^{g_1}] = 1$ as well as $\mu_2[e^{g_2}] = 1$. Using the first formula of Lemma 4.6 we obtain:

$$\mu_1 \otimes \mu_2\big[fg\big] = \mu_2\big[\mu_1[fg_1]\big] + \mu_1\big[\mu_2[fg_2]\big] \leq \mu_1 \otimes \mu_2\Big[\mathrm{Ent}_{\mu_1}(f) + \mathrm{Ent}_{\mu_2}(f)\Big]$$

And by taking the supremum over all admissible g:

$$\mathrm{Ent}_{\mu_1 \otimes \mu_2}(f) \leq \mu_1 \otimes \mu_2\Big[\mathrm{Ent}_{\mu_1}(f) + \mathrm{Ent}_{\mu_2}(f)\Big]$$

We replace f by f^2 and apply the logarithmic Sobolev inequalities for μ_1 and μ_2 to this expression:

$$\mathrm{Ent}_{\mu_1 \otimes \mu_2}(f^2) \leq 2\alpha \cdot \mu_1 \otimes \mu_2\big[f(-\mathcal{L}_1)f + f(-\mathcal{L}_2)f\big],$$

which is the desired inequality. □

To conclude this section we once again comment on the similarity between the spectral gap inequality and the logarithmic Sobolev inequality. We confirm these similarities by proving that any measure satisfying a logarithmic Sobolev inequality also satisfies a spectral gap inequality.

Theorem 4.7: *Assume μ satisfies a LSI with respect to Γ_1 and constant $\alpha > 0$. Then μ satisfies a spectral gap inequality with constant λ and $\lambda > \frac{1}{\alpha}$.*

Proof. We follow the approach of [15, Thm. 4.9].
Fix $f \in \mathcal{A}$ bounded. Without loss of generality, we assume $\mu\big[f\big] = 0$. Because f is not a non-negative function, we cannot apply the LSI straight away. Instead we consider the function

$$g_\varepsilon = 1 + \varepsilon f,$$

with $\varepsilon > 0$ small enough for g_ε to be positive. Notice that $\Gamma_1(g_\varepsilon, g_\varepsilon) = \varepsilon^2 \Gamma_1(f, f)$. We now apply the LSI to g_ε^2 which reads as follows:

$$\mu\big[2(1+\varepsilon f)^2 \log\big(1+\varepsilon f\big)\big] \leq 2\alpha\varepsilon^2 \mu\big[\Gamma_1(f,f)\big] + \mu\big[(1+\varepsilon f)^2\big]\mu\big[2\log\big(1+\varepsilon f\big)\big] \tag{4.2}$$

We expand the inequality above in powers of ε by using

$$\begin{aligned}\mu\big[(1+\varepsilon f)^2 \log\big(1+\varepsilon f\big)\big] &= \mu\big[(1+\varepsilon f)^2(\varepsilon f - \tfrac{1}{2}\varepsilon^2 f^2 + \mathcal{O}(\varepsilon^3))\big] \\ &= \underbrace{\mu\big[\varepsilon f\big]}_{=0} + \tfrac{3}{2}\varepsilon^2 \mu\big[f^2\big] + \mathcal{O}(\varepsilon^3).\end{aligned}$$

And similarly $\mu\big[(1+\varepsilon f)^2\big]\mu\big[\log\big(1+\varepsilon f\big)\big] = \varepsilon^2\mu\big[f^2\big]+\mathcal{O}(\varepsilon^3)$. After dividing by ε^2, (4.2) becomes:

$$3\mu\big[f^2\big] + \mathcal{O}(\varepsilon) \le 2\alpha\Gamma_1(f,f) + \mu\big[f^2\big] + \mathcal{O}(\varepsilon)$$

By letting $\varepsilon \to 0$, we read off $\mu\big[f^2\big] \le \alpha\Gamma_1(f,f)$ which is the desired spectral gap inequality. □

Remark 4.8: We have shown that the logarithmic Sobolev inequality is stronger than the spectral gap inequality. In fact they are not equivalent: Consider the family of measures on $\mathbb{R}$ given by

$$\mu_a(dx) = \frac{1}{Z_a}e^{-|x|^a}dx,$$

where $a > 0$ and $Z_a = \int e^{-|x|^\alpha}dx$. [1, Cor. 6.4.5] shows that μ_a satisfies a logarithmic Sobolev inequality if and only if $a \ge 2$ and a spectral gap inequality if and only if $a \ge 1$. In particular $\{\mu_a, 1 \le a < 2\}$ satisfy a spectral gap inequality but not a logarithmic Sobolev inequality.

The proof of this result is not easy and we will not carry it out. The statement about logarithmic Sobolev inequalities is based on Theorem 5.11 and the statement about the spectral gap inequality uses similar techniques.

4.2 Examples

In this sections we prove some logarithmic Sobolev inequalities. Of special interest is the proof of the inequality for the standard Gaussian measure. Instead of following the historic approach by Gross [14], we present a proof which will be the basis of the curvature criterion introduced in the next section.

Gaussian measures

Without further ado, we present the LSI for the standard Gaussian measure.

Theorem 4.9: *The standard Gaussian measure γ on $\mathbb{R}$ satisfies the logarithmic Sobolev inequality* (4.1)*:*

$$Ent_\gamma(f^2) \le 2\int_{\mathbb{R}}(f'(x))^2\gamma(dx), \quad f \in \mathcal{A}. \tag{4.3}$$

This inequality is sharp, achieved by $f(x) = e^{\lambda x}, \lambda \in \mathbb{R}$ and therefore $\alpha_{LS} = 1$.

Remark 4.10: *This justifies the factor* 2 *in the definition of the LSI* (4.1)*.*

We follow the approach which can be found in [1, Chapter 5.2]. We need one lemma regarding the semi-group of the Ornstein-Uhlenbeck process. This lemma will later be replaced by 4.22.

Lemma 4.11: *We have the following representation of the semi-group $(P_t)_{t\ge0}$ of the Ornstein-Uhlenbeck process:*

$$P_tf(x) = \int_{\mathbb{R}} f\big(e^{-t}x + \sqrt{1-e^{-2t}}y\big)\gamma(dy), \tag{4.4}$$

for $f \in \mathcal{A}$. In particular, this yields

$$\frac{d}{dx}P_tf(x) = e^{-t}P_tf'(x). \tag{4.5}$$

Proof of Lemma 4.11. This lemma can be found in [1, Chapter 2.3].

Because we introduced the Ornstein-Uhlenbeck process via its generator $\mathcal{L} = \frac{d^2}{dx^2} - x\frac{d}{dx}$, we need to check that $\lim_{t\to0}\frac{P_tf-f}{t} = \mathcal{L}f$ for $f \in \mathcal{A}$, hence fix $f \in \mathcal{A}$.

For $z \in \mathbb{R}$, we approximate $f(e^{-t}x + z)$ by its Taylor series:

$$f(e^{-t}x+z) = f(e^{-t}x) + zf'(e^{-t}x) + \frac{1}{2}z^2f''(e^{-t}x) + \mathcal{O}(|z|^3)$$

And hence:

$$\begin{aligned}
&\frac{1}{t}\Big(\int_{\mathbb{R}} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\gamma(dy)-f(x)\Big)\\
=&\frac{1}{t}\Big(\int_{\mathbb{R}} \big(f(e^{-t}x)-f(x)\big)+\sqrt{1-e^{-2t}}f'(e^{-t}x)+\frac{1}{2}(1-e^{-2t})y^2 f''(e^{-t}x)\gamma(dy)\Big)+\mathcal{O}(t^{1/2})
\end{aligned}$$

We use that $\int y\gamma(dy)=0$ and $\int y^2\gamma(dy)=1$ to simplify the above expression

$$\frac{1}{t}\Big(\int_{\mathbb{R}} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\gamma(dy)-f(x)\Big)=\frac{f(e^{-t}x)-f(x)}{t}+\frac{1}{t}\frac{1-e^{-2t}}{2}f''(e^{-t}x)+\mathcal{O}(t^{1/2}).$$

Here we can take the limit $t\to 0$ to obtain

$$\begin{aligned}
\lim_{t\to 0}\frac{1}{t}\Big(\int_{\mathbb{R}} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\gamma(dy)-f(x)\Big)&=\frac{d}{dt}f(e^{-t}x)\Big|_{t=0}+\frac{d}{dt}\frac{1-e^{-2t}}{2}f''(e^{-tx})\Big|_{t=0}\\
&=-xf'(x)+f''(x).
\end{aligned}$$

This proves (4.4). (4.5) follows immediately from this representation

$$\frac{d}{dx}P_tf(x)=\frac{d}{dx}\int_{\mathbb{R}} f\big(e^{-t}x+\sqrt{1-e^{-2t}}y\big)\gamma(dy)=e^{-t}P_tf'(x).$$

□

Proof of Theorem 4.9. This proof is the content of [1, Chapter 5.2]. In the following, let $f\in\mathcal{A}$ with $f\geq 0$. We first need to make the qualitative observation

$$\operatorname{Ent}_\gamma(P_tf)\overset{t\to\infty}{\longrightarrow}0. \tag{4.6}$$

There are two ways to check this: On the one hand, we could just invoke Lemma 4.11. To avoid this lemma which is specific for this process, we appeal to Theorem 3.7 and Theorem 3.2 which tell us that P_tf converges in $L^2(\gamma)$ to $\gamma[f]$, hence also in distribution which implies 4.6.

This allows us to write:

$$\operatorname{Ent}_\gamma(f)=\operatorname{Ent}_\gamma(P_0f)=-\int_0^\infty\partial_s[\operatorname{Ent}_\gamma(P_sf)]ds=-\int_0^\infty\int_{\mathbb{R}}\partial_s[P_sf\log P_sf]d\gamma ds,$$

where we assume $\gamma[f]=\gamma[P_sf]=1$, without loss of generality by the homogeneity of the entropy, Lemma 4.2.

Remembering that $\partial_sP_sf=\mathcal{L}P_sf=P_s\mathcal{L}f$ and $\gamma[g(-\mathcal{L})h]=\gamma[g'h']$, we continue with the computation:

$$\begin{aligned}
-\int_0^\infty\int_{\mathbb{R}}\partial_s[P_sf\log P_sf]d\gamma ds&=-\int_0^\infty\int_{\mathbb{R}}(1+\log P_sf)\mathcal{L}P_sf d\gamma ds\\
&=\int_0^\infty\int_{\mathbb{R}}\partial_x[1+\log P_sf(x)]\cdot\partial_x[P_sf(x)]\gamma(dx)ds\\
&=\int_0^\infty\int_{\mathbb{R}}\frac{(\partial_xP_sf(x))^2}{P_sf(x)}\gamma(dx)ds
\end{aligned}$$

Remark that this quotient is well defined $\gamma-$almost-everywhere. Further, it is nonnegative, as P_s is positivity preserving. We are now at the point to apply Lemma 4.11:

$$\frac{(\partial_xP_sf(x))^2}{P_sf(x)}=e^{-2s}\frac{(P_sf'(x))^2}{P_sf(x)} \tag{4.7}$$

Furthermore, we apply the Cauchy Schwarz inequality:

$$e^{-2s}\frac{(P_sf'(x))^2}{P_sf(x)} = e^{-2s}\frac{1}{P_sf(x)}P_s\left(\frac{f'\sqrt{f}}{\sqrt{f}}\right)^2(x) \le e^{-2s}P_s\left(\frac{f'^2}{f}\right)(x)$$

Combing all of the above, we arrive at the final estimate:

$$\begin{aligned}\mathrm{Ent}_\gamma(f) &\le \int_0^\infty \int_{\mathbb{R}} e^{-2s}P_s\left(\frac{f'^2}{f}\right)d\gamma ds = \int_0^\infty e^{-2s}\gamma P_s\left[\frac{f'^2}{f}\right]ds \\ &= \int_0^\infty e^{-2s}\gamma\left[\frac{f'^2}{f}\right]ds = \frac{1}{2}\gamma\left[\frac{f'^2}{f}\right],\end{aligned}$$

where we used the P_s invariance of γ.

Lastly, we apply the above estimate to f^2 to obtain $\mathrm{Ent}_\gamma(f^2) \le 2\gamma[f'^2]$ which completes the proof. □

It is evident that the only time we actually used some specific information about $(P_t)_{t\ge 0}$ is when we commuted P_s and ∂_x in (4.7). In the next section the curvature criterion allows us to replace precisely this equation.

As a corollary, we can immediately derive that the standard Gaussian measure on $\mathbb{R}^d$ satisfies a logarithmic Sobolev inequality:

Corollary 4.12: *The standard Gaussian measure on $\mathbb{R}^d$, namely $\gamma^{\otimes d}$, satisfies a LSI:*

$$Ent_{\gamma^{\otimes d}}(f^2) \le 2\int_{\mathbb{R}^d} |\nabla f|^2 \, d\gamma^{\otimes d}$$

Proof. This is precisely the product property of logarithmic Sobolev inequalities. □

Brownian motion on the circle and XY-model

Besides the Gaussian measure, another measure satisfying a logarithmic Sobolev inequality is the Lebesgue measure on $[0,1]$ with respect to the generator $\mathcal{L}f = f''$ defined on $\mathcal{A} = C^\infty_{per}([0,1])$, the periodic smooth functions.

Theorem 4.13: *The Lebesgue measure ν on $[0,1]$ satisfies the logarithmic Sobolev inequality*

$$Ent_\nu(f^2) \le \frac{1}{2\pi^2}\int_0^1 \left(f'(x)\right)^2 dx, \quad f \in C^\infty_{per}([0,1]).$$

We prove a more general statement later, namely Proposition 5.10. For this statement with the optimal constant, see [4, Prop. 5.7.5].

Nevertheless, we can already mention a consequence for the XY-model:

Corollary 4.14: *The XY-model the finite volume $\Lambda \Subset \mathbb{Z}^d$ satisfies a logarithmic Sobolev inequality uniformly in the boundary condition:*

$$Ent_{\mu^\eta_\Lambda}(f^2) \le \frac{1}{2\pi^2}e^{8\beta|\Lambda|}\mu^\eta_\Lambda\left[(\nabla f)^2\right],$$

for any differentiable f, any boundary condition η.

Proof. This follows directly from the perturbation and tensoration properties of logarithmic Sobolev inequalities, Theorems 4.4 and 4.5. □

4.3 Decay of the entropy

In this section we assume an additional property of Γ_1: we assume it to satisfy the Leibniz rule. This has several nice consequences, in particular the entropy decays exponentially fast. We also present the Herbst argument which links logarithmic Sobolev inequalities to a concentration of measure phenomenon.

We say that Γ_1 satisfies the Leibniz rule if

$$\Gamma_1(f, gh) = \Gamma_1(f, g)h + \Gamma_1(f, h)g. \tag{4.8}$$

This is for example the case if $\Gamma_1(f, g) = f'g'$ for $f, g \in C_c^2(\mathbb{R})$. We use this to derive the following property of Γ_1:

$$\Gamma_1(f, g^n) = \Gamma_1(f, g^{n-1})g + \Gamma_1(f, g)g^{n-1} = \Gamma_1(f, g)ng^{n-1}$$

by induction. This can be lifted to analytic functions in the following form, [15, Lemma 4.12]:

$$\Gamma_1(f, \phi(g)) = \Gamma_1(f, g)\phi'(g) \tag{4.9}$$

for any analytic ϕ and $f, g \in \mathcal{A}$. We use this multiple times in this and the following section. Under this additional assumption we obtain the exponential decay of the entropy:

Theorem 4.15: *Assume that μ satisfies a logarithmic Sobolev inequality with respect to $\mathcal{L}$ with constant α. Assume additionally that Γ_1 satisfies the Leibniz rule. We then have*

$$Ent_\mu(P_t f) \leq e^{-\frac{2t}{\alpha}} Ent_\mu(f), \quad f \in \mathcal{A}, f \geq 0,$$

for any $t \geq 0$.

Proof. This statement is part of Gross' integration lemma, as formulated in [14]. We discuss the general version of the integration lemma shortly after. The proof can also be found in [1, Thm. 2.6.7][4, Thm. 5.2.1].

Combining (4.9) and the logarithmic Sobolev inequality reads as follows:

$$\mu\left[\frac{\Gamma_1(g, g)}{g}\right] = \mu\left[4\Gamma_1(g^{1/2}, g^{1/2})\right] \geq \frac{2}{\alpha}\mathrm{Ent}_\mu(g), \tag{4.10}$$

for any $g \in \mathcal{A}$ with $g \geq 0$. We use this to derive an inequality for $\frac{d}{dt}\mathrm{Ent}_\mu(P_t f)$, assume without loss of generality $\mu[f] = 1$:

$$\begin{aligned}
\frac{d}{dt}\mathrm{Ent}_\mu(P_t f) &= \int \frac{d}{dt}(P_t f)\log(P_t f)d\mu = \int (1 + \log(P_t f))\mathcal{L}(P_t f)d\mu = \int \log(P_t f)\mathcal{L}(P_t f)d\mu \\
&= -\int \Gamma_1(P_t f, \log P_t f)d\mu = -\int \frac{\Gamma_1(P_t f, P_t f)}{P_t f}d\mu \\
&\leq -\frac{2}{\alpha}\mathrm{Ent}_\mu(P_t f)
\end{aligned}$$

where we used (4.10) and (4.9), applied to $\Gamma_1(g, \log g)$ for $g = P_t f$. Grönwall's lemma now yields:

$$\mathrm{Ent}_\mu(P_t f) \leq e^{-\frac{2t}{\alpha}}\mathrm{Ent}_\mu(P_0 f) = e^{-\frac{2t}{\alpha}}\mathrm{Ent}_\mu(f),$$

which is the desired statement. □

This exponential decay of entropy implies the convergence to equilibrium in a strong sense, namely with respect to the total variation distance. We follow the exposition of [4, p.244]. For two probability measures μ, ν

$$\|\mu - \nu\|_{TV} = \sup_A |\mu(A) - \nu(A)|$$

is called the total variation distance. The supremum ranges over all measurable sets A. Further, Pinskers's inequality links the total variation distance to the entropy

$$\|\mu - \nu\|_{TV}^2 \leq \frac{1}{2}\text{Ent}(\nu|\mu). \tag{4.11}$$

With these statements we immediately obtain the following corollary from the previous theorem:

Corollary 4.16: *Under the assumptions of the previous theorem: Assume we are given $\nu(dx) = \rho(x)dx$ and assume that $Ent(\nu|\mu) < \infty$. For $\nu_t = \nu P_t$ we then have*

$$\|\nu_t - \mu\|_{TV} \leq \frac{1}{\sqrt{2}} e^{-\frac{t}{\alpha}} \sqrt{Ent(\nu|\mu)},$$

for all $t \geq 0$

Proof. This follows from Theorem 4.15 and (4.11) by observing that $\nu_t(dx) = P_t\rho(x)\mu(dx)$:

$$\|\nu_t - \mu\|_{TV}^2 \leq \frac{1}{2}\text{Ent}(\nu|\mu) \leq \frac{1}{2}e^{-\frac{2t}{\alpha}}\text{Ent}(\nu_0|\mu)$$

□

Next we present a version of Gross' integration Lemma [14]. This characterisation is key to proving uniform ergodicity of the dynamical XY-model later. Note that for this statement we do not need assume that Γ_1 satisfies the Leibniz rule.

Proposition 4.17 (Gross' integration Lemma)**:** *A reversible probability measure μ satisfies a logarithmic Sobolev inequality with respect to a semi-group $(P_t)_t$ with constant $\alpha > 0$ if and only if for all $1 < p < \infty$ and all $p \leq q \leq q(t, p, \alpha) = 1 + (p-1)e^{\frac{2t}{\alpha}}$ we have for all $t \geq 0$*

$$\|P_t\|_{p,q} \leq 1.$$

Remark 4.18: *This property is called hypercontractivity: namely above P_t is hypercontractive with contraction function $q(t, p, \alpha)$, [1, Def. 2.7.1].*

Proof. We keep the proof brief by leaving out lengthy computations of derivatives, see [15, Thm. 4.1] for more details and [1, Thm. 2.8.2] for a concise presentation.

Fix α, p and $f \in \mathcal{A}$ with $f \geq 0$, abbreviate $q(t) = q(t, p, \alpha)$. The key to the proof is to consider the derivative of $\Phi(t) = \log \|P_tf\|_{q(t)}$. One can compute

$$\frac{d}{dt}\Phi(t) = \frac{q'(t)}{q^2(t)\mu\big[(P_tf)^{q(t)}\big]}\left[\text{Ent}_\mu\big((P_tf)^{q(t)}\big) + \frac{q^2(t)}{q'(t)}\mu\left[(P_tf)^{q(t)-1}\mathcal{L}P_tf\right]\right]. \tag{4.12}$$

Assume first that $\|P_t\|_{p,q} \leq 1$ holds and choose $p = 2$. We then have

$$e^{\Phi(t)} = \|P_tf\|_{q(t)} \leq \|f\|_2 = \|P_0f\|_2 = e^{\Phi(0)}. \tag{4.13}$$

This entails that $\frac{d}{dt}e^{\Phi(t)}\big|_{t=0} \leq 0$ which in turn implies $\Phi'(0) \leq 0$. We evaluate (4.12) at $t = 0$, $q(0) = 2$ and $q'(0) = \frac{2}{\alpha}$:

$$\frac{2}{4\alpha\mu[f^2]}\left[\text{Ent}_\mu(f^2) + \frac{4\alpha}{2}\mu\big[f\mathcal{L}f\big]\right] = \frac{1}{2\alpha\mu[f^2]}\left[\text{Ent}_\mu(f^2) - 2\alpha\Gamma_1(f,f)\right] \leq 0$$

This yields

$$\text{Ent}_\mu\big(f^2\big) \leq 2\alpha\mu\big[\Gamma_1(f,f)\big],$$

which is the desired logarithmic Sobolev inequality. The other implication holds true as well: one can show, [1, Lemma 2.8.1] that a logarithmic Sobolev inequality implies that

$$\text{Ent}_\mu\big((P_tf)^{q(t)}\big) + \frac{q^2(t)}{q'(t)}\mu\left[(P_tf)^{q(t)-1}\mathcal{L}P_tf\right] \leq 0,$$

which in turn implies (4.13). This is the desired estimate $\|P_t\|_{p,q(t)} \leq 1$. □

To conclude this section, we present one more consequence of the Leibniz rule for Γ_1: it links the logarithmic Sobolev inequality to a concentration of measure inequality.

Theorem 4.19: *Assume that μ satisfies a logarithmic Sobolev inequality with respect to $\mathcal{L}$ with constant α. Assume that Γ_1 satisfies the Leibniz rule. Then for all $f \in \mathcal{A}$ with $\|\Gamma_1(f,f)\|_\infty \leq 1$ we have*[3]

$$\mu\big(\{|f - \mu[f]| \geq r\}\big) \leq 2e^{-\frac{r^2}{2\alpha}},$$

for all $r > 0$.

Proof. This is a combination of [1, Thm. 7.4.1] which discusses the special case of $\mathbb{R}^n$ and [15, Exercise 4.8].

Denote $H(\lambda) = \mu\big[e^{\lambda f}\big]$ the Laplace transform of f. We apply the logarithmic Sobolev inequality to $e^{\lambda f}$:

$$\begin{aligned}\mathrm{Ent}_\mu(e^{\lambda f}) &= \mu\big[\lambda f e^{\lambda f}\big] - \mu\big[e^{\lambda f}\big]\log\mu\big[e^{\lambda f}\big] = \lambda H'(\lambda) - H(\lambda)\log H(\lambda)\\ &\leq 2\alpha\int \Gamma_1(e^{\frac{\lambda}{2}f}, e^{\frac{\lambda}{2}f})d\mu\end{aligned}$$

Here we can use (4.9): $\Gamma_1(e^{\frac{\lambda}{2}f}, e^{\frac{\lambda}{2}f}) = \frac{\lambda^2}{4}\Gamma_1(f,f)e^{\lambda f}$. The above inequality hence becomes

$$\lambda H'(\lambda) - H(\lambda)\log H(\lambda) \leq \frac{\alpha}{2}\lambda^2\int \Gamma_1(f,f)e^{\lambda f}d\mu \leq \frac{\alpha\|\Gamma_1(f,f)\|_\infty}{2}H(\lambda) \leq \frac{\alpha}{2}\lambda^2 H(\lambda),$$

by using our assumption $\|\Gamma_1(f,f)\|_\infty \leq 1$. Rearranging this differential inequality for $H(\lambda)$:

$$\frac{H'(\lambda)}{\lambda H'(\lambda)} - \frac{\log H(\lambda)}{\lambda^2} \leq \frac{\alpha}{2}$$

Defining $K(\lambda) = \frac{\log H(\lambda)}{\lambda}$ for $\lambda > 0$, this is equivalent to

$$K'(\lambda) \leq \frac{\alpha}{2}. \tag{4.14}$$

Further, we observe by l'Hôpital's rule:

$$\lim_{\lambda\to 0} K(\lambda) = \lim_{\lambda\to 0}\frac{\log\mu[e^{\lambda f}]}{\lambda} = \lim_{\lambda\to 0}\frac{\mu[fe^{\lambda f}]}{\mu[e^{\lambda f}]} = \mu[f]$$

This allows us to obtain an estimate for $H(\lambda)$ by using (4.14):

$$K(\lambda) - K(0) = \int_0^\lambda K'(s)ds \leq \lambda\frac{\alpha}{2} \implies H(\lambda) \leq e^{\lambda^2\frac{\alpha}{2} + \lambda\mu[f]} \tag{4.15}$$

The desired statement now follows from an application of Markov's inequality:

$$\mu\big(\{f - \mu[f] \geq r\}\big) \leq \inf_{\lambda > 0} H(\lambda)e^{-\lambda r}e^{-\lambda\mu[f]} \leq \inf_{\lambda>0} e^{-\lambda r + \lambda^2\frac{\alpha}{2}} = e^{-\frac{r^2}{2\alpha}}$$

Using the same inequality for $-f$ yields the desired statement.

□

[3] This requirement is for example satisfied when $\Gamma_1(f,f) = |\nabla f|^2$ on $\mathbb{R}^n$ and $f \in C^1(\mathbb{R}^n)$ is Lipschitz continuous with Lipschitz constant ≤ 1.

4.4 Curvature criterion

In this section we introduce the curvature criterion which implies the logarithmic Sobolev inequality for a large class of measures. It is essentially a generalisation of the proof of Theorem 4.9 where the criterion replaces Lemma 4.11. Throughout this section we need to assume that Γ_1 satisfies the Leibniz rule, just like in the previous section.

This criterion was introduced in [2] by Bakry and Emery and is often called the Bakry-Emery criterion as well.

Before we can introduce the criterion, we need a definition:

Definition 4.20: *For a generator $\mathcal{L}$ with carré du champ Γ_1, we define the iterated carré du champ*

$$\Gamma_2(f,f) = \frac{1}{2}\big[\mathcal{L}\Gamma_1(f,f) - 2\Gamma_1(f,\mathcal{L}f)\big],$$

for $f \in \mathcal{A}$.

We can now introduce the curvature criterion:

Definition 4.21: *$\mathcal{L}$ satisfies the curvature criterion (denoted (CC)) with constant $\alpha > 0$ if*

$$\Gamma_2(f,f) \geq \frac{1}{\alpha}\Gamma_1(f,f), \tag{4.16}$$

for any $f \in \mathcal{A}$.

Here is a characterisation of the curvature criterion:

Proposition 4.22: *Condition (CC) is satisfied if and only if*

$$\Gamma_1(P_tf, P_tf) \leq e^{-\frac{2}{\alpha}t} P_t\Gamma_1(f,f), \tag{4.17}$$

for all $t \geq 0$ and all $f \in \mathcal{A}$.

Proof. This is [15, Thm. 4.15] and we follow their proof. Check [1, Chapter 5.4] for slightly different characterisations with similar proofs.

We first assume that (CC) is satisfied, fix f and consider the function

$$F(s) = e^{\frac{2}{\alpha}s} P_{t-s}\Gamma_1(P_sf, P_sf) \tag{4.18}$$

for $s \in [0,t]$. We compute the derivative of F:

$$\begin{aligned}
\frac{d}{ds}F(s) &= e^{\frac{2}{\alpha}s} P_{t-s}\bigg(\frac{2}{\alpha}\Gamma_1(P_sf,P_sf) - \mathcal{L}\Gamma_1(P_sf,P_sf) + \frac{d}{ds}\Gamma_1(P_sf,P_sf)\bigg) \\
&= e^{\frac{2}{\alpha}s} P_{t-s}\bigg(\frac{2}{\alpha}\Gamma_1(P_sf,P_sf) - \mathcal{L}\Gamma_1(P_sf,P_sf) + 2\Gamma_1(P_sf,\mathcal{L}P_sf\bigg) \\
&= 2e^{\frac{2}{\alpha}s} P_{t-s}\bigg(\underbrace{\frac{1}{\alpha}\Gamma_1(P_sf,P_sf) - \Gamma_2(P_sf,P_sf)}_{\leq 0}\bigg) \leq 0
\end{aligned}$$

by our assumption (4.16). This means that F is a decreasing function, in particular $F(0) \geq F(t)$. This reads as

$$F(t) = e^{\frac{2}{\alpha}t}\Gamma_1(P_tf,P_tf) \leq P_t\Gamma_1(f,f) = F(0),$$

which is (4.17).

For the converse statement, assume (4.17) to be satisfied. This entails that for any $t > 0$ we have:

$$\frac{1}{t}\big[e^{-\frac{2}{\alpha}t}P_t\Gamma_1(f,f) - \Gamma_1(P_tf,P_tf)\big] = \frac{e^{-\frac{2}{\alpha}t}-1}{t}P_t\Gamma_1(f,f) + \frac{1}{t}\big[P_t\Gamma_1(f,f) - \Gamma_1(P_tf,P_tf)\big] \geq 0$$

and taking the limit $t \to 0$ yields:

$$\begin{aligned}
\lim_{t\to 0} \frac{e^{-\frac{2}{\alpha}t}-1}{t} P_t\Gamma_1(f,f) &= -\frac{2}{\alpha}\Gamma_1(f,f) \\
\lim_{t\to 0}\frac{1}{t}\big[P_t\Gamma_1(f,f)-\Gamma_1(P_tf,P_tf)\big] &= \lim_{t\to 0}\frac{1}{t}\frac{1}{2}\Big(P_t\big(\mathcal{L}(f^2)\big)-\mathcal{L}\big((P_tf)^2\big)\Big) \\
&\quad - \lim_{t\to 0}\frac{1}{t}\Big(P_t\big(f\mathcal{L}f\big)-(P_tf)(\mathcal{L}P_tf)\Big) \\
&= \bigg[\frac{d}{dt}\bigg(\frac{P_t\big(\mathcal{L}(f^2)\big)}{2}-\frac{\mathcal{L}\big((P_tf)^2\big)}{2} \\
&\qquad - P_t\big(f\mathcal{L}f\big)+(P_tf)(\mathcal{L}P_tf)\bigg)\bigg]_{t=0} \\
&= \frac{1}{2}\mathcal{L}^2(f^2)-\mathcal{L}\big(f\mathcal{L}f\big)-\mathcal{L}\big(f\mathcal{L}f\big)+\big(\mathcal{L}f\big)^2+f\mathcal{L}^2f \\
&= 2\Gamma_2(f,f)
\end{aligned}$$

In total we get that

$$-\frac{2}{\alpha}\Gamma_1(f,f)+\Gamma_2(f,f)\geq 0,$$

which is the condition (CC). □

The reason why the curvature criterion is useful for us is the following theorem: it implies a logarithmic Sobolev inequality for the reversible measure under some mild assumptions.

Theorem 4.23: *Let $\mathcal{L}$ be the generator of the semi-group $(P_t)_{t\geq 0}$. Assume that the carré du champ Γ_1 satisfies the Leibniz rule and let μ be reversible. Assume that the semi-group is weakly ergodic* (2.7)*. Then the curvature criterion* (4.16) *with constant α implies that μ satisfies a logarithmic Sobolev inequality with the same constant:*

$$Ent_\mu(f^2)\leq 2\alpha\mu\big[\Gamma_1(f,f)\big],$$

for all $f \in \mathcal{A}$.

Proof. We follow the presentation of [15, Thm. 4.16], an alternative presentation can for example be found in [1, Chapter 5]. As mentioned before, the proof is very similar to the proof of Theorem 4.9.

Before we start the proof, we need to make some observations about Γ_1. The first one is general, namely Γ_1 satisfies a Cauchy-Schwarz inequality

$$\Gamma_1(f,g)\leq \Gamma_1(f,f)^{1/2}\Gamma_1(g,g)^{1/2}.$$

This follows from its bilinearity and the fact that $\Gamma_1(f,f)\geq 0$. The second observation is a consequence of the Leibniz rule that we have seen before, recall (4.9):

$$\Gamma_1(f,\phi(g))=\Gamma_1(f,g)\phi'(g),$$

for any analytic ϕ and $f,g\in\mathcal{A}$.

Without loss of generality, let $\mathcal{A}\ni f\geq 0$ with $\mu[f]=1$, we start again with the qualitative observation:

$$\mathrm{Ent}_\mu(P_tf)\xrightarrow{t\to\infty}0,$$

due to our assumption of weak ergodicity and the dominated convergence theorem. This allows us to write

$$\mathrm{Ent}_\mu(f)=-\int_0^\infty\frac{d}{ds}\mathrm{Ent}_\mu(P_sf)ds. \tag{4.19}$$

We compute the derivative:

$$
\begin{aligned}
-\frac{d}{ds}\mathrm{Ent}_\mu(P_s f) &= -\mu\big[\mathcal{L}P_s(f)\log P_s f\big] - \mu\big[\mathcal{L}(P_s f)\big] = -\mu\big[\mathcal{L}P_s(f)\log P_s f\big] \\
&= \mu\big[\Gamma_1(f, P_s(\log P_s f))\big]
\end{aligned}
$$

where we further used that P_s is a symmetric operator on $L^2(\mu)$. We estimate this by use of the Cauchy-Schwarz inequality for Γ_1 and μ respectively:

$$
\begin{aligned}
\mu\big[\Gamma_1(f, P_s(\log P_s f))\big] &\le \mu\left[\frac{\Gamma_1(f,f)^{1/2}}{f^{1/2}} f^{1/2}\Gamma_1(P_s(\log P_s f), P_s(\log P_s f))^{1/2}\right] \\
&\le \mu\left[\frac{\Gamma_1(f,f)}{f}\right]^{1/2} \mu\Big[f\Gamma_1\big(P_s(\log P_s f), P_s(\log P_s f)\big)\Big]^{1/2}
\end{aligned}
$$

We use Proposition 4.22 to estimate the second term:

$$
\begin{aligned}
\mu\big[f\Gamma_1\big(P_s(\log P_s f), P_s(\log P_s f)\big)\big] &\le \mu\big[fP_s\Gamma_1\big(\log P_s f, \log P_s f\big)\big] \\
&= e^{-\frac{2}{\alpha}s}\mu\big[(P_s f)\Gamma_1\big(\log P_s f, \log P_s f\big)\big] \\
&= e^{-\frac{2}{\alpha}s}\mu\big[\Gamma_1\big(P_s f, \log P_s f\big)\big] \\
&= e^{-\frac{2}{\alpha}s}\mu\big[\Gamma_1\big(f, P_s(\log P_s f)\big)\big]
\end{aligned}
$$

where we used the symmetry of P_s twice and (4.9) once. In total we have shown the estimate

$$
\mu\big[\Gamma_1\big(f, P_s(\log P_s f)\big)\big] \le e^{-\frac{1}{\alpha}s}\mu\left[\frac{\Gamma_1(f,f)}{f}\right]^{1/2} \mu\big[\Gamma_1\big(f, P_s(\log P_s f)\big)\big]^{1/2},
$$

and by observing that $f^{-1}\Gamma(f,f) = 4\Gamma_1(f^{1/2}, f^{1/2})$, again by (4.9), we obtain

$$
\mu\big[\Gamma_1\big(f, P_s(\log P_s f)\big)\big] \le 4e^{-\frac{2}{\alpha}s}\Gamma_1(f^{1/2}, f^{1/2}).
$$

We recall that the left-hand side of this inequality was $-\frac{d}{ds}\mathrm{Ent}_\mu(P_s f)$ and thus by plugging this estimate into (4.19) we get:

$$
\mathrm{Ent}_\mu(f) \le \int_0^\infty 4e^{-\frac{2}{\alpha}s}\Gamma_1(f^{1/2}, f^{1/2})ds = 2\alpha\Gamma_1(f^{1/2}, f^{1/2})
$$

This is the desired logarithmic Sobolev inequality. $\square$

This theorem allows us to quickly check logarithmic Sobolev inequalities as the following result shows:

Theorem 4.24: *Let $\Phi \in C^2(\mathbb{R}^n)$ and assume that $\int e^{-\Phi}dx = 1$. Define the generator*

$$
\mathcal{L}f = \Delta f - \nabla\Phi \cdot \nabla f,
$$

for $f \in C_c^2(\mathbb{R}^n)$. Then $\mu(dx) = e^{-\Phi(x)}dx$ is reversible with respect to $\mathcal{L}$. If additionally

$$
Hess(\Phi) \ge \frac{1}{\alpha}\mathbb{1}
$$

as quadratic form for some $\alpha > 0$ where $Hess(\Phi)$ is the Hessian, then μ satisfies a logarithmic Sobolev inequality with constant α, i.e. for any $f \in C_c^2(\mathbb{R}^n)$:

$$
Ent_\mu(f^2) \le 2\alpha\mu\big[|\nabla f|^2\big]
$$

Before we start with the proof let us state a corollary which yields a logarithmic Sobolev inequality for a wide class of measures on $\mathbb{R}^n$.

Corollary 4.25: *This theorem still holds when Φ is replaced by $\Phi + U$ where U is a bounded function. The constant of the inequality changes to $\alpha e^{2osc(U)}$.*

Proof. This follows from the theorem and the perturbation property of logarithmic Sobolev inequalities, Theorem 4.4. □

Proof of the Theorem. This theorem again goes back to [2] and is one of the most prominent application of the curvature criterion. It can also be found in [1, Cor. 5.5.2] and [15, Exercise 4.18]. Its proof consists of checking reversibility of μ and checking Theorem 4.23 for μ to satisfy the desired inequality.

We start by checking the reversibility of μ, let $f, g \in C_c^2(\mathbb{R}^n)$ and by partial integration:

$$\begin{aligned}\int f(\mathcal{L}g)d\mu &= \sum_{i=1}^n \int \left(f \frac{d^2 g}{dx_i^2} e^{-\Phi} - f \frac{dg}{dx_i} \frac{d\Phi}{dx_i} e^{-\Phi} \right) dx \\ &= \sum_{i=1}^n \int \left(- \frac{df}{dx_i} \frac{dg}{dx_i} e^{-\Phi} + \frac{df}{dx_i} g \frac{d\Phi}{dx_i} e^{-\Phi} - \frac{df}{dx_i} g \frac{d\Phi}{dx_i} e^{-\Phi} \right) dx \\ &= -\int \nabla f \cdot \nabla g d\mu = \int (\mathcal{L}f) g d\mu\end{aligned}$$

Hence μ is reversible. A computation similar to Example 2.18 and the above shows that $\Gamma_1(f, g) = \nabla f \cdot \nabla g$.

Next, we want to compute $\Gamma_2(f, f)$:

$$\Gamma_2(f, f) = \sum_{i,j=1}^n \left(\frac{d^2 f}{dx_i dx_j} \right)^2 + \langle \nabla f, \text{Hess}(\Phi) \nabla f \rangle \tag{4.20}$$

Assume that (4.20) holds true. Due to our additional assumption we then have

$$\Gamma_2(f, f) \geq \langle f, \text{Hess}(\Phi) \nabla f \rangle \geq \alpha |\nabla f|^2 = \alpha \Gamma_1(f, f),$$

which means that the curvature criterion (4.16) holds with constant α. Let us prove (4.20), we compute:

$$\begin{aligned}\mathcal{L}\Gamma_1(f, f) &= \mathcal{L}\big(|\nabla f|^2\big) \\ &= \sum_{i=1}^n \frac{d^2}{dx_i^2} \sum_{j=1}^n \left(\frac{df}{dx_j} \right)^2 - \sum_{i=1}^n \frac{d\Phi}{dx_i} \frac{d}{dx_i} \sum_{j=1}^n \left(\frac{df}{dx_j} \right)^2 \\ &= 2 \sum_{i,j=1}^n \left(\frac{d^2 f}{dx_i dx_j} \right)^2 + 2 \sum_{i,j=1}^n \frac{df}{dx_j} \frac{d^3}{dx_i^2 dx_j} - 2 \sum_{i,j=1}^n \frac{df}{dx_j} \frac{d^2 f}{dx_i dx_j} \frac{d\Phi}{dx_i}, \\ \Gamma_1(f, \mathcal{L}f) &= \sum_{i=1}^n \frac{df}{dx_i} \frac{d}{dx_i} \left(\sum_{j=1}^n \frac{d^2 f}{dx_j^2} - \frac{df}{dx_j} \frac{d\Phi}{dx_j} \right) \\ &= \sum_{i,j=1}^n \frac{df}{dx_i} \frac{d^3 f}{dx_i dx_j^2} - \sum_{i,j=1}^n \frac{df}{dx_i} \frac{d^2 f}{dx_i dx_j} \frac{d\Phi}{dx_j} - \sum_{i,j=1}^n \frac{df}{dx_i} \frac{df}{dx_j} \frac{d^2 \Phi}{dx_i dx_j},\end{aligned}$$

and by definition $\Gamma_2(f,f) = \frac{1}{2}\big[\mathcal{L}\Gamma_1(f,f) - 2\Gamma_1(f,\mathcal{L}f)\big]$, we relabel some indices:

$$\begin{aligned}
\Gamma_2(f,f) &= \frac{1}{2}\big[\mathcal{L}\Gamma_1(f,f) - 2\Gamma_1(f,\mathcal{L}f)\big] \\
&= \sum_{i,j=1}^{n}\left(\frac{d^2 f}{dx_i dx_j}\right)^2 + \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{d^3}{dx_i^2 dx_j} - \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{d^2 f}{dx_i dx_j}\frac{d\Phi}{dx_i} \\
&\quad - \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{d^3 f}{dx_j dx_i^2} + \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{d^2 f}{dx_j dx_i}\frac{d\Phi}{dx_i} + \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{df}{dx_i}\frac{d^2\Phi}{dx_j dx_i} \\
&= \sum_{i,j=1}^{n}\left(\frac{d^2 f}{dx_i dx_j}\right)^2 + \sum_{i,j=1}^{n}\frac{df}{dx_j}\frac{df}{dx_i}\frac{d^2\Phi}{dx_j dx_i} \\
&= \sum_{i,j=1}^{n}\left(\frac{d^2 f}{dx_i dx_j}\right)^2 + \big\langle f, \mathrm{Hess}(\Phi)\nabla f\big\rangle
\end{aligned}$$

This proves (4.20).

The last requirement of Theorem 4.23 which we need to check is that $(P_t)_t$ is weakly ergodic. This follows from the subsequent proposition, [3, Prop. 2.2], and the fact that we have $\Gamma_1(f,f) = |\nabla f|^2$ in our case. □

Proposition 4.26: *Let $\mathcal{L}$ be a generator and μ a reversible probability measure. Assume that all invariant functions, i.e. functions with $P_t f = f$, are contained in $\mathcal{A}$ and the only functions $g \in \mathcal{A}$ with $\Gamma_1(g,g) = 0$ are the constant functions. Then the associated semi-group $(P_t)_t$ is weakly ergodic.*

Proof. For the proof of this proposition we refer to [3, Prop. 2.2], it uses a spectral decomposition of $\mathcal{L}$. □

Remark 4.27: Here we want to remark that the computation in Theorem 4.24 can also be done on a compact Riemannian manifold (M,g). One can then compute for the Laplace-Beltrami operator $\mathcal{L} = \Delta_g$:

$$\Gamma_1(f,f) = g^{-1}(\nabla f, \nabla f), \quad \Gamma_2(f,f) = Ric(\nabla f, \nabla f) + \|\mathrm{Hess} f\|_2^2,$$

where Ric is the Ricci tensor. Furthermore, because we assume M to be compact there exists $\rho \in \mathbb{R}$ such that $Ric(\nabla f, \nabla f) \geq \rho\Gamma_1(f,f)$. Hence Δ_g satisfies the curvature criterion if $\rho > 0$. This is for example the case for the $n-$spheres $(\mathbb{S}^{n-1}, n \geq 3)$ where $\rho = n$. Details can be found in [1] and [4].

To conclude this section we want to show that a slightly weaker assumption than the curvature criterion implies a slightly weaker inequality, namely the spectral gap inequality. We have the following characterisation:

Theorem 4.28: *Let $\mathcal{L}$ be a generator and μ a reversible probability measure. Assume that the semi-group is weakly ergodic and let $\lambda > 0$. The following statements are equivalent:*

(a) $\forall f \in \mathcal{A} : \mu\big[\Gamma_2(f,f)\big] \geq \lambda\mu\big[\Gamma_1(f,f)\big]$.

(b) $\forall f \in A :$ $Var_\mu(f) \leq \frac{1}{\lambda}\mu\big[\Gamma_1(f,f)\big]$.

Proof. This is [1, Prop. 5.5.4] and we follow their proof.

Assume that (a) is true and fix $f \in \mathcal{A}$. Our assumption of weak ergodicity yields:

$$\mathrm{Var}_\mu(f) = \mu\big[P_0(f)^2\big] - \big[\lim_{t\to\infty} P_t(f)\big]^2 = -\int\int_0^\infty \frac{d}{ds}P_s(f)^2 ds d\mu$$

Further we have $\frac{d}{ds}P_s(f)^2 = 2(P_sf)\mathcal{L}(P_sf)$, hence:

$$\mathrm{Var}_\mu(f) = -2\int_0^\infty \int (P_sf)\mathcal{L}(P_sf)d\mu ds = 2\int_0^\infty \underbrace{\mu\big[\Gamma_1(P_sf,P_sf)\big]}_{:=\Psi(s)} ds \tag{4.21}$$

For $\Psi'(s)$ we observe:

$$\begin{aligned}\Psi'(s) &= 2\int (P_sf)\mathcal{L}^2(P_sf)d\mu = 2\int \Gamma_1(P_sf,\mathcal{L}P_sf)d\mu - \underbrace{\int \mathcal{L}(\Gamma_1(P_sf,P_sf)d\mu}_{=0} \\ &= -2\int \Gamma_2(P_sf,P_sf)d\mu \\ &\le -2\lambda\int \Gamma_1(P_sf,P_sf)d\mu = -2\lambda\Psi(s)\end{aligned}$$

where we have used our assumption (a). Grönwall's lemma implies $\Psi(s) \le e^{-2\lambda s}\Psi(0)$. Plugging this into (4.21) we get:

$$\mathrm{Var}_\mu(f) = 2\int_0^\infty \Psi(s)ds \le \int_0^\infty 2e^{-2\lambda s}ds\Psi(0) = \frac{1}{\lambda}\mu\big[\Gamma_1(f,f)\big]$$

This is the desired estimate.

For the converse implication, assume (b) to be true and fix $f \in \mathcal{A}$. With the reversibility of μ we obtain:

$$\mu\big[\Gamma_2(f,f)\big] = \tfrac{1}{2}\mu\big[\mathcal{L}\Gamma_1(f,f)\big] - \mu\big[\Gamma_1(f,\mathcal{L}f)\big] = -\mu\big[\Gamma_1(f,\mathcal{L}f)\big] = \mu\big[\big(\mathcal{L}f\big)^2\big]$$

And with $\mu\big[f\big]\mu\big[\mathcal{L}f\big] = 0$ we observe:

$$\mu\big[\Gamma_1(f,f)\big]^2 = \mu\big[-f\mathcal{L}f\big]^2 = \mu\big[\big(f-\mu[f]\big)\mathcal{L}f\big]^2 \le \mathrm{Var}_\mu(f)\mu\big[\big(\mathcal{L}f\big)^2\big],$$

where we used the Cauchy-Schwarz inequality. Using $\mu\big[\big(\mathcal{L}f\big)^2\big] = \mu\big[\Gamma_2(f,f)\big]$ and our assumption (b) we obtain:

$$\mu\big[\Gamma_1(f,f)\big]^2 \ge \lambda\mathrm{Var}_\mu(f)\mu\big[\Gamma_1(f,f)\big]$$

Dividing this inequality by $\lambda\mu\big[\Gamma_1(f,f)\big]$ yields the desired spectral gap inequality. □

5 Logarithmic Sobolev inequality in infinite volume

We recall the setting: In this section we want to perceive the circle as $[0,1]/0\sim 1 = \Omega_0$ where ν denotes the Lebesgue-measure on $[0,1]$ as reference measure. We work with the d-dimensional lattice $\mathbb{Z}^d$, the space of infinite configurations is denoted by $\Omega = \Omega_0^{\mathbb{Z}^d}$. $\mathcal{F}_\Lambda$ denotes the Borel-σ-algebra on Ω_0^Λ, implicitly $\mathcal{F}_\Lambda$ is embedded in $\mathcal{F} = \mathcal{F}_{\mathbb{Z}^d}$. For a measurable function $f\colon \Omega \to \mathbb{R}$, $\Lambda(f)$ is the smallest subset of $\mathbb{Z}^d$ such that f is $\mathcal{F}_{\Lambda(f)}$-measurable.

The Hamiltonian of the XY-model is

$$\begin{aligned} H_\Lambda(\omega_\Lambda\eta_{\Lambda^c}) &= -\beta \sum_{\substack{i\in\Lambda, j\in\mathbb{Z}^d \\ |i-j|=1}} \cos(2\pi((\omega_\Lambda\eta_{\Lambda^c})_i - (\omega_\Lambda\eta_{\Lambda^c})_j)) \\ &= -\beta \sum_{\substack{i,j\in\Lambda \\ |i-j|=1}} \cos(2\pi(\omega_i - \omega_j)) - \beta \sum_{\substack{i\in\Lambda, j\in\Lambda^c \\ |i-j|=1}} \cos(2\pi(\omega_i - \eta_j)), \end{aligned}$$

where $\omega \in \Omega_\Lambda$ for $\Lambda \Subset \mathbb{Z}$ and $\eta \in \Omega$ is the boundary condition. The associated Boltzmann measures are given by

$$\mu_\Lambda^\eta(d\omega) = \frac{e^{-H_\Lambda(\omega_\Lambda\eta_{\Lambda^c})}}{Z_{\Lambda^\eta}} \nu_\Lambda(d\omega),$$

where Z_Λ^η is an appropriate normalising constant. We want to perceive the μ_Λ^η as kernels which we can concatenate with each other and with measures, compare to Lemma 2.5. The measures on Ω that we are interested in are the associated Gibbs measures, we denote them by π - see Definition 2.6.

For later convenience, we further decompose the Hamiltonian:

$$H_\Lambda(\omega_\Lambda\eta_{\Lambda^c}) = \underbrace{-\beta \sum_{\substack{i,j\in\Lambda \\ |i-j|=1}} \cos(2\pi(\omega_i - \omega_j))}_{:=U_\Lambda(\omega)} \underbrace{-\beta \sum_{\substack{i\in\Lambda, j\in\Lambda^c \\ |i-j|=1}} \cos(2\pi(\omega_i - \eta_j))}_{:=W_\Lambda(\omega_\Lambda\eta_{\Lambda^c})} \tag{5.1}$$

This is to say that U_Λ are the interactions between spins in Λ and W_Λ are the interactions between spins in Λ and spins in Λ^c.

5.1 One dimension

The goal of this section is to show a logarithmic Sobolev inequality for the one-dimensional XY-model for all $\beta > 0$.

Our goal is the following:

Theorem 5.1: *The Gibbs measure π of the $XY-$model on $\mathbb{Z}$ satisfies a logarithmic Sobolev inequality.*

The proof of this theorem is long and requires the entire section. It follows a general scheme which is used to prove logarithmic Sobolev inequalities, we follow its presentation in [15, Chapter 5.2]. We leave the spaces on which the measures and kernels are defined in this proposition implicit because it works in a more general setup. Nevertheless, we assume that the carré du champ that we are interested in is $\Gamma_1(f,f) = |\nabla f|^2$.

5.1.1 The general scheme

Proposition 5.2: *Assume we have a probability measure μ and an auxiliary probability kernel $\boldsymbol{E}^\omega(\cdot)$ so that the following conditions are satisfied for some $\tilde\alpha < \infty$ and some $\tilde\kappa \in (0,1)$:*

$$(C1) \qquad \mu\boldsymbol{E} = \mu.$$

$$(C2) \qquad \mu\big[Ent_{\boldsymbol{E}}(f)\big] \le 2\tilde\alpha\mu\big[|\nabla\sqrt{f}|^2\big].$$

$$(C3) \qquad \mu\big[|\nabla\sqrt{\boldsymbol{E}^\omega[f]}|^2\big] \le \tilde\kappa\mu\big[|\nabla\sqrt{f}|^2\big].$$

$$(C4) \qquad \begin{cases} f_0 = f \\ f_{n+1} = \boldsymbol{E}[f_n] \end{cases} \qquad \text{we then have} \quad \lim_{n\to\infty} f_n = \mu\big[f\big].$$

for all $f \in \mathcal{A}$, compare to Hypothesis 2.15. Then μ satisfies a logarithmic Sobolev inequality as follows:

$$Ent_\mu(f) \le 2\frac{\tilde\alpha}{1-\tilde\kappa}\mu\big[|\nabla\sqrt{f}|^2\big]$$

Proof. Fix f bounded, measurable and non-negative. As our conditions already suggest, we consider the sequence

$$f_0 = f \quad f_{n+1}(\omega) = \mathbf{E}^\omega[f],$$

for which every f_n is again bounded and non-negative. A direct consequence of $(C1)$ is

$$\mu\big[f_{n+1}\big] = \mu\mathbf{E}\big[f_n\big] \overset{(C1)}{=} \mu\big[f_n\big] = \mu\big[f\big]$$

by induction. We can also use $(C1)$ to rewrite $\mathrm{Ent}_\mu(f)$ as follows:

$$\begin{aligned}\mathrm{Ent}_\mu(f) &= \mu\big[f\log f\big] - \mu\big[f\big]\log\mu\big[f\big] \\ &= \mu\mathbf{E}\big[f\log f\big] - \mu\big[\mathbf{E}[f]\log\mathbf{E}[f]\big] + \mu\big[\underbrace{\mathbf{E}[f]}_{=f_1}\log\underbrace{\mathbf{E}[f]}_{=f_1}\big] - \underbrace{\mu\mathbf{E}\big[f\big]}_{=\mu[f_1]}\log\underbrace{\mu\mathbf{E}\big[f\big]}_{=\mu[f_1]} \\ &= \mu\big[\mathrm{Ent}_{\mathbf{E}}(f_0)\big] + \mathrm{Ent}_\mu(f_1)\end{aligned}$$

and by iterating this we get

$$\mathrm{Ent}_\mu(f) = \sum_{n=0}^{N-1}\mu\big[\mathrm{Ent}_{\mathbf{E}}(f_n)\big] + \mathrm{Ent}_\mu(f_N)$$

for any $N \in \mathbb{N}$. We now use $(C2)$ and $(C3)$ to estimate $\mu\big[\mathrm{Ent}_{\mathbf{E}}(f_n)\big]$:

$$\mu\big[\mathrm{Ent}_{\mathbf{E}}(f_n)\big] \overset{(C2)}{\le} 2\tilde\alpha\mu\big[|\nabla\sqrt{f_n}|^2\big] \overset{(C3)}{\le} 2\tilde\alpha\tilde\kappa\mu\big[|\nabla\sqrt{f_{n-1}}|^2\big] \le 2\tilde\alpha\tilde\kappa^n\mu\big[|\nabla\sqrt{f}|^2\big]$$

where we used induction for the last inequality. Therefore:

$$\sum_{n=0}^{N-1}\mu\big[\mathrm{Ent}_{\mathbf{E}}(f_n)\big] \le \sum_{n=0}^{\infty}2\tilde\alpha\tilde\kappa^n\mu\big[|\nabla\sqrt{f}|^2\big] = \frac{2\tilde\alpha}{1-\tilde\kappa}\mu\big[|\nabla\sqrt{f}|^2\big]$$

It remains to show that $\mathrm{Ent}_\mu(f_N) \to 0$ as $N \to \infty$. This is where we use $(C4)$:

$$\mathrm{Ent}_\mu(f_N) \xrightarrow{N\to\infty} \mu\big[\mu[f]\log\mu[f]\big] - \mu\big[f\big]\log\mu\big[f\big] = 0,$$

where we further used the dominated convergence theorem and the fact that $(f_n)_{n\in\mathbb{N}_0}$ is uniformly bounded. This completes the proof. □

5.1.2 The technical lemmata

Before we construct the auxiliary kernel and before we check the conditions of Proposition 5.2, we state and prove a few technical lemmata. Lemma 5.3 is a typical property of one-dimensional models: the dependence of $\mu_\Lambda[f]$ on the boundary condition decays exponentially fast when f is fixed and Λ grows. The two following lemmata are equivalent to [15, Lemma 5.3] which concerns itself with the Ising model. We adapt the proof to the XY-model, the notable difference being that here the derivative satisfies the Leibniz rule unlike the discrete derivative.

For later convenience we define, $\Delta \subset \mathbb{Z}$ and f measurable:

$$osc_\Delta(f) = \sup_{\sigma_{\Delta^c} = \tilde{\sigma}_{\Delta^c}} |f(\sigma) - f(\tilde{\sigma})|, \tag{5.2}$$

the maximal oscillation of f when the spins outside Δ are fixed.

Lemma 5.3: *There exists $\gamma > 0$ such that for any continuous function f with $\Lambda(f) = \{0, ..., N\}$ for some N we have:*

$$\sup_{x,y,z \in [0,1]} \left| \mu^{\{z,x\}}_{\{1,...,n\}}[f] - \mu^{\{z,y\}}_{\{1,...,n\}}[f] \right| \leq C(f) e^{-\gamma(n-N)},$$

for all $n \geq N$ and for some $C(f)$ which depends only on f. Here, $\mu^{\{z,x\}}_{\{1,...,n\}}$, denotes the finite volume Gibbs measure with boundary condition z for the spin at 0 and x for the spin at $n+1$. The above inequality can also be formulated as

$$osc_{\{1,...,n+1\}}(\mu^{\cdot}_{\{1,...,n\}}[f]) \leq C(f) e^{-\gamma(n-N)}.$$

Lemma 5.4: *For any finite $\Lambda \Subset \mathbb{Z}$ there exists $B(\Lambda))$ such that for any $i \in \mathbb{Z}$:*

$$\pi\big[\big(\nabla_i \sqrt{\mu_\Lambda[f]}\big)^2\big] \leq 2\pi\big[\big(\nabla_i \sqrt{f}\big)^2\big] + B(\Lambda)\pi\big[\big(\nabla_\Lambda \sqrt{f}\big)^2\big], \tag{5.3}$$

for any local $f \in C^1(\Omega)$ with $f \geq 0$ and any Gibbs measure π. Further, $B(l) = \sup_{|\Lambda| \leq l} B(\Lambda) < \infty$.

Lemma 5.5: *There exist $L_0 \in \mathbb{N}$ and $\kappa > 0$ with $\kappa \max\{2, B(l)\} < 1$ where $l = 2L + 2$ and $L \geq L_0$ such that for all intervals $\Lambda \Subset \mathbb{Z}, |\Lambda| = l$ and $\Delta \subset \Lambda$ with $dist(\Delta, \Lambda^c) \geq L - 1$ we have*

$$\pi\big[\big(\nabla_i \sqrt{\mu_\Lambda[f]}\big)^2\big] \leq \kappa \pi\big[\big(\nabla_\Lambda \sqrt{f}\big)^2\big],$$

for any $i \in \mathbb{Z}$ and any local $f \in C^1(\Omega)$ with $f \geq 0$ and any Gibbs measure π. Furthermore we must assume that f does not depend on the spins in $\Lambda \backslash \Delta$, i.e. f is assumed to be $\mathcal{F}_{\Delta \cup \Lambda^c}$-measurable.

Proof of Lemma 5.3. The basic idea of the proof is to uncover the spins one at the time starting from the right. Every time we do this, the new spin equals a uniform spin in distribution with probability $e^{-2\beta}$, hence with probability $e^{-4\beta}$ we can couple the spins of $\mu^{z,x}$ and $\mu^{z,y}$ so that the effect of the boundary condition is lost. We formalise this idea but instead of coupling the spins we decompose the expectations.

We first define a sequence of intervals

$$\Delta_k = \{1, ..., N + k\} \quad \text{for} \quad 0 \leq k \leq n - N.$$

Observe that $\mu^{\{z,\sigma_{N+k+2}\}}_{\Delta_{k+1}}$ has a density with respect to $\mu^{\{z,\sigma_{N+k+1}\}}_{\Delta_k} \otimes \nu$:

$$\mu^{\{z,\sigma_{N+k+2}\}}_{\Delta_{k+1}}(d\sigma) = \underbrace{\frac{e^{\beta\cos(2\pi(\sigma_{N+k+2} - \sigma_{N+k+1}))}}{\int e^{\beta\cos(2\pi(\sigma_{N+k+2} - \omega))}\nu(d\omega)}}_{=\varphi(\sigma_{N+k+1}, \sigma_{N+k+2})} \mu^{\{z,\sigma_{N+k+1}\}}_{\Delta_k}(d\sigma)\nu(d\sigma_{N+k+1}),$$

by a slight abuse of notation letting $\sigma = (\sigma_1, ..., \sigma_\ell)$ for appropriate $\ell \in \mathbb{Z}$. Note that for any $\sigma_{N+k+2} \in [0,1]$ we have $\int \varphi(\omega, \sigma_{N+k+2})\nu(d\omega) = 1$. More importantly, we have

$$\varphi(\sigma_{N+k+1}, \sigma_{N+k+2}) \geq e^{-2\beta}.$$

Abbreviate $\psi(\sigma_{N+k+1}, \sigma_{N+k+2}) = \varphi(\sigma_{N+k+1}\sigma_{N+k+2}) - e^{-2\beta}$. To make the following more readable, we suppress some dependencies: write $\mu_{\Delta_{k+1}}(d\sigma)$ for $\mu_{\Delta_{k+1}}^{\{z,\sigma_{N+k+2}\}}(d\sigma)$ and $\tilde{\mu}_{\Delta_{k+1}}(d\tilde{\sigma})$ for $\mu_{\Delta_{k+1}}^{\{z,\tilde{\sigma}_{N+k+2}\}}(d\tilde{\sigma})$ - and similarly $f, \tilde{f}, \psi, \tilde{\psi}$. We then have:

$$\begin{aligned}\left|\mu_{\Delta_{k+1}}[f] - \tilde{\mu}_{\Delta_{k+1}}[\tilde{f}]\right| &= \left|\iint (f - \tilde{f})\mu_{\Delta_{k+1}}(d\sigma)\tilde{\mu}_{\Delta_{k+1}}(d\tilde{\sigma})\right| \\ &= \left|\iint (f - \tilde{f})(\psi\tilde{\psi} + e^{-2\beta}\psi + e^{-2\beta}\tilde{\psi})\nu(d\sigma)\nu(d\tilde{\sigma})\mu_{\Delta_k}(d\sigma)\tilde{\mu}_{\Delta_k}(d\tilde{\sigma})\right. \\ &\qquad \left. + \iint (f - \tilde{f})e^{-4\beta}\nu(d\sigma)\nu(d\tilde{\sigma})\mu_{\Delta_k}(d\sigma)\tilde{\mu}_{\Delta_k}(d\tilde{\sigma})\right|\end{aligned}$$

The last term evaluates to 0 because both μ_{Δ_k} and $\tilde{\mu}_{\Delta_k}$ have the same boundary condition. Hence, using that f does not depend on the spin σ_{N+k+1}:

$$\begin{aligned}\left|\mu_{\Delta_{k+1}}[f] - \tilde{\mu}_{\Delta_{k+1}}[\tilde{f}]\right| &= \left|\iint (f - \tilde{f})(\psi\tilde{\psi} + e^{-2\beta}\psi + e^{-2\beta}\tilde{\psi})\nu(d\sigma)\nu(d\tilde{\sigma})\mu_{\Delta_k}(d\sigma)\tilde{\mu}_{\Delta_k}(d\tilde{\sigma})\right| \\ &\leq \left|\iint (\psi\tilde{\psi} + e^{-2\beta}\psi + e^{-2\beta}\tilde{\psi})\nu(d\sigma)\nu(d\tilde{\sigma})\right| \sup_{x,y}\left|\mu_{\Delta_k}^{\{z,x\}}[f] - \mu_{\Delta_k}^{\{z,y\}}[f]\right| \\ &= \left(1 - e^{-4\beta}\right)\sup_{x,y}\left|\mu_{\Delta_k}^{\{z,x\}}[f] - \mu_{\Delta_k}^{\{z,y\}}[f]\right|\end{aligned}$$

Taking suprema, we have shown

$$\sup_{x,y,z}\left|\mu_{\Delta_{k+1}}^{\{z,x\}}[f] - \mu_{\Delta_{k+1}}^{\{z,y\}}[f]\right| \leq \left(1 - e^{-4\delta}\right)\sup_{x,y,z}\left|\mu_{\Delta_k}^{\{z,x\}}[f] - \mu_{\Delta_k}^{\{z,y\}}[f]\right|.$$

The desired statement now follows by induction and $\gamma = -\log(1 + e^{-4\delta})$ and

$$C(f) = \sup_{x,y,z}\left|\mu_{\{1,\dots,N\}}^{\{z,x\}}[f] - \mu_{\{1,\dots,N\}}^{\{z,y\}}[f]\right|.$$

□

Proof of Lemma 5.4. First we remark that because f is in $C^1(\Omega)$ and local, we may always interchange derivatives and ν integrals.

Let us observe:

$$\nabla_i\sqrt{\mu_\Lambda^\omega[f]} = \frac{1}{2\sqrt{\mu_\Lambda^\omega[f]}}\nabla_i\mu_\Lambda^\omega[f],$$

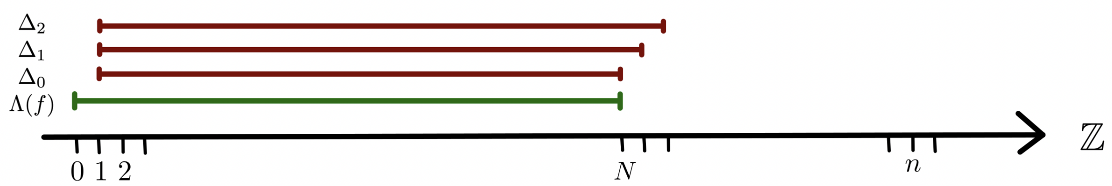


Figure 3: Some of the different subsets of $\mathbb{Z}$ in the proof of Lemma 5.3.

hence we thrive to find an estimate of the form $\nabla_i\mu_\Lambda^\omega[f] \le \sqrt{\mu_\Lambda^\omega[f]}\cdot(...)$.

To handle μ_Λ expectations, we write $\rho_\Lambda(\sigma_\Lambda\omega_{\Lambda^c}) = \frac{\exp(-H_\Lambda(\sigma_\Lambda\omega_{\Lambda^c}))}{\nu_\Lambda\left[\exp(-H_\Lambda(\cdot\omega_{\Lambda^c})\right]}$ which is the density of μ_Λ^ω with respect to ν_Λ. We then have

$$\nabla_i\mu_\Lambda^\omega[f] = \nu_\Lambda[\nabla_i\rho_\Lambda f] = \nu_\Lambda[\rho_\Lambda\nabla_i f] + \nu_\Lambda[f\nabla_i\rho_\Lambda] = \mu_\Lambda^\omega[\nabla_i f] + \mu_\Lambda^\omega[f\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]. \tag{5.4}$$

The first term is just an application of the Cauchy-Schwarz inequality away from the desired form. Write $f = f^{1/2}f^{1/2}$, then:

$$\mu_\Lambda^\omega[\nabla_i f] = 2\mu_\Lambda^\omega[f^{1/2}\nabla_i f^{1/2}] \le 2\mu_\Lambda^\omega[f]^{1/2}\mu_\Lambda^\omega[(\nabla_i f^{1/2})^2]^{1/2} \tag{5.5}$$

The estimate of the second term of (5.4) is a bit more work. Observe:

$$\mu_\Lambda^\omega[\rho_\Lambda^{-1}\nabla_i\rho_\Lambda] = \nu_\Lambda\left[\nabla_i\frac{e^{-H_\Lambda}}{\nu_\Lambda[e^{-H_\Lambda}]}\right] = \frac{\nu_\Lambda[\nabla_i e^{-H_\Lambda}]}{\nu_\Lambda[e^{-H_\Lambda}]} - \frac{\nu_\Lambda[e^{-H_\Lambda}]\nu_\Lambda[\nabla_i e^{-H_\Lambda}]}{\nu_\Lambda[e^{-H_\Lambda}]^2} = 0$$

This allows us to estimate:

$$\begin{aligned}
\left|\mu_\Lambda^\omega[f\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\right| &= \left|\mu_\Lambda^\omega[(f-\mu_\Lambda^\omega[f])\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\right| && (5.6)\\
&= \frac{1}{2}\left|\iint(\rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\sigma) - \rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\tilde\sigma))(f(\sigma)-f(\tilde\sigma))\mu_\Lambda^\omega(d\sigma)\mu_\Lambda^\omega(d\tilde\sigma)\right|\\
&\le \sup_{\sigma_{\Lambda^c}=\tilde\sigma_{\Lambda^c}}\left|\rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\sigma) - \rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\tilde\sigma)\right| \underbrace{\iint\left|f(\sigma)-f(\tilde\sigma)\right|\mu_\Lambda^\omega(d\sigma)\mu_\Lambda^\omega(d\tilde\sigma)}_{=:A} && (5.7)
\end{aligned}$$

We estimate the two factors separately. Uniformly in Λ we have:

$$\begin{aligned}
&\sup_{\sigma_{\Lambda^c}=\tilde\sigma_{\Lambda^c}}\left|\rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\sigma) - \rho_\Lambda^{-1}\nabla_i\rho_\Lambda(\tilde\sigma)\right| =\\
&\qquad= \sup_{\sigma_{\Lambda^c}=\tilde\sigma_{\Lambda^c}}\left|\big(-\nabla_i H_\Lambda(\sigma) + \nabla_i H_\Lambda(\tilde\sigma)\big) + \underbrace{\left(\frac{\nu[\nabla_i e^{-H_\Lambda}]}{\nu[e^{-H_\Lambda}]}(\sigma) - \frac{\nu[\nabla_i e^{-H_\Lambda}]}{\nu[e^{-H_\Lambda}]}(\tilde\sigma)\right)}_{=0 \text{ because } \sigma_{\Lambda^c}=\tilde\sigma_{\Lambda^c}}\right|\\
&\qquad\le 2\left\|\nabla_i H_\Lambda\right\|_\infty\\
&\qquad\le 4\beta =: M && (5.8)
\end{aligned}$$

For the second factor of (5.7), A, we write $f(\sigma)-f(\tilde\sigma) = \big(f^{1/2}(\sigma)-f^{1/2}(\tilde\sigma)\big)\big(f^{1/2}(\sigma)+f^{1/2}(\tilde\sigma)\big)$:

$$\begin{aligned}
A^2 &= \left(\iint\left|\big(f^{1/2}(\sigma)-f^{1/2}(\tilde\sigma)\big)\big(f^{1/2}(\sigma)+f^{1/2}(\tilde\sigma)\big)\right|\mu_\Lambda^\omega(d\sigma)\mu_\Lambda^\omega(d\tilde\sigma)\right)^2\\
&\le \mu_\Lambda^\omega\otimes\mu_\Lambda^\omega\left[\big(f(\sigma)-f(\tilde\sigma)\big)^2\right] + \mu_\Lambda^\omega\otimes\mu_\Lambda^\omega\left[\big(f(\sigma)+f(\tilde\sigma)\big)^2\right]\\
&= \big(2\mu_\Lambda^\omega[f] - 2\mu_\Lambda^\omega[f^{1/2}]^2\big)\big(2\mu_\Lambda^\omega[f] + 2\underbrace{\mu_\Lambda^\omega[f^{1/2}]^2}_{\le\mu_\Lambda^\omega[f]}\big)\\
&\le 8\mu_\Lambda^\omega[f]\operatorname{Var}_{\mu_\Lambda^\omega}(f^{1/2}) && (5.9)
\end{aligned}$$

where we used first the Chauchy-Schwarz inequality and then Jensen's inequality. The occurrence of the variance in (5.9) allows us to use the spectral gap inequality for μ_Λ^ω. By Proposition 3.9 we obtain

$$8\mu_\Lambda^\omega[f]\operatorname{Var}_{\mu_\Lambda^\omega}(f^{1/2}) \le \frac{2e^{24\beta}}{\pi}|\Lambda|\mu_\Lambda^\omega[f]\mu_\Lambda^\omega\left[(\nabla_\Lambda f^{1/2})^2\right].$$

Combining these estimates with (5.8) and (5.5), and applying them to (5.4) yields:

$$\nabla_i\mu_\Lambda^\omega[f] \le 2\mu_\Lambda^\omega[f]^{1/2}\mu_\Lambda^\omega\left[(\nabla_i f^{1/2})^2\right]^{1/2} + e^{12\beta}\sqrt{|\Lambda|}M\mu_\Lambda^\omega[f]^{1/2}\mu_\Lambda^\omega\left[(\nabla_\Lambda f^{1/2})^2\right]^{1/2} \tag{5.10}$$

And therefore:

$$\nabla_i(\mu_\Lambda^\omega[f])^{1/2} \leq \frac{1}{2(\mu_\Lambda^\omega[f])^{1/2}}\nabla_i\mu_\Lambda^\omega[f] \leq \mu_\Lambda^\omega\big[(\nabla_i f^{1/2})^2\big]^{1/2} + \frac{1}{2}e^{12\beta}\sqrt{|\Lambda|}M\mu_\Lambda^\omega[f]^{1/2}\mu_\Lambda^\omega\big[(\nabla_\Lambda f^{1/2})^2\big]^{1/2}$$

Squaring this equation, using $(a+b)^2 \leq 2a^2 + 2b^2$ and taking π-expectations yields:

$$\pi\big[\big(\nabla_i\sqrt{\mu_\Lambda[f]}\big)^2\big] \leq 2\pi\big[\big(\nabla_i\sqrt{f}\big)^2\big] + \frac{e^{24\beta}}{2}|\Lambda|M^2\pi\big[\big(\nabla_\Lambda\sqrt{f}\big)^2\big], \tag{5.11}$$

which is the desired equation with $B(\Lambda) = \frac{e^{24\beta}}{2}|\Lambda|M^2$.

□

Proof of Lemma 5.5. This lemma is an improvement over the previous lemma under some additional assumptions. Hence we want amend the its proof to apply Lemma 5.3 to obtain the smallness of κ.

We start by reconsidering (5.4). First, we note that if $i \in \Lambda$, then we have $\nabla_i\mu_\Lambda^\omega[f] = 0$ because $\mu_\Lambda^\omega[f]$ does not depend on the spins in Λ. Hence we can restrict ourselves to $i \in \Lambda^c$. If we assume f to be localised in $\Delta \subset \Lambda$, then we further have $\nabla_i f = 0$ and thus:

$$\nabla_i\mu_\Lambda^\omega[f] = \mu_\Lambda^\omega[f\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]$$

We can now improve (5.6) because f does not depend on $\Lambda\backslash\Delta$:

$$\big|\mu_\Lambda^\omega[f\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\big| = \big|\mu_\Lambda^\omega\big[(f - \mu_\Lambda^\omega[f])\mu_{\Lambda\backslash\Delta}[\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\big]\big|$$

where we used $\mu_\Lambda\mu_{\Lambda\backslash\Delta} = \mu_\Lambda$. (5.7) then becomes

$$\big|\mu_\Lambda^\omega[f\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\big| \leq osc_\Lambda\big(\mu_{\Lambda\backslash\Delta}[\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\big)A,$$

where osc_Λ is defined in (5.2). By reusing the estimates for A we arrive at the following equivalent to (5.11):

$$\pi\big[\big(\nabla_i\sqrt{\mu_\Lambda[f]}\big)^2\big] \leq \underbrace{\frac{e^{24\beta}}{2}|\Lambda|\Big(osc_\Lambda\big(\mu_{\Lambda\backslash\Delta}[\rho_\Lambda^{-1}\nabla_i\rho_\Lambda]\big)\Big)^2}_{=:\kappa_\Lambda}\pi\big[\big(\nabla_\Lambda\sqrt{f}\big)^2\big]$$

Lemma 5.3 now implies that $\kappa_\Lambda \xrightarrow{L\to\infty} 0$ and hence we can choose L_0 large enough such that κ_Λ satisfies the smallness constraints for any interval Λ with $|\Lambda| \geq l$. □

5.1.3 The construction of the auxiliary kernel and the proof

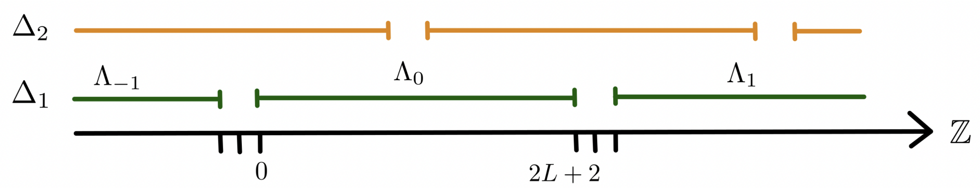


Figure 4: Δ_1 and Δ_2 used for the construction of the auxiliary kernel.

Having dealt with some technicalities, we now construct the kernel $\mathbf{E}$ which we use in combination with Proposition 5.2 to show that the Gibbs measure of the XY-model satisfies a logarithmic Sobolev inequality.

The basic idea is that we partition $\mathbb{Z}$ into disjoint intervals of the same length which will be updated according to the local Gibbs measures μ_Λ. To guarantee that every spin gets updated at least once, we need to do this twice for different subsets of $\mathbb{Z}$. Let us define this formally:

Let $L \in \mathbb{N}$ - we fix it later when we use Lemma 5.5. Define $\Lambda_0 = \{0, ..., 2L+2\}$ and $\Lambda_k = \Lambda_0 + 2k(L+2)$, let $\Delta_1 = \bigcup_{k\in\mathbb{Z}} \Lambda_k$ and $\Delta_2 = \Delta_1 + (L+2)$. Further we define two kernels: $\mathbf{E}_1^\omega = \bigotimes_{k\in\mathbb{Z}} \mu_{\Lambda_k}^\omega$ and analogously $\mathbf{E}_2^\omega$ using the building blocks of Δ_2. Lastly, we set $\mathbf{E}^\omega = \mathbf{E}_2^\omega \mathbf{E}_1$ as the concatenation.

$\mathbf{E}$ does not quite satisfy a LSI in the sense of condition $(C2)$ of Proposition 5.2 straight away. Nevertheless we observe that for any fixed ω, $\mathbf{E}_1^\omega$ is a product measure of the μ_Λ of which we know that they satisfy a LSI with the same constant $\alpha \le e^{16\beta L}$ by the perturbation property, Theorem 4.4. Hence, by the product property, Theorem 4.5, $\mathbf{E}_1^\omega$ satisfies a LSI of the form

$$\mathrm{Ent}_{\mathbf{E}_1^\omega}(f^2) \le 2\alpha \mathbf{E}_1^\omega\big[(\nabla_{\Delta_1} f)^2\big]. \tag{5.12}$$

Further, we can choose α uniformly in the boundary condition ω. The same holds for $\mathbf{E}_2^\omega$ with respect to ∇_{Δ_2} and the same constant α.

We now prove Theorem 5.1 by checking conditions $(C1)-(C4)$ of Proposition 5.2. In the following $f \in C^\infty(\Omega)$ with $f \ge 0$ shall always be a local function, i.e. depending only on finitely many spins. This allows us to always interchange expectations and derivatives or infinite sums. π shall always be the Gibbs measure.

Proof of $(C1)$. This follows directly from the definition of Gibbs measures, $\pi\mu_\Lambda = \pi$ and therefore $\pi\mathbf{E}_1 = \pi = \pi\mathbf{E}_2$. We then have

$$\pi\big[\mathbf{E}[f]\big] = \pi\big[\mathbf{E}_2\mathbf{E}_1[f]\big] = \pi\big[\mathbf{E}_1[f]\big] = \pi\big[f\big].$$

□

Proof of $(C2)$. Recall our goal:

$$\pi\big[\mathrm{Ent}_{\mathbf{E}}(f)\big] \le 2\tilde{\alpha}\pi\big[\big(\nabla_{\mathbb{Z}} f^{1/2}\big)^2\big]$$

for some $\tilde{\alpha}$. Using $\pi\mathbf{E}_2 = \pi$, we can decompose $\pi\big[\mathrm{Ent}_{\mathbf{E}}(f)\big]$:

$$\begin{aligned}\pi\big[\mathrm{Ent}_{\mathbf{E}}(f)\big] &= \pi\big[\mathbf{E}[f\log f] - \mathbf{E}[f]\log\mathbf{E}[f]\big] \\ &= \pi\big[\mathbf{E}_1[f\log f] - \mathbf{E}_1[f]\log\mathbf{E}_1[f]\big] + \pi\big[\mathbf{E}_1[f]\log\mathbf{E}_1[f] - \mathbf{E}_2\mathbf{E}_1[f]\log\mathbf{E}_2\mathbf{E}_1[f]\big] \\ &= \pi\big[\mathrm{Ent}_{\mathbf{E}_1}(f)\big] + \pi\big[\mathrm{Ent}_{\mathbf{E}_2}(\mathbf{E}_1[f])\big]\end{aligned} \tag{5.13}$$

This puts us in a place to apply the LSI of $\mathbf{E}_1$ and $\mathbf{E}_2$. Using $\pi\mathbf{E}_1 = \pi$:

$$\pi\big[\mathrm{Ent}_{\mathbf{E}_1}(f)\big] \le 2\alpha\pi\big[\big(\nabla_{\Delta_1} f^{1/2}\big)^2\big], \tag{5.14}$$

where α comes from (5.12). We do the same for the second term of (5.13) and use $\pi\mathbf{E}_2 = \pi$:

$$\pi\big[\mathrm{Ent}_{\mathbf{E}_2}(\mathbf{E}_1[f])\big] \le 2\alpha\pi\big[\big(\nabla_{\Delta_2}(\mathbf{E}_1[f])^{1/2}\big)^2\big] \tag{5.15}$$

This is not quite what we want, we need f instead of $\mathbf{E}_1[f]$ on the right-hand side. First, notice that $\mathbf{E}_1[f]$ does not depend on the spins in Δ_1, hence $\nabla_{\Delta_2}(\mathbf{E}_1[f])^{1/2} = \nabla_{\Delta_2\backslash\Delta_1}(\mathbf{E}_1[f])^{1/2}$. Fix $i \in \Delta_2\backslash\Delta_1$. There exists a set $\Lambda^{(i)} \subset \Delta_1$ with $|\Lambda^{(i)}| \le 2L+2$ such that:

$$\nabla_i \mathbf{E}_1[f] = \nabla_i \mathbf{E}_1 \mu_{\Lambda^{(i)}}[f] = \mathbf{E}_1 \nabla_i \mu_{\Lambda^{(i)}}[f] \tag{5.16}$$

And with that:

$$\pi\big[\big(\nabla_i(E_1[f])^{1/2}\big)^2\big] \le \pi\big[\big(\nabla_i(\mu_{\Lambda^{(i)}}[f])^{1/2}\big)^2\big],$$

which puts us in a place to apply Lemma 5.4:

$$\pi\big[\big(\nabla_i(\mu_{\Lambda^{(i)}}[f])^{1/2}\big)^2\big] \leq 2\pi\ \big[\big(\nabla_i f^{1/2}\big)^2\big] + B(l)\pi\big[\big(\nabla_{\Lambda^{(i)}} f^{1/2}\big)^2\big]$$

Combining the above estimates yields:

$$\begin{aligned}\pi\big[\mathrm{Ent}_{\mathbf{E}_2}(\mathbf{E}_1[f])\big] &\leq 2\alpha \sum_{i\in\Delta_2\backslash\Delta_1} 2\pi\big[\big(\nabla_i f^{1/2}\big)^2\big] + B(l)\pi\big[\big(\nabla_{\Lambda^{(i)}} f^{1/2}\big)^2\big] \\ &\leq 2\alpha \sum_{i\in\Delta_2\backslash\Delta_1} 2\pi\big[\big(\nabla_i f^{1/2}\big)^2\big] + 2\alpha B(l) l \sum_{i\in\Delta_0} 2\pi\big[\big(\nabla_i f^{1/2}\big)^2\big]\end{aligned}$$

And together with (5.13) and (5.14) we obtain:

$$\pi\big[\mathrm{Ent}_{\mathbf{E}}(f)\big] \leq 2\alpha(1 + lB(l)) \sum_{i\in\mathbb{Z}} \pi\big[\big(\nabla_i f^{1/2}\big)^2\big],$$

which is $(C2)$ with $\tilde{\alpha} = \alpha(1 + lB(l))$.

□

Proof of $(C3)$. Recall our goal:

$$\pi\big[\big|\nabla\sqrt{\mathbf{E}[f]}\big|^2\big] \leq \tilde{\kappa}\pi\big[|\nabla\sqrt{f}|^2\big],$$

for a $\tilde{\kappa}$ sufficiently small, namely $\tilde{\kappa} < 1$.

First of all, we note that $\mathbf{E}[f] = \mathbf{E}_2[\mathbf{E}_1[f]]$ only depends on the spins in $\Delta_1^c = \Delta_2\backslash\Delta 1$, hence:

$$\pi\big[\big|\nabla\mathbf{E}[f]^{1/2}\big|^2\big] = \sum_{i\in\Delta_2\backslash\Delta_1} \pi\big[\big|\nabla_i\mathbf{E}[f]^{1/2}\big|^2\big]$$

Next, we fix $i \in \Delta_2\backslash\Delta_1$. Let $\Lambda^{(i)}$ be like in (5.16), except that we now have $\Lambda^{(i)} \subset \Delta_2$, and like before we have

$$\pi\big[\big|\nabla_i(\mathbf{E}_2\mathbf{E}_1[f])^{1/2}\big|^2\big] \leq \pi\big[\big(\nabla_i(\mu_{\Lambda^{(i)}}\mathbf{E}_1[f])^{1/2}\big)^2\big].$$

This is the place where we can apply Lemma 5.5 namely to $g = \mathbf{E}_1[f]$. We are allowed to do this because $\mathbf{E}_1[f]$ does not depend on the spins in Δ_1. Hence:

$$\pi\big[\big(\nabla_i(\mu_{\Lambda^{(i)}}\mathbf{E}_1[f])^{1/2}\big)^2\big] \leq \kappa\pi\big[\big(\nabla_{\Lambda^{(i)}}(\mathbf{E}_1[f])^{1/2}\big)^2\big]$$

For each $j \in \Lambda^{(i)}$ let $\tilde{\Lambda}^{(j)} \subset \Delta_1$ be, again like in (5.16), such that we have:

$$\begin{aligned}\kappa\pi\big[\big(\nabla_j(\mathbf{E}_1[f])^{1/2}\big)^2\big] &\leq \kappa\pi\big[\big(\nabla_j(\mu_{\Lambda^{(j)}}[f])^{1/2}\big)^2\big] \\ &\leq 2\kappa\pi\big[\big(\nabla_j f\big)^2\big] + \kappa B(l)\pi\big[\big(\nabla_{\Lambda^{(j)}} f^{1/2}\big)^2\big],\end{aligned}$$

where we applied Lemma 5.4 once again. Collecting all our estimates we obtain:

$$\pi\big[\big(\nabla(\mathbf{E}[f])^{1/2}\big)\big] \leq 2\kappa \sum_{i\in\Delta\backslash\Delta_1} \pi\big[\big(\nabla_i f^{1/2}\big)^2\big] + \kappa B(l) \sum_{i\in\Delta_1} \pi\big[\big(\nabla_i f^{1/2}\big)^2\big]$$

Hence $(C3)$ holds with $\tilde{\kappa} = \kappa\max\{2, B(l)\}$ and $\tilde{\kappa} < 1$ according to Lemma 5.5 which we used to choose κ. □

Proof of $(C4)$. Recall our goal: define a sequence of functions, $f_0 = f$ and $f_{n+1} = \mathbf{E}[f_n]$. We want to show that $(f_n)_{n\in\mathbb{N}}$ converges π-almost surely to 0.

First, let us note that by Theorem 4.7 both $\mathbf{E}_1$ and $\mathbf{E}_2$ satisfy a spectral gap inequality with constant $1/\alpha$ where α comes from (5.12). By using $(a+b)^2 \leq 2a^2 + 2b^2$ we obtain:

$$\begin{aligned}
\pi\big[\big(f - \mathbf{E}[f]\big)^2\big] &= \pi\big[\big((f - \mathbf{E}_1[f]) + (\mathbf{E}_1[f] - \mathbf{E}[f])\big)^2\big] \\
&\leq 2\pi\big[\big(f - \mathbf{E}_1[f]\big)^2\big] + 2\pi\big[\big(\mathbf{E}_1[f] - \mathbf{E}_2[\mathbf{E}_1[f]]\big)^2\big] \\
&= 2\pi\big[\mathrm{Var}_{\mathbf{E}_1}(f)\big] + 2\pi\big[\mathrm{Var}_{\mathbf{E}_2}(\mathbf{E}_1[f])\big] \\
&\leq 2\alpha\pi\big[\big(\nabla_{\Delta_1} f\big)^2\big] + 2\alpha\pi\big[\big(\nabla_{\Delta_2}\mathbf{E}_1[f]\big)^2\big]
\end{aligned}$$

By reasoning that we have seen before in the proofs of $(C2)$ and $(C3)$ we obtain, by using Lemma 5.4, the following estimate:

$$\pi\big[\big(\nabla_{\Delta_2}\mathbf{E}_1[f]\big)^2\big] \leq K\pi\big[\big(\nabla_{\Delta_2} f\big)^2\big],$$

for some K large enough. Combined:

$$\pi\big[\big(f - \mathbf{E}[f]\big)^2\big] \leq 2\alpha(K+1)\pi\big[\big(\nabla f\big)^2\big], \tag{5.17}$$

where $\nabla = \nabla_{\mathbb{Z}}$. Furthermore, $(C3)$ implies

$$\pi\big[\big(\nabla f_n^{1/2}\big)^2\big] \leq \tilde{\kappa}\pi\big[\big(\nabla f_{n-1}^{1/2}\big)^2\big] \leq \tilde{\kappa}^n\pi\big[\big(\nabla f^{1/2}\big)^2\big]. \tag{5.18}$$

Combining these two estimates, we can show that f_n converges to a constant by the Borel-Cantelli Lemma, let $\varepsilon > 0$:

$$\begin{aligned}
\sum_{n=0}^{\infty} \pi\big[\{|f_{n+1} - f_n| > \varepsilon\}\big] &\leq \frac{1}{\varepsilon^2}\sum_{n=0}^{\infty}\pi\big[\underbrace{(f_n - f_{n+1})^2}_{=(f_n - \mathbf{E}[f])^2}\big] \\
&\overset{(5.17)}{\leq} \frac{2\alpha(K+1)}{\varepsilon^2}\sum_{n=0}^{\infty}\pi\big[(\nabla f_n^{1/2} f_n^{1/2})^2\big] \\
&= \frac{4\alpha(K+1)}{\varepsilon^2}\sum_{n=0}^{\infty}\pi\big[f_n^{1/2}\nabla f_n^{1/2}\big]
\end{aligned}$$

The last expectation can be estimated by the Cauchy-Schwarz inequality:

$$\pi\big[f_n^{1/2}\nabla f_n^{1/2}\big] \leq \pi\big[f_n\big]^{1/2}\pi\big[(\nabla f_n^{1/2})^2\big]^{1/2} \overset{(5.18)}{\leq} \tilde{\kappa}^{n/2}\pi\big[\big(\nabla f^{1/2}\big)^2\big]^{1/2}\pi\big[f\big]$$

where we again used $(C1)$ as $\pi[f_n] = \pi[f]$. This proves that $\pi\big[\{|f_{n+1} - f_n| > \varepsilon\}\big]$ is summable and hence f_n converges π-almost surely to a constant. Because $\pi[f_n] = \pi[f]$ for all $n \geq 0$, we have $\lim n \to \infty f_n = \pi[f]$ π-almost surely. □

Proof of Theorem 5.1. The statement of the theorem now follows directly from Proposition 5.2. □

5.2 Higher dimensions and high temperature

In the previous section we have proven a logarithmic Sobolev inequality for the XY-model on a one-dimensional lattice for arbitrary $\beta > 0$. In this section we want to prove a similar result for the $d-$dimensional lattice, $d \geq 2$, for sufficiently large β.

We do this by proving a more general result for Boltzmann measure on $(\mathbb{S}^1)^\Lambda$ where Λ is an arbitrary finite index set. In this section $\mathbb{S}^1$ is the circle as a subset of $\mathbb{R}^2$ unless stated otherwise. Let us fix some notation: ν denotes the normalised Lebesgue on $\mathbb{S}^1$ and $\nu_\Lambda = \nu^{\otimes \Lambda}$. $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^n$. For $h \in \mathbb{R}^2$ we let

$$\nu^h(d\sigma) = \frac{1}{Z(h)} e^{\langle h, \sigma \rangle} \nu(d\sigma) \tag{5.19}$$

be the probability measure on $\mathbb{S}^1$ with external field h, $Z(h)$ is a normalising constant.

For a symmetric matrix $M \in \mathbb{R}^{\Lambda \times \Lambda}$ we let

$$\mu_M(d\sigma) = \frac{1}{Z_M} e^{-\frac{1}{2}\langle \sigma, M\sigma \rangle} \nu_\Lambda(d\sigma) \quad \text{with} \quad \langle \sigma, M\sigma \rangle = \sum_{i,j \in \Lambda} M_{ij} \langle \sigma_i, \sigma_j \rangle, \tag{5.20}$$

where Z_M is chosen such that μ_M is a probability measure on $(\mathbb{S}^1)^\Lambda$. Observe that by choosing M and Λ appropriately we can retrieve the $XY-$model with free boundary conditions from μ_M. Let $\|M\| = |\lambda^+ - \lambda^-|$ denote the difference between the largest and smallest eigenvalues of M, λ^+ and λ^- respectively.

Before we can state the theorem concerning μ_M, we need to state a result about the ν^h: they satisfy a logarithmic Sobolev inequality uniformly in h.

Proposition 5.6: *There exists a constant $\alpha_0 < \infty$ such that for any $h \in \mathbb{R}^2$, we have*

$$Ent_{\nu^h}(f^2) \leq 2\alpha_0 \nu^h\left[|\nabla f|^2\right],$$

for any differentiable f.

We postpone the proof of this proposition to the end of this section. The main goal of this section is the following theorem:

Theorem 5.7 ([5])**:** *Assume $\|M\| < 1$. Then μ_M satisfies a logarithmic Sobolev inequality of the form:*

$$Ent_{\mu_M}(f^2) \leq 2\alpha_0 \left(1 + \frac{2\|M\|}{1 - \|M\|}\right) \mu_M\left[|\nabla_\Lambda f|^2\right],$$

for any differentiable f.

Before turning to the proof of the above theorem, let us quickly describe how to obtain a logarithmic Sobolev inequality for the Gibbs measure of the XY-model at high temperature as a corollary:

Theorem 5.8: *For $d \geq 2$ and $\beta < \min\{\frac{1}{4d}, \beta_c(d)\}$ the unique Gibbs measure π of the XY-model satisfies a LSI:*

$$Ent_\pi(f^2) \leq 2\alpha_0 \frac{1 + 4d\beta}{1 - 4d\beta} \pi\left[|\nabla f|^2\right],$$

where f is any local and differentiable function. α_0 is given by Proposition 5.6 and $\beta_c(d)$ by Theorem 2.8.

Proof. Due to our assumptions on β Theorem 2.8 applies and hence π is unique. We can make use of this uniqueness: by inspecting the proof of Theorem 2.7 we get that also the finite volume measures with free boundary conditions converge weakly to π. This is the following sequence of measures:

$$\mu_k(d\omega) = \frac{1}{Z_k} e^{-H_k(\omega)} \nu_{\{-k,\dots,k\}^d}(d\omega),$$

where $H_k(\omega) = -\sum_{|i-j|=1} \beta\langle\sigma_i,\sigma_j\rangle$. This measure is of the form given in (5.20) with $\Lambda = \{-k,...,k\}^d$ and

$$M_{ij} = \beta \mathbb{1}_{|i-j|=1}.$$

A basic estimate for the eigenvalues of M yields $|\lambda| \le 2d\beta$ for any eigenvalue λ and hence $\|M\| \le 4d\beta$. Thus if we choose $\beta < \frac{1}{4d}$ we can apply Theorem 5.7.

Let f be local and differentiable and choose K such that $\Lambda(f) \subset \{-K,...,K\}^d$. Theorem 5.7 combined with the dominated convergence theorem yields:

$$\begin{aligned}\mathrm{Ent}_\pi(f^2) &= \lim_{\substack{k\to\infty\\ k\ge K}} \mathrm{Ent}_{\mu_k}(f^2) &(5.21)\\ &\le \lim_{\substack{k\to\infty\\ k\ge K}} 2\alpha_0 \frac{1+4d\beta}{1-4d\beta}\mu_k\big[|\nabla_{\Lambda(f)}|^2\big] = 2\alpha_0\frac{1+4d\beta}{1-4d\beta}\pi\big[|\nabla f|^2\big] &(5.22)\end{aligned}$$

□

Let us turn to the proof of Theorem 5.7, we follow the proof of [5]. We split their proof into two parts. First, we present a lemma in which we prove a logarithmic Sobolev inequality for an auxiliary measure:

Lemma 5.9: *Assume additionally M to be positive definite and assume we have $M^{-1} = c^{-1}\mathbb{1} + B^{-1}$ where $c > 0$, $\mathbb{1}$ the identity matrix and B a symmetric, positive definite matrix. For $y \in \mathbb{R}^2$ let*

$$V(y) = -\log \int e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma) \quad \text{and} \quad \mu^r(dx) = \frac{1}{Z_r} e^{-\frac{1}{2}\langle x,Bx\rangle - \sum_{i\in\Lambda} V(x_i)}dx,$$

where Z_r is chosen in such a way that μ^r is a probability measure on $\mathbb{R}^{2\Lambda}$. Then μ^r satisfies a logarithmic Sobolev inequality with constant $\alpha_r = (c-c^2)^{-1}$:

$$Ent_{\mu^r}(f^2) \le 2\alpha_r \mu^r\big[|\nabla f|^2\big]$$

Proof. To prove this lemma we check Theorem 4.24. Because B is positive definite and due to the product structure of $\exp\big(-\sum_{i\in\Lambda} V(x_i)\big)$ it suffices to check that

$$\langle y, \mathrm{Hess}(V)y\rangle \ge \frac{1}{\alpha_r}\mathbb{1} \tag{5.23}$$

as quadratic form for any $y \in \mathbb{R}^2$. We therefore compute the second derivatives of $V(y)$:

$$\begin{aligned}\frac{d^2V}{dy_1dy_2} &= -\frac{c^2\int(y_1-\sigma_1)(y_2-\sigma_2)e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)}{\int e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)}\\ &= -\frac{c^2\int(y_1-\sigma_1)(y_2-\sigma_2)e^{c(y_1\sigma_1+y_2\sigma_2)}\nu(d\sigma)}{\int e^{c(y_1\sigma_1+y_2\sigma_2)}\nu(d\sigma)}\\ &= -c^2\nu^{cy}\big[(y_1-\sigma_1)(y_2-\sigma_2)\big]\\ \frac{d^2V}{dy_1^2} &= \frac{c\int e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)}{\int e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)} - \frac{c^2\int(y_1-\sigma_1)^2e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)}{\int e^{-\frac{c}{2}|y-\sigma|^2}\nu(d\sigma)}\\ &= c - c^2\nu^{cy}\big[(y_1-\sigma_1)^2\big]\end{aligned}$$

And analogously $\frac{d^2V}{dy_2^2} = c - c^2\nu^{cy}\big[(y_2-\sigma_2)^2\big]$. Observe:

$$\langle y, \mathrm{Hess}(V)y\rangle = c|y|^2 - c^2\mathrm{Var}_{\nu^{cy}}(\langle y,\sigma\rangle) \ge (c-c^2)|y|^2 \tag{5.24}$$

where we used the bound $\mathrm{Var}_{\nu^{cy}}(\langle y,\sigma\rangle) \le |y|$ which follows from the fact that $|\sigma|^2 = 1$. This proves (5.23) with $\alpha_r^{-1} = c - c^2$ and thus completes the proof of the lemma. □

Proof of Theorem 5.7. We continue following [5].

The key to the proof is expressing μ_M as a combination of ν^h for some appropriate h and μ^r to use their respective logarithmic Sobolev inequalities.

But before we do that, let us observe that the measure μ_M does not change when we replace M by $M + \delta\mathbb{1}$ where $\mathbb{1}$ denotes the identity matrix: this only changes the normalising constant Z_M. Hence we assume that M is in fact a positive definite matrix which allows us to use Lemma 5.9 later.

Using that M is a positive definite matrix, we can find another positive definite matrix B such that $M^{-1} = c^{-1}\mathbb{1} + B^{-1}$ for any $c < \|M\|$. With this decomposition of M we can express the law of a multivariate Gaussian random variable with covariance M^{-1} as the convolution of two Gaussians with covariance B^{-1} and $c^{-1}\mathbb{1}$ respectively. On the level of densities this reads

$$e^{-\frac{1}{2}\langle\varphi,M\varphi\rangle} = C\int_{\mathbb{R}^{2\Lambda}} e^{-\frac{1}{2}c\langle x-\varphi,x-\varphi\rangle}e^{-\frac{1}{2}\langle x,Bx\rangle}dx, \tag{5.25}$$

where C is a normalising constant which we do not need to specify. This equation allows us to relate μ_M with ν_h and μ^r. To see this, we write $\nu_x = \bigotimes_{i\in\Lambda}\nu^{x_i}$ for $x \in \mathbb{R}^{2\Lambda}$. We then have $\mu_M\big[f(\sigma)\big] = \mu^r\big[\nu_{cx}[f(\sigma)]\big]$ in the following sense:

$$\int f(\sigma)\mu_M(d\sigma) = \iint f(\sigma)\nu_{cx}(d\sigma)\mu^r(dx) \tag{5.26}$$

To see why this is true, we first observe:

$$e^{V(\psi)}e^{-\frac{1}{2}c|\psi-\sigma|}\nu(d\sigma) \overset{|\sigma|^2=1}{=} \underbrace{e^{V(\psi)}e^{-\frac{c}{2}|\psi|^2}}_{=Z(c\psi)}e^{c\langle\psi,\sigma\rangle}\nu(d\sigma) = \nu^{c\psi}(d\sigma)$$

Next, we write $\propto$ if both sides of the equation are equal up to some normalising constant:

$$\begin{aligned}\int f(\sigma)\mu_M(d\sigma) &\propto \int f(\sigma)e^{-\frac{1}{2}\langle\sigma,M\sigma\rangle}\nu_\Lambda(d\sigma)\\ &\overset{(5.25)}{\propto} \iint f(\sigma)e^{-\frac{1}{2}c\langle x-\sigma,x-\sigma\rangle}e^{-\frac{1}{2}\langle x,Bx\rangle}\prod_{i\in\Lambda}\big(e^{V(x_i)}e^{-V(x_i)}\big)dx\ \nu_\Lambda(d\sigma)\\ &\propto \iint f(\sigma)\nu_{cx}(d\sigma)\mu^r(dx)\end{aligned}$$

This shows (5.26).

We are now in a position to apply Proposition 5.6 and Lemma 5.9. Note that by Lemma 5.9 and the tensoration property, Theorem 4.5, each ν_{cx} satisfies a logarithmic Sobolev inequality with constant α_0 as well, uniformly in x. Let $f\colon (\mathbb{S}^1)^\Lambda \to \mathbb{R}$ be differentiable and abbreviate $g(x) = \nu_{cx}\big[f(\sigma)^2\big]^{1/2}$, we then have:

$$\begin{aligned}\mathrm{Ent}_{\mu_M}(f^2) &= \mu^r\big[\nu_{cx}\big[f^2\log f^2\big]\big] - \mu^r\big[\nu_{cx}[f^2]\log\nu_{cx}[f^2]\big]\\ &\quad + \mu^r\big[\nu_{cx}[f^2]\log\nu_{cx}[f^2]\big] - \mu^r\big[\nu_{cx}[f^2]\big]\log\mu^r\big[\nu_{cx}[f^2]\big]\\ &= \mu^r\big[\mathrm{Ent}_{\nu_{cx}}(f^2)\big] + \mathrm{Ent}_{\mu^r}(g^2)\\ &\le \mu^r\big[2\alpha_0\nu_{cx}\big[|\nabla_\Lambda f|^2\big]\big] + 2\alpha_r\mu_r\big[|\nabla_{\mathbb{R}^{2\Lambda}}g|^2\big]\\ &= 2\alpha_0\mu_M\big[|\nabla_\Lambda f|^2\big] + 2\alpha_r\mu_r\big[|\nabla_{\mathbb{R}^{2\Lambda}}g|^2\big]\end{aligned} \tag{5.27}$$

where we used the logarithmic Sobolev inequalities for μ^r and ν_{cx} (uniformly in x) and (5.26).

The only term left to estimate is $\mu_r\big[|\nabla_{\mathbb{R}^{2\Lambda}}g|^2\big]$ for which we proceed in a very similar way to Lemma 5.4. Our goal is to derive the following bound:

$$\mu_r\big[|\nabla_{\mathbb{R}^{2\Lambda}}g|^2\big] \le 2c^2\alpha_0\mu^r\big[\nu_{cx}\big[|\nabla_\Lambda f|^2\big]\big] \overset{(5.26)}{=} 2c^2\alpha_0\mu_M\big[|\nabla_\Lambda f|^2\big] \tag{5.28}$$

In the following, let ∇_i denote the $\mathbb{R}^2$-gradient acting on x_i for some $i \in \Lambda$. We then have:

$$\begin{aligned}\nabla_i g(x) &= \frac{\nabla_i g(x)^2}{2g(x)} = \frac{\nabla_i \nu_{cx}[f^2]}{2\nu_{cx}[f^2]^{1/2}} = \frac{1}{2\nu_{cx}[f^2]^{1/2}} \nabla_i \frac{\int f^2(\sigma) e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)}{\int e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)} \\ &= \frac{1}{2\nu_{cx}[f^2]^{1/2}}\Bigg(\frac{\int f^2(\sigma)\cdot c\sigma_i e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)}{\int e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)} \\ &\qquad - \frac{\int c\sigma_i e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)\int f^2(\sigma)e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)}{\left(\int e^{c\sum_{j\in\Lambda}\langle x_j,\sigma_j\rangle}\nu_\Lambda(d\sigma)\right)^2}\Bigg) \\ &= \frac{c}{2\nu_{cx}[f^2]^{1/2}}\mathrm{Cov}_{\nu_{cx}}\big(f^2(\sigma),\sigma_i\big)\end{aligned}$$

We bound the norm of the covariance. Fix the values of σ_j for $j \in \Lambda\backslash\{i\}$ and consider first $\mathrm{Cov}_{\nu^{cx_i}}\big(f^2(\sigma),\sigma_i\big)$. We use a similar approach as in (5.6), namely we duplicate the measure ν^{cx_i}:

$$\begin{aligned}\left|\mathrm{Cov}_{\nu^{cx_i}}\big(f^2(\sigma),\sigma_i\big)\right| &= \frac{1}{2}\left|(\nu^{cx_i}\otimes\nu^{cx_i})\big[(f(\sigma_i)-f(\tilde\sigma_i))(f(\sigma_i)+f(\tilde\sigma_i))(\sigma_i-\tilde\sigma_i)\big]\right| \\ &\le \mathrm{Var}_{\nu^{cx_i}}(f)^{1/2}\left(\frac{1}{2}(\nu^{cx_i}\otimes\nu^{cx_i})\big[(f(\sigma)+f(\tilde\sigma))^2|\sigma_i-\tilde\sigma_i|^2\big]\right)^{1/2} \\ &\le \mathrm{Var}_{\nu^{cx_i}}(f)^{1/2}\big(8\nu^{cx_i}[f^2(\sigma)]\big)^{1/2},\end{aligned}$$

where we used the Cauchy-Schwarz inequality and afterwards that $|\sigma-\tilde\sigma|\le 2$ as well as $(a+b)^2 \le 2a^2+2b^2$. To got back to $\mathrm{Cov}_{\nu_{cx}}$ we use that $\nu_{cx}[\cdot] = \nu_{cx}[\nu^{cx_i}[\cdot]]$ as well as the fact that under ν_{cx} all the $\{\sigma_j\}_{j\in\Lambda}$ are independent. We obtain:

$$\left|\mathrm{Cov}_{\nu_{cx}}\big(f^2(\sigma),\sigma_i\big)\right| \le 8\nu_{cx}\big[\mathrm{Var}_{\nu^{cx_i}}(f)^{1/2}\nu^{cx_i}[f^2(\sigma)]^{1/2}\big]^2 \le 8\nu_{cx}\big[\mathrm{Var}_{\nu^{cx_i}}(f)\big]\nu_{cx}\big[f^2\big]$$

where we used the Cauchy-Schwarz inequality again. Lastly we use the spectral gap inequality for ν^{cx_i} which is satisfied with constant α_0 according to Proposition 5.6 and Theorem 4.7:

$$\nu_{cx}\big[\mathrm{Var}_{\nu^{cx_i}}(f)\big] \le \alpha_0\nu_{cx}\big[\nu^{cx_i}\big[|\nabla_{\sigma_i}f|^2\big]\big] = \alpha_0\nu_{cx}\big[|\nabla_{\sigma_i}f|^2\big]$$

Collecting all our our estimates we obtain:

$$|\nabla_i g|^2 \le 2c^2\alpha_0\nu_{cx}\big[|\nabla_{\sigma_i}f|^2\big]$$

Summing over $i\in\Lambda$ and taking μ^r-expectation yields (5.28).

To complete the proof, we combine (5.27) and (5.28):

$$\begin{aligned}\mathrm{Ent}_{\mu_M}(f^2) &\le 2\alpha_0\mu_M\big[|\nabla_\Lambda f|^2\big] + 2\alpha_r\mu_r\big[|\nabla_{\mathbb{R}^{2\Lambda}}g|^2\big] \\ &\le 2\alpha_0\big(1+2c^2\alpha_r\big)\mu_M\big[|\nabla_\Lambda f|^2\big]\end{aligned}$$

Lastly we use that by Lemma 5.9 we have $\alpha_r = (c-c^2)^{-1}$. Further, we chose c in such a way that $\|M\| > c$ and by letting $c\to\|M\|$ we obtain the statement of the theorem. □

The proof of Proposition 5.6

To conclude this section and in particular the proof of Theorem 5.8, we need to show Proposition 5.6. Before we start discussing its proof, we rephrase it slightly by perceiving the circle[4] as ${}^{[-\pi,\pi]}/_{-\pi\sim\pi}$ and perceiving the measure $\{\nu^h, h\in\mathbb{R}^2\}$ as measures defined on $[-\pi,\pi]$. Furthermore, due to the rotational invariance of ν in the definition of ν^h, we can restrict ourselves to $h=\beta(1,0)^T$ for $\beta\ge 0$.

[4]This is slightly different compared to the rest of this thesis where we perceive the circle as ${}^{[0,1]}/_{0\sim 1}$. We do this so that the median of the measures becomes 0 which makes the technical estimates more readable.

Proposition 5.10: *Let $\beta \geq 0$, define a measure on $[-\pi, \pi]$ by*

$$\nu^\beta(dx) = \frac{1}{Z(\beta)} e^{\beta \cos(x)} dx,$$

where dx is the Lebesgue measure and $Z(\beta) = \int_{-\pi}^{\pi} e^{\beta\cos(x)}dx$ is a normalising constant. Then there exists a constant $\alpha_0 < \infty$ such that

$$Ent_{\nu^\beta}(f^2) \leq 2\alpha_0 \nu^\beta\big[(f')^2\big],$$

for any $\beta \geq 0$ and any differentiable f.

We discuss the proof of this proposition instead of Proposition 5.6. The constants α_0 in two propositions differ only by a constant factor.

At first it may seem that the proof could be very short and simply an application of the Bakry-Emery criterion. This is unfortunately not the case as $\inf_{x\in[0,1]} \partial_x^2 \beta\cos(2\pi x) = -4\pi^2\beta$ is not uniformly bounded from below in β. Hence we cannot apply Theorem 4.24.

We need to use a different approach. Fortunately, the instances in which probability measures on $\mathbb{R}$ which satisfy logarithmic Sobolev inequalities are well characterised. Here is a version of [6, Thm. 3]:

Theorem 5.11: *Let $\mu(dx) = \rho(x)dx$ be an absolutely continuous probability measure on $\mathbb{R}$. Let m be a median on μ, i.e. $\mu((-\infty, m]) = 1/2$. Then μ satisfies the logarithmic Sobolev inequality*

$$Ent_\mu(f^2) \leq 4\max\{B_+, B_-\}\mu\big[(f')^2\big],$$

where f is smooth and B_+, B_- are given by

$$B_+ = \sup_{x>m} \mu([x,\infty)) \int_m^x \frac{1}{\rho(y)} dy \log\left(1 + \frac{e^2}{\mu([x,\infty))}\right),$$
$$B_- = \sup_{x<m} \mu((-\infty,x]) \int_x^m \frac{1}{\rho(y)} dy \log\left(1 + \frac{e^2}{\mu((-\infty,x])}\right),$$

provided that they are finite.

Remark 5.12: [6, Thm. 3] also gives a lower bound for the optimal logarithmic Sobolev constant with expressions very similar to B_+ and B_-.

This theorem is an improvement over a theorem by [7] which is presented in [1, Chapter 6]. The proof is quite long and technical at times so we only discuss some of the methods involved:

Discussion of the proof of Theorem 5.11. The approach of [1, Chapter 6] consists of three steps: reducing the logarithmic Sobolev inequality to some Orlicz norm which then can be reduced to some Sobolev inequality which in turn is proven by a Hardy inequality. We describe the rough sequence of steps without proofs.

For an appropriate positive function $\phi : \mathbb{R} \to [0, \infty]$ we define the Orlicz space associated to ϕ

$$L^\phi(\mu) = \left\{ f\colon \mathbb{R} \to \mathbb{R} : \exists a > 0 \text{ with } \int \phi(af)d\mu < \infty \right\}$$

and

$$\|f\|_\phi = \sup\left\{ \int |fg|d\mu; g\colon \mathbb{R}\to\mathbb{R}, \int \phi(g)d\mu \leq 1 \right\}.$$

Under appropriate assumptions on ϕ, $\|f\|_\phi$ is a norm and renders $L^\phi(\mu)$ a Banach space. This is for example the case for

$$\Phi(x) = |x|\log(1+|x|) \quad \text{and} \quad \Theta(x) = x^2\log(1+x^2).$$

The norm $\|\cdot\|_\Theta$ is then used to bound $\mathrm{Ent}_\mu(f^2)$ as followed, [1, Lemma 6.3.1]. In the following, f is smooth and sufficiently integrable:

$$\mathrm{Ent}_\mu(f^2) \leq \sup_{a\in\mathbb{R}} \mathrm{Ent}_\mu((f+a)^2) \leq \frac{5}{2}\|f-\mu[f]\|_\Theta^2$$

Further comparing the norms of $\|\cdot\|_\Theta$ and $\|\cdot\|_\Phi$ they arrive at a criterion - [1, Prop. 6.3.2] - for μ to satisfy a logarithmic Sobolev inequality with constant $180d$:

$$\|f_-^2\|_\Phi + \|f_+^2\|_\Phi \leq d\int (f')^2 d\mu$$

for $f_+ = f\mathbb{1}_{[m,\infty)}$ and $f_- = f\mathbb{1}_{(-\infty,m]}$ where m is a median of μ. This inequality is then proven by use of the Hardy inequality [1, Thm. 6.2.1],

$$\int_0^\infty \left(\int_0^x f(t)dt\right)^2 \mu(dx) \leq 4B\int_0^\infty f(x)^2\mu(dx), \tag{5.29}$$

where $B = \sup_{x\geq 0}\mu([x,\infty))\int_0^x \frac{1}{\rho(x)}dx$ under the additional assumption that ρ is strictly positive. The inequality (5.29) essentially follows from applying the Cauchy-Schwarz inequality to

$$\int_0^x f(t)dt = \int_0^x f(t)\frac{g(t)}{g(t)}dt,$$

where $g(t) = \left(\int_0^t \frac{1}{\rho(s)}ds\right)^{1/2}$ and integrating with respect to $\mu(dx)$ afterwards.

This completes our discussion of the proof of Theorem 5.11. □

So let us see how to deduce Proposition 5.10 from Theorem 5.11. Proposition 5.6, which we recall is equivalent to Proposition 5.10, is actually very close to [19, Thm. 2.1]. The notable difference is that their theorem concerns itself with measures on $(\mathbb{S}^n, n\geq 2)$, where as we are interested in measures on $\mathbb{S}^1$. Nevertheless, our proof is similar and leaves out the step which is needed to reduce the measures on $\mathbb{S}^n$ to measures on $\mathbb{R}$. We establish similar estimates as [18, Thm 5.1] where this proposition can also be found.

Proof of Proposition 5.10. Thanks to Theorem 5.11 the only thing we need to do is to check that B_+ and B_- are uniformly bounded in β. In our setting we have

$$\rho(x) = \frac{e^{\beta\cos(x)}\mathbb{1}_{[-\pi,\pi]}(x)}{\int_{-\pi}^{\pi} e^{\beta\cos(y)}dy}$$

and because ρ is symmetric around 0 we have $m=0$ for any β. This symmetry of ρ also implies that $B_+ = B_-$ which is why we restrict ourselves to B_+.

Further we can restrict ourselves to $\beta \geq 1$. The logarithmic Sobolev inequality for $\{\nu^\beta, 0\leq \beta < 1\}$ follows from Theorem 4.4 with the constant worsening by a factor of at most e^2. Let us write out B_+:

$$B_+ = \sup_{x\in[0,\pi]}\int_x^\pi e^{\beta\cos(t)}dt\int_0^x e^{-\beta\cos(t)}dt\log\left(1+\frac{e^2 Z(\beta)}{\int_x^\pi e^{\beta\cos(t)}dt}\right)$$

Let us start with the technical estimates. First we observe two basic estimates:

$$2\sqrt{r} \leq \int_0^r t^{-1/2}e^t dt \leq 2e^r\min\{1,\sqrt{r}\}; \qquad \int_r^\infty t^{-1/2}e^{-t}dt \leq 2e^{-r} \tag{5.30}$$

for any $r>0$. These follow from bounding either $t^{-1/2}$ or e^{-t} by 1 from above or e^t by 1 from below.

The first estimate we need is:

$$Z(\beta) = \int_{-\pi}^{\pi} e^{\beta\cos(t)}dt \le 2\pi e^{\beta} \tag{5.31}$$

Next, we turn to $\int_0^x e^{-\beta\cos(t)}dt$. Assume that $0 \le x \le \pi/2$. We substitute $s = \cos(t)$:

$$\int_0^x e^{-\beta\cos(t)}dt = \int_{\cos(x)}^1 e^{-\beta s}\frac{1}{\sqrt{1-s^2}}ds \le \int_{\cos(x)}^1 e^{-\beta s}\frac{1}{\sqrt{1-s}}ds$$

because $1 - s^2 \ge 1 - s > 0$ on $(0,1]$. Substitute $u = \beta(1-s)$:

$$\begin{aligned}\int_{\cos(x)}^1 e^{-\beta s}\frac{1}{\sqrt{1-s}}ds = \frac{e^{-\beta}}{\sqrt{\beta}}\int_0^{\beta(1-\cos(x))} u^{-1/2}e^{u}du &\le \frac{2}{\sqrt{\beta}}e^{-\beta\cos(x)}\min\left\{1, \sqrt{\beta(1-\cos(x))}\right\} \\ &\le \frac{2}{\sqrt{\beta}}e^{-\beta\cos(x)}\min\left\{1, \sqrt{\beta(1+\cos(x))}\right\}\end{aligned} \tag{5.32}$$

where we used (5.30). For $\pi/2 \le x \le \pi$ we observe:

$$\int_0^x e^{-\beta\cos(t)}dt = \int_0^{\pi/2} e^{-\beta\cos(t)}dt + \int_{\pi/2}^x e^{-\beta\cos(t)}dt \le \frac{2}{\sqrt{\beta}} + \int_{\pi/2}^x e^{-\beta\cos(t)}dt$$

according to the previous estimate. With similar substitutions as before we obtain:

$$\begin{aligned}\int_{\pi/2}^x e^{-\beta\cos(t)}dt &= \int_0^{\cos(x)} e^{-\beta s}\frac{1}{\sqrt{1-s^2}}ds \le \int_0^{\cos(x)} e^{-\beta s}\frac{1}{\sqrt{1-s}}ds \\ &= \frac{e^{-\beta}}{\sqrt{\beta}}\int_{\beta(1-\cos(x))}^{\beta} u^{-1/2}e^{-u}du \le \frac{2}{\sqrt{\beta}}e^{-\beta\cos(x)}\end{aligned}$$

where we used (5.30) for the last bound. In total we obtain

$$\int_0^x e^{-\beta\cos(t)} \le \frac{2}{\sqrt{\beta}}\left(1 + e^{-\beta\cos(x)}\right)dt \le \frac{4}{\sqrt{\beta}}e^{-\beta\cos(x)} \tag{5.33}$$

for $\pi/2 \le x \le \pi$. Analogously, we obtain the following bounds for $\int_x^\pi e^{\beta\cos(t)}dt$:

$$\begin{cases}\int_x^\pi e^{\beta\cos(t)}dt \le \frac{4}{\sqrt{\beta}}e^{\beta\cos(x)} & \text{for} \quad 0 \le x \le \pi/2 \\ \int_x^\pi e^{\beta\cos(t)}dt \le \frac{2}{\sqrt{\beta}}e^{\beta\cos(x)}\min\left\{1, \sqrt{\beta(1+\cos(x))}\right\} & \text{for} \quad \pi/2 \le x \le \pi\end{cases} \tag{5.34}$$

Lastly, we also need a lower bound for $\int_x^\pi e^{\beta\cos(t)}dt$. We first substitute $s = \cos(t)$, $0 \le x \le \pi$:

$$\int_x^\pi e^{\beta\cos(t)}dt = \int_{-1}^{\cos(x)} e^{\beta s}\frac{1}{\sqrt{1-s^2}}ds \ge \frac{1}{\sqrt{2}}\int_{-1}^{\cos(x)} e^{\beta s}\frac{1}{\sqrt{1+s}}ds$$

where we used that $1 - s^2 \le 2 + 2s$ on $[-1,1]$. We substitute $u = \beta(1+y)$ and use (5.30):

$$\frac{1}{\sqrt{2}}\int_{-1}^{\cos(x)} e^{\beta s}\frac{1}{\sqrt{1+s}}ds = \frac{e^{-\beta}}{\sqrt{2\beta}}\int_0^{\beta(1+\cos(x))} u^{-1/2}e^{u}du \ge \sqrt{2}e^{-\beta}\sqrt{1+\cos(x)}$$

In total we have the estimate, $0 \le x \le \pi$:

$$\int_x^\pi e^{\beta\cos(t)}dt \ge \sqrt{2}e^{-\beta}\sqrt{1+\cos(x)} \tag{5.35}$$

This concludes the technical estimates and we can bound B_+. We use the estimates (5.31) - (5.35):

$$
\begin{aligned}
B_+ &= \sup_{x\in[0,\pi]} \int_x^\pi e^{\beta\cos(t)}dt \int_0^x e^{-\beta\cos(t)}dt \log\left(1+\frac{e^2 Z(\beta)}{\int_x^\pi e^{\beta\cos(t)}dt}\right)\\
&\le \sup_{x\in[0,\pi]} \frac{8}{\beta}\min\left\{1, \sqrt{\beta(1+\cos(x))}\right\}\log\left(1+\frac{e^2 2\pi e^\beta}{\sqrt{2}e^{-\beta}\sqrt{1+\cos(x)}}\right)\\
&\le \sup_{r\in[0,\sqrt{2}]} \frac{8}{\beta}\min\left\{1,\sqrt{\beta}r\right\}\log\left(1+\frac{\sqrt{2}e^2\pi e^{2\beta}}{r}\right)\\
&\le \frac{8}{\beta}\sup_{r\in[0,\sqrt{2}]}\min\left\{1,\sqrt{\beta}r\right\}\log\left(\frac{e^4 e^{2\beta}}{r}\right)\\
&= \frac{32}{\beta}+16+\frac{8}{\sqrt{\beta}}\sup_{r\in[0,\sqrt{2}]} r\log(r^{-1})\\
&\le 52,
\end{aligned}
$$

where we used $\beta \ge 1$ and $\sup_{r\in[0,\sqrt{2}]} r\log(r^{-1}) = e^{-1} \le 1/2$ for the last inequality.

This completes the proof of Proposition 5.10 with $\alpha_0 = 52e^2$.

□

5.3 Ergodicity

Having established logarithmic Sobolev inequalities for the XY-model at various temperatures and dimensions, we now turn to the ergodicity of the associated semi-groups. In fact, Theorem 4.7 combined with Theorem 3.2 imply that the semi-group is $L^2(\pi)$-ergodic for any Gibbs measure π. We aim for a stronger statement: in this section we want to show that the semi-group is uniformly ergodic.

Before we can do that, we actually need to define the semi-group in infinite volume and prove some of its properties. Most importantly, we want it to have the exponential approximation property which states that it is well approximated by the semi-groups in finite volume.

Let us introduce the finite volume generators as follows, $\Lambda \Subset \mathbb{Z}^d$

$$\mathcal{L}^{(\Lambda)} = \sum_{i\in\mathbb{Z}^d} \Delta_i - \nabla_i H_\Lambda \cdot \nabla_i$$

and denote the induced semi-group on Ω by $(P_t^{(\Lambda)})_t$. Further for differentiable $f\colon \Omega \to \mathbb{R}$ we define

$$|||f||| = \sum_{i\in\mathbb{Z}^d} \|\nabla_i f\|_\infty .$$

Clearly, all local and differential functions f satisfy $|||f||| < \infty$. Our goal is the following property:

Definition 5.13 (Exponential approximation property)**:** *In our setting a semi-group $(P_t)_t$ is said to have the exponential approximation property if for all $A > 0$ there is a $B \geq 0$ such that*

$$\left\|P_t f - P_t^{(\Lambda)} f\right\|_\infty \leq e^{-At}\, |||\, f|||$$

holds for all local and differentiable functions f as long as $\Lambda(f) \subset \Lambda$ and $dist(\Lambda(f), \Lambda^c) \geq Bt$.

But we have yet to specify how $P_t f$ is defined. For a local and differentiable function we define

$$P_t f = \lim_{\Lambda\uparrow\mathbb{Z}^d} P_t^{(\Lambda)} f, \tag{5.36}$$

if the limit exists in the supremum norm. Before we prove that this is always the case, we need a lemma:

Lemma 5.14: *For $f\colon \Omega \to \mathbb{R}$ local and smooth, $\Lambda \Subset \mathbb{Z}^d$,$i \in \mathbb{Z}^d$ and any $t \geq 0$, we have the estimate*

$$\left\|\nabla_i P_t^{(\Lambda)} f\right\|_\infty \leq \frac{(Dt)^{N_i}}{N_i!} e^{Dt}\, |||\, f|||,$$

where $N_i = dist(i, \Lambda(f))$ and $D < \infty$ is a constant which depends only on β.

In particular, this lemma implies that $\|\nabla_i P_t^{(\Lambda)} f\|_\infty \to 0$ if $dist(i, \Lambda(f)) \to \infty$. This lemma is similar to part of the proof of [15, Thm. 8.2]. Their theorem deals with semi-groups on $\{\pm 1\}^{\mathbb{Z}^d}$ and the discrete derivative, we adapt their approach to our setting.

Proof. Let f be local and smooth, the statement for differentiable f follows by a density argument. First, we observe:

$$\nabla_i P_t^{(\Lambda)} f - P_t^{(\Lambda)} \nabla_i f = \int_0^t \frac{d}{ds} P_{t-s}^{(\Lambda)} \nabla_i P_s^{(\Lambda)} f ds = \int_0^t \frac{d}{ds} P_{t-s}^{(\Lambda)} \left[\nabla_i, \mathcal{L}^{(\Lambda)}\right] P_s^{(\Lambda)} f ds,$$

where $\left[\nabla_i, \mathcal{L}^{(\Lambda)}\right] = \nabla_i \mathcal{L}^{(\Lambda)} - \mathcal{L}^{(\Lambda)} \nabla_i$ is the commutator. Rearranging the above equation yields:

$$\nabla_i P_t^{(\Lambda)} f = P_t^{(\Lambda)} \nabla_i f + \int_0^t \frac{d}{ds} P_{t-s}^{(\Lambda)} \left[\nabla_i, \mathcal{L}^{(\Lambda)}\right] P_s^{(\Lambda)} f ds \tag{5.37}$$

Later, we want to use (5.37) recursively, hence we need to estimate $\|[\nabla_i, \mathcal{L}^{(\Lambda)}]F\|_\infty$ for any smooth F. Write $V_{j,k}$ for the interaction of the spins at j and k, note that $V_{j,k} \equiv 0$ unless $|j-k| = 1$ in the $XY-$model. An elementary calculation yields:

$$\|[\nabla_i, \mathcal{L}^{(\Lambda)}]F\|_\infty = \Big\| \sum_{j\in\Lambda} \sum_{k:|j-k|=1} (\underbrace{\nabla_j \nabla_i V_{j,k}}_{=0 \text{ if } i\notin\{j,k\}})(\nabla_j F)\Big\|_\infty \leq D \sum_{j\in B(i,1)} \|\nabla_j F\|_\infty,$$

where $B(i,n) = \{j \in \mathbb{Z}^d : dist(i,j) \leq n\}$ is the ball of radius n around i in $\mathbb{Z}^d$ in the path-metric. The constant D is given by $2d \sup \nabla_j \nabla_i V_{j,k} = 8\pi^2 d\beta$ in the XY-model.

We use this estimate to estimate (5.37) - after using that $\|P_s^{(\Lambda)}\|_\infty = 1$ for all s twice:

$$\begin{aligned} \|\nabla_i P_t^{(\Lambda)} f\|_\infty &\leq \|\nabla_i f\|_\infty + \int_0^t \|[\nabla_i, \mathcal{L}^{(\Lambda)}]P_s^{(\Lambda)} f\|_\infty ds \\ &\leq \|\nabla_i f\|_\infty + \int_0^t D \sum_{j\in B(i,1)} \|\nabla_j P_s^{(\Lambda)} f\|_\infty ds \end{aligned} \tag{5.38}$$

Notice that we can plug (5.38) into itself. Iterating this yields:

$$\begin{aligned} \|\nabla_i P_t^{(\Lambda)} f\|_\infty &\leq \|\nabla_i f\|_\infty + \sum_{n=1}^\infty \int_{A_n} D^n \sum_{j\in B(i,n)} \|\nabla_j P_s^{(\Lambda)} f\|_\infty ds_1 ... ds_n \\ &= \sum_{n=0}^\infty \frac{(Dt)^n}{n!} \sum_{j\in B(i,n)} \|\nabla_j f\|_\infty, \end{aligned} \tag{5.39}$$

where $A_n = \{s \in \mathbb{R}^n : 0 \leq s_n \leq ... \leq s_1 \leq t\}$, $|A_n| = \frac{t^n}{n!}$. Because we have $\nabla_j f \equiv 0$ whenever $j \notin \Lambda(f)$, the sum $\sum_{j\in B(i,n)} \|\nabla_j f\|_\infty$ equals zero whenever $B(i,n) \cap \Lambda(f) = \emptyset$. (5.39) thus becomes:

$$\begin{aligned} \|\nabla_i P_t^{(\Lambda)} f\|_\infty &\leq \sum_{n=N_i}^\infty \frac{(Dt)^n}{n!} \sum_{j\in B(i,n)} \|\nabla_j f\|_\infty \\ &\leq \sum_{n=N_i}^\infty \frac{(Dt)^n}{n!} \,|||\, f|||\\ &= \frac{(Dt)^{N_i}}{N_i!} e^{Dt} \,|||\, f|||, \end{aligned}$$

which concludes the proof of the lemma. □

Having established the technical lemma, we can prove the existence of the limit (5.36). The proof also yields the exponential approximation property.

Proposition 5.15: *Let f be local and differentiable, then the limit*

$$P_t f = \lim_{\Lambda\uparrow\mathbb{Z}^d} P_t^{(\Lambda)} f$$

exists along sequences $(\Lambda_n)_n$ for which $\Lambda_n \backslash \Lambda_{n-1}$ is a translation of Λ_0 for all n. Furthermore, $(P_t)_t$ possesses the exponential approximation property.

This theorem is similar to [15, Thm. 8.2] which deals with models on $\{\pm 1\}^{\mathbb{Z}^d}$ and the discrete derivative, we adapt their approach to our setting. See also [10] where more estimates of similar flavor are proven.

Proof. Let f be smooth and local, the statement for differentiable f again follows by a density argument. We want to compare $P_t^{(\Lambda_1)}$ and $P_t^{(\Lambda_2)}$ for two different $\Lambda_i \Subset \mathbb{Z}^d$. Let $\Lambda_1 \subset \Lambda_2 \Subset \mathbb{Z}^d$. We observe:

$$P_t^{(\Lambda_2)} f - P_t^{(\Lambda_1)} f = \int_0^t \frac{d}{ds} P_{t-s}^{(\Lambda_1)} P_s^{(\Lambda_2)} f ds = \int_0^t P_{t-s}^{(\Lambda_1)} \big(\mathcal{L}^{(\Lambda_2)} - \mathcal{L}^{(\Lambda_1)}\big) P_s^{(\Lambda_2)} f ds$$

We have $\big(\mathcal{L}^{(\Lambda_2)} - \mathcal{L}^{(\Lambda_1)}\big) = \sum_{i \in \Lambda_2 \setminus \Lambda_1} \nabla_i H_{\Lambda_2} \cdot \nabla_i$. Using $\|P_{t-s}^{(\Lambda_1)}\|_\infty = 1$ for all s we obtain:

$$\big\|P_t^{(\Lambda_2)} f - P_t^{(\Lambda_1)} f\big\|_\infty \leq \int_0^t \big\|\big(\mathcal{L}^{(\Lambda_2)} - \mathcal{L}^{(\Lambda_1)}\big) P_s^{(\Lambda_2)} f\big\|_\infty ds \leq C_0 \sum_{i \in \Lambda_2 \setminus \Lambda_1} \int_0^t \big\|\nabla_i P_s^{(\Lambda_2)} f\big\|_\infty ds,$$

where $C_0 = \sup |\nabla_i H_\Lambda| = 4d\pi\beta < \infty$ in the $XY-$model. Here we are in a position to apply Lemma 5.14:

$$\big\|P_t^{(\Lambda_2)} f - P_t^{(\Lambda_1)} f\big\|_\infty \leq C_0 \sum_{i \in \Lambda_2 \setminus \Lambda_1} \int_0^t \frac{(Ds)^{N_i}}{N_i!} e^{Ds} \, |||\, f ||| \tag{5.40}$$

Let $A > 0$ be given from Definition 5.13 and choose B large enough such that $2 - \log B + \log D + \frac{D}{B} \leq -2A$. Assume now that $N_i \geq Bt$ for all $i \in \Lambda_2 \setminus \Lambda_1$. Using the rough estimate $k! \geq k^k e^{-2k}$ for all $k \geq 1$ we obtain for $0 \leq s \leq t$:

$$\frac{(Ds)^{N_i}}{N_i!} e^{Ds} \leq \frac{(Dt)^{N_i}}{N_i!} e^{Dt} \leq \exp\big(N_i(\log Dt - \log N_i + 2) + Dt\big) \leq e^{-At - AN_i}$$

With this estimate (5.40) becomes:

$$\begin{aligned}\big\|P_t^{(\Lambda_2)} f - P_t^{(\Lambda_1)} f\big\|_\infty &\leq C_0 \sum_{i \in \Lambda_2 \setminus \Lambda_1} \int_0^t e^{-At - AN_i} \, |||\, f \,|||\, ds \\ &\leq C_0 t |\Lambda_2 \setminus \Lambda_1| e^{-At - A \cdot \mathrm{dist}(\Lambda_2 \setminus \Lambda_1, \Lambda(f))} \, |||\, f ||| \end{aligned} \tag{5.41}$$

Here it becomes evident that $P_t^{(\Lambda_n)} f$ is a Cauchy sequence in the supremum norm when $(\Lambda_n)_n$ is chosen in such a way that $\Lambda_n \setminus \Lambda_{n-1}$ is a translation of Λ_0. This means in particular that $|\Lambda_n \setminus \Lambda_{n-1}| = |\Lambda_0|$ is constant. Hence the limit $\lim_{n \to \infty} P_t^{(\Lambda_n)} f$ exists. Further, we can see that the limit does not depend on the precise choice of the sequence $(\Lambda_n)_n$.

It remains to show the exponential approximation property, let Λ be given like in Definition 5.13. Without loss of generality we assume that $\Lambda = \Lambda_N$ for some sequence $(\Lambda_n)_n$ like above. We then have:

$$\begin{aligned}\big\|P_t f - P_t^{(\Lambda)} f\big\|_\infty &\leq \sum_{n=N}^{\infty} \big\|P_t^{(\Lambda_{n+1})} f - P_t^{(\Lambda_n)} f\big\|_\infty \\ &\overset{(5.41)}{\leq} e^{-At} \left(C_0 t |\Lambda_0| \sum_{n=N}^{\infty} e^{-A \cdot \mathrm{dist}(\Lambda_{n+1} \setminus \Lambda_n, \Lambda(f))} \right) |||\, f |||\end{aligned}$$

By choosing B larger we implicitly increase N which allows us make $\big(...\big)$ in the inequality above as small as we want, in particular smaller than 1. This completes the proof of the exponential approximation property. □

Now that we know how to relate P_t and $P_t^{(\Lambda)}$ we can prove that a logarithmic Sobolev inequality implies ergodicity in a uniform sense. We need an additional assumption, (5.42), which states that the semi-group in finite volume behaves somewhat reasonably.

Theorem 5.16: *In our setting: Assume $(P_t)_t$ is a semi-group satisfying the exponential approximation property and assume π is a Gibbs measure. We assume that π satisfies a logarithmic Sobolev inequality with constant $\alpha > 0$.*

Further assume that there exists $c > 0$ such that we have for any $\Lambda \Subset \mathbb{Z}^d$ with $|\Lambda| \geq |\partial\Lambda|$ where $\partial\Lambda$ is the set of edges between Λ and Λ^c:

$$\left\|P_1^{(\Lambda)} g\right\|_\infty \leq \sup_\eta e^{c|\Lambda|^2} \mu_\Lambda^\eta[g], \tag{5.42}$$

for any measurable $g \geq 0$ localised in Λ.

Under these assumptions we then have for any local differentiable function f and any $\theta \in (0,1)$ that there exists $C(\theta, \Lambda(f)) < \infty$ and $m \geq \frac{1}{\alpha}$ such that for all $t \geq 0$:

$$\left\|P_t f - \pi[f]\right\|_\infty \leq C(\theta, \Lambda(f)) e^{-\theta m t} \,|||\, f |||$$

Proof. We follow the presentation of [15, Thm. 8.5].

Before we start with the proof, let us observe a small inequality for later use: let $g \geq 0$, $\Lambda \Subset \mathbb{Z}^d$, $\Lambda(g) \subset \Lambda$ and η a boundary condition. We then have, uniformly in η:

$$\begin{aligned}
\pi\left[g\right] &= \int \mu_\Lambda^\omega\left[g\right] \pi(d\omega) = \int \nu_\Lambda\left[e^{-W_\Lambda(\cdot\ \omega_{\Lambda^c})} e^{-U_\Lambda} g\right] \mu(d\omega) \\
&\geq e^{-4\|W_\Lambda\|_\infty} \int \mu_\Lambda^\eta\left[g\right] \pi(d\omega) = e^{-4\|W_\Lambda\|_\infty} \mu_\Lambda^\eta\left[g\right],
\end{aligned} \tag{5.43}$$

where W_Λ is the energy of the interaction between Λ and Λ^c as seen in (5.1). In the case of the XY-model we have $\|W_\Lambda\|_\infty = \beta|\partial\Lambda|$.

Fix f as required and let us now start estimating $\left\|P_t f - \pi[f]\right\|_\infty$. First, we use the triangle inequality to insert $P_t^{(\Lambda)}$ for some Λ which we fix later:

$$\left\|P_t f - \pi[f]\right\|_\infty \leq \left\|P_t^{(\Lambda)} f - \pi[f]\right\|_\infty + \left\|P_t f - P_t^{(\Lambda)} f\right\|_\infty \tag{5.44}$$

The second term is later estimated by the the exponential approximation property. For the first term we proceed as follows, assume $t \geq 1$ and let $q \geq 1$ to be chosen later:

$$\begin{aligned}
\left|P_t^{(\Lambda)} f - \pi[f]\right| &\leq \left|P_1^{(\Lambda)}\left[P_{t-1}^{(\Lambda)} f - \pi[f]\right]\right| \\
&= \left|P_1^{(\Lambda)}\left[\left|P_{t-1}^{(\Lambda)} f - \pi[f]\right|^q\right]\right|^{1/q} \\
&\leq e^{\frac{c|\Lambda|^2}{q}} \mu_\Lambda\left[\left|P_{t-1}^{(\Lambda)} f - \pi[f]\right|^q\right]^{1/q},
\end{aligned}$$

where we first used Jensen's inequality and then our assumption (5.42). By further applying (5.43) to $\left|P_{t-1}^{(\Lambda)} f - \pi[f]\right|^q$ we obtain:

$$\left|P_t^{(\Lambda)} f - \pi[f]\right| \leq e^{\frac{c|\Lambda|^2}{q}} e^{\frac{4\beta|\partial\Lambda|}{q}} \pi\left[\left|P_{t-1}^{(\Lambda)} f - \pi[f]\right|^q\right]^{1/q} = e^{\frac{c|\Lambda|^2}{q} + \frac{4\beta|\partial\Lambda|}{q}} \left\|P_{t-1}^{(\Lambda)} f - \pi[f]\right\|_q$$

Having used our assumption (5.42), we now want to go back from $P_{t-1}^{(\Lambda)}$ to P_{t-1}. We do this the following way:

$$\left\|P_{t-1}^{(\Lambda)} f - \pi[f]\right\|_q \leq \left\|P_{t-1} f - \pi[f]\right\|_q + \left\|P_{t-1}^{(\Lambda)} f - P_{t-1} f\right\|_q \leq \left\|P_{t-1} f - \pi[f]\right\|_q + \left\|P_{t-1}^{(\Lambda)} f - P_{t-1} f\right\|_\infty$$

Again, the last term should become small when we use the exponential approximation property. Before we actually use it, let us briefly collect the estimates we have thus far, namely the one above and (5.44):

$$\left\|P_t f - \pi[f]\right\|_\infty \leq \left\|P_t f - P_t^{(\Lambda)} f\right\|_\infty + e^{\frac{c|\Lambda|^2}{q} + \frac{4\beta|\partial\Lambda|}{q}} \left(\left\|P_{t-1} f - \pi[f]\right\|_q + \left\|P_{t-1}^{(\Lambda)} f - P_{t-1} f\right\|_\infty\right) \tag{5.45}$$

Now we want to use the exponential approximation property. Keep in mind that we want to have an estimate which is valid for all t large, it is evident that we cannot choose the same Λ for all t as the conditions of the exponential approximation property would not be satisfied. Instead we need to choose Λ in dependence of t:

$$\Lambda_t = \{-\lfloor \lambda t \rfloor, ..., \lfloor \lambda t \rfloor\}^d$$

for some parameter $\lambda > 0$. We then have $|\Lambda_t| \leq (4\lambda t)^d$ and $|\partial\Lambda_t| \leq 4d(\lambda t)^{d-1}$. Now let $A > 0$ and let λ be chosen implicitly so that the condition $dist(\Lambda(f), \Lambda_t^c) \geq Bt$ is satisfied for t large enough, this choice of λ depends only on A and $\Lambda(f)$. By the use of the exponential approximation property (5.45) becomes:

$$\left\|P_t f - \pi[f]\right\|_\infty \leq e^{-At} \; ||| \, f \, ||| + \underbrace{e^{\frac{c(4\lambda t)^{2d}}{q} + \frac{16\beta d(\lambda t)^{d-1}}{q}}}_{:=\varphi(t)} \left(\left\|P_{t-1} f - \pi[f]\right\|_q + e^{-A(t-1)} \; ||| \, f \, ||| \right) \quad (5.46)$$

There are two things left to do before we are done: we need to show that $\varphi(t)$ is uniformly bounded in t and that $\left\|P_{t-1} f - \pi[f]\right\|_q$ decays fast enough. Clearly these goals cannot be achieved if we choose a constant q, hence we choose q depending on t. In light of Gross' integration Lemma, Proposition 4.17, we choose

$$q = q\left(\tfrac{t-1}{\theta t}, 2, \alpha\right) = 1 + e^{\frac{2(1-\theta)t-2}{\alpha}},$$

where $\theta \in (0, 1)$. With this choice $\varphi(t) \to 1$ as $t \to \infty$ and hence $\varphi(t)$ is uniformly bounded in t. To estimate $\left\|P_{t-1} f - \pi[f]\right\|_q$ we use Proposition 4.17 which we are allowed to do because π satisfies a logarithmic Sobolev inequality with constant α by assumption:

$$\left\|P_{t-1} f - \pi[f]\right\|_q = \left\|P_{t-1-\theta t}\left[P_{\theta t} f - \pi[f]\right]\right\|_q \leq \left\|P_{\theta t} f - \pi[f]\right\|_2$$

Lastly, we use the fact that π also satisfies a spectral gap inequality with constant α^{-1} combined with the $L^2(\pi)$-ergodicity, Theorem (3.2):

$$\left\|P_{\theta t} f - \pi[f]\right\|_2 \leq e^{-\theta\alpha^{-1} t}\left\|f - \pi[f]\right\|_2 \leq \alpha e^{-\theta\alpha^{-1}} \; ||| \, f |||$$

Looking at (5.46) again, we can see that we have derived the desired estimate for large enough t if we choose A accordingly. The desired estimate holds true for all t by choosing $C(\theta, \Lambda(f))$ big enough. □

Combining everything we have showed so far allows us to deduce:

Theorem 5.17: *The dynamical XY-model is ergodic in the sense of the previous theorem if*

- $d = 1$ *and* $0 < \beta < \infty$
- $d \geq 2$ *and* $0 < \beta < \beta_c(d)$.

Proof. Our proof is based on checking the conditions of Theorem 5.16. The exponential approximation property is satisfied according to Proposition 5.15 and the logarithmic Sobolev inequality holds according to Theorem 5.1 and Theorem 5.8 respectively. The only thing left to show is that (5.42) is satisfied, i.e. we need to show that

$$\left\|P_1^{(\Lambda)} g\right\|_\infty \leq \sup_\eta e^{c|\Lambda|} \mu_\Lambda^\eta[g]$$

is satisfied for some $c > 0$, all $\Lambda \Subset \mathbb{Z}^d$ with $|\Lambda| \geq |\partial\Lambda|$, all $g \geq 0$ and all boundary conditions η. This is the content of the following lemma which then completes the proof of this theorem. □

Lemma 5.18: *The semi-groups of the XY model in finite volume satisfy* (5.42) *as described above.*

Proof. The proof is inspired by the instructions given in [15, Exercise 3.8].

Fix $\Lambda \Subset \mathbb{Z}^d$ with $|\Lambda| \geq |\partial\Lambda|$, $t \geq 0$, a boundary condition η and $g \geq 0$. In this proof we use a probabilistic approach, let X_t denote the process associated to the generator $\mathcal{L} = \Delta_\Lambda - \nabla_\Lambda H^\eta_\Lambda \cdot \nabla_\Lambda$. Equivalently, see [4], X_t is the process satisfying the following stochastic differential equation

$$dX_t = -\nabla_\Lambda H^\eta_\Lambda(X_t)dt + \sqrt{2}\, dB_t,$$

where B_t is a Brownian motion on $[0,1)^\Lambda$ with periodic boundary conditions. It can be defined as $B_t = \lfloor \tilde{B}_t \rfloor = \big(\lfloor \tilde{B}_t^{(i)} \rfloor\big)_{i\in\Lambda}$ where $\tilde{B}_t$ is a Brownian motion on $\mathbb{R}^\Lambda$. Write $\mathbb{P}^x$ for the measure under which $\mathbb{P}^x(X_0 = x) = 1$ and $\mathbb{E}^x$ for the corresponding expectation. Further we define the process

$$Z_t = \exp\bigg(\int_0^t \frac{1}{\sqrt{2}} \nabla_\Lambda H^\eta_\Lambda(X_s)dB_s - \frac{1}{2}\int_0^t \bigg| \frac{1}{\sqrt{2}} \nabla_\Lambda H^\eta_\Lambda(X_s) \bigg|^2 dt \bigg).$$

We have following estimate for Z_t by using that $\|\nabla_\Lambda H^\eta_\Lambda\| \leq 2\pi\beta\big(|\Lambda| + |\partial\Lambda|\big) \leq 4\pi\beta|\Lambda|$:

$$\exp\big(-\sqrt{8}t\beta\pi|\Lambda| - \sqrt{32}t\pi^2\beta^2|\Lambda|^2 \big) \leq Z_t \leq \exp\big(\sqrt{8}t\pi\beta|\Lambda|\big)$$

By Girsanov's theorem and Novikov's criterion (for example [16]) it follows that for any $T < \infty$ $(X_t, 0 \leq t \leq T)$ has the same distribution as $(\sqrt{2}B_t, 0 \leq t \leq T)$ under the measure $\mathbb{Q}^x_T$ which is defined by $\mathbb{Q}^x_T(A) = \mathbb{E}^x[\mathbb{1}_A Z_T]$. This entails $\mathbb{E}^x[g(X_t)Z_t] = \mathbb{E}^x\big[g\big(\sqrt{2}B_t\big)\big]$. We can then deduce following estimates for any $g \geq 0$ and $x \in [0,1)$:

$$\begin{aligned} P^{(\Lambda)}_t g(x) &= \mathbb{E}^x[g(X_t)] = \mathbb{E}^x[g(X_t)Z_tZ_t^{-1}] \\ &\leq e^{-\sqrt{8}t\beta\pi|\Lambda| - \sqrt{32}t\pi^2\beta^2|\Lambda|^2}\ \mathbb{E}^x[g(X_t)Z_t] \\ &\leq e^{\tilde{c}t|\Lambda|^2}\mathbb{E}^x\big[g\big(\sqrt{2}B_t\big)\big], \end{aligned}$$

for some appropriate $\tilde{c} < \infty$ which does not depend on Λ or on t. Next we use the fact that the distribution of $B^{(1)}_t$, the first coordinate of B_t, is absolutely continuous with respect to the Lebesgue measure on $[0,1)$ and uniformly bounded in x. Hence there exists a function c_t such that

$$\mathbb{E}^x\big[g\big(\sqrt{2}B_t\big)\big] \leq c_t^{|\Lambda|}\nu_\Lambda[g],$$

where ν_Λ is the Lebesgue measure on $[0,1)^\Lambda$. Recall that $\mu^\eta_\Lambda(d\omega) = \nu_\Lambda[e^{-H^\eta_\Lambda}]^{-1}e^{-H^\eta_\Lambda}(\omega)\nu_\Lambda(d\omega)$. Combining the previous estimates thus yields:

$$P^{(\Lambda)}_t g(x) \leq e^{\tilde{c}t|\Lambda|^2}c_t^{|\Lambda|}\nu_\Lambda[g] \leq e^{\tilde{c}t|\Lambda|^2}c_t^{|\Lambda|}e^{osc(H^\eta_\Lambda)}\mu^\eta_\Lambda[g] \leq e^{\tilde{c}t|\Lambda|^2}c_t^{|\Lambda|}e^{4\beta|\Lambda|}\mu^\eta_\Lambda[g]$$

where we used Lemma 3.5 for the last inequality to relate the expectations of ν_Λ and μ^η_Λ. Plugging $t = 1$ into the above inequality yields the desired statement for some sufficiently large c.

□

6 Conclusion

In this thesis, I have proven the Theorems 5.1, 5.8 and 5.17 that together show that Glauber dynamics converges uniformly to the Gibbs measure. The main conditions required for this strong result are the compactness of the circle for the spin degrees of freedom and the finite range of interactions combined with either the one dimensional structure of the lattice or with high temperature ($d > 1$). The ergodicity proven here means that the dynamical XY-model will reach equilibrium exponentially fast. Furthermore, the proofs presented here provide explicit estimates of the constants of the logarithmic Sobolev inequalities of the respective Gibbs measures: at high temperature, they grow in proportion to $\frac{1+4d\beta}{1-4d\beta}$, whereas in one dimension the estimate behaves like e^{β}. One can not hope to show a logarithmic Sobolev at high dimensions at all temperatures because the resulting uniform ergodicity would contradict the non-uniqueness of Gibbs measures. Lastly, we want to mention that a logarithmic Sobolev inequality at high temperature can also be proven by using Proposition 5.2 and certain mixing conditions, see [15]. In particular, this leads us to assume that in two dimensions ($d = 2$) there exists a regime where the Gibbs measure is unique yet we do not have uniform ergodicity of the dynamical XY-model. This is suggested by the Berezinskii-Kosterlitz-Thouless phase transition (see [9]) which states the existence of a regime in $d = 2$ where the Gibbs measure is unique but correlations decay slowly.

Although most of the proofs given here are similar to the corresponding proofs for the Ising model, there are important differences resulting from the continuous nature of the spin states in the XY-model. In particular, we cannot use that the spin system is attractive as in [12]. Instead we can make use of the fact that Γ_1 satisfies the Leibniz rule, which in turn amplifies the usefulness of logarithmic Sobolev inequalities.

It is interesting to consider potential extensions of this model. For example, one might be able to obtain similar results for a XY-model in which the interactions between neighbouring spins vary randomly, which yields a so-called spin glass model [15]. An interesting Hamiltonian to consider might be

$$H(\sigma) = \sum_{|i-j|=1} X_{i,j} \langle \sigma_i, \sigma_j \rangle,$$

where the $\{X_{i,j}\}_{i,j \in \mathbb{Z}^d}$ are identically and independent real random variables, for example centered Gaussians with variance β. If the $X_{i,j}$ are almost-surely uniformly bounded, our earlier results still hold in an almost-sure sense, as they depend either on the one-dimensional structure of the underlying lattice or on a spectral condition in Theorem 5.7. Further, one can replace $\mathbb{S}^1$ by a general compact Riemannian manifold M. Again, our results should still hold. In particular, if one were to choose ($\mathbb{S}^{n-1}, n \geq 3$), Theorem 5.7 still holds, see [5].

Acknowledgements

There are many things for which I would like to thank Prof. Heydenreich: for introducing me to the world of probability theory and stochastic processes, for supervising first my bachelor and now my master thesis, and for the many interesting lectures and seminars of his which I was able to attend during my very enjoyable time at LMU Munich. I also would like to thank Prof. Bauerschmidt for helpful correspondence and references regarding logarithmic Sobolev inequalities. Then I would like to thank my parents for their strong support during the years of studying, and my brother Felix for being such a good and close friend. Many thanks also go to all my friends who supported me during my studies, in particular my girlfriend Melina.

7 References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffler; *Sur Les Inégalités de Sobolev Logarithmiques*; Panoramas et Synthèses, Société Mathématique de France, Paris, 2000.

[2] D. Bakry, M. Emery; Diffusions hypercontractivies; *Séminaire de probabilités de Strasbourg*, Tome 19, pp. 177-206, 1985.

[3] D. Bakry; *L'hypercontractivité et son utilisation en théorie des semigroupes*; Lectures on probability theory, École d'été de probabilités de St-Flour 1992, Springer, Berlin, 1994.

[4] D. Bakry, I. Gentil, M. Ledoux; *Analysis and Geometry of Markov Diffusion Operators*; Grundlehren der mathematischen Wissenschaften 348, Springer, Heidelberg, 2014.

[5] R. Bauerschmidt, T. Bodineau; A very simple proof of the LSI for high temperature spin systems; *Journal of Functional Analysis*, 276:8, pp. 2582-2588, 2019.

[6] F. Barthe, C. Roberto; Sobolev inequalities for probability measures on the real line; *Studia Mathematica*, 159(3):481-497, 2003.

[7] S. G. Bobkov, F. Götze; Exponential integrability and transportation cost related to logarithmic Sobolev inequalities; *Journal of Functional Analysis*, 163, pp. 1-28, 1999.

[8] H. Chernoff; A Note on an Inequality Involving the Normal Distribution; *Ann. Probab.*, Vol. 3, pp. 533-535, 1981.

[9] S. Friedli, Y. Velenik; *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*; Cambridge University Press, Cambridge, 2017.

[10] E. Laroche; Hypercontractivité pour des systèmes de spins de portée infinie; *Prob. Theory and Rel. fields*, Vol. 101(1), pp. 89-132, 1995.

[11] A. Levin, Y. Peres, E. Wilmer; *Markov Chains and Mixing Times*; Amer. Math. Soc., Providence, 2017.

[12] T. Liggett; *Continuous Time Markov Processes*; Graduate Studies in Mathematics: Volume 113, Amer. Math. Soc., Providence, 2010.

[13] H.-O. Georgii; *Gibbs measures and phase transitions*; de Gruyter, Berlin, 2011.

[14] L. Gross; Logarithmic Sobolev Inequalities; *American Journal of Mathematics*, 97, pp. 1061 - 1083, 1976.

[15] A. Guionnet, B. Zergalinski; Lectures on logarithmic Sobolev inequalities; *Séminaire de Probabilités*, XXXVI, volume 1801 of *Lecture Notes in Math.*, pp. 120-216. Springer, Berlin, 2003.

[16] I. Karatzas, S. Shreve; *Brownian motion and stochastic calculus*; Graduate Texts in Mathematics 113, Springer, New York, 2007.

[17] L. Saloff-Coste; *Lectures on Probability Theory and Statistics: Lectures on finite Markov Chains*; Springer, Berlin, Heidelberg, 1997.

[18] J. Sieber; Formulae for the Derivative of the Poincaré Constant of Gibbs Measures; *Preprint*, arxiv.org/abs/1910.08826, 2019.

[19] Z. Zhang, B. Qian, Y. Ma; Uniform Logarithmic Sobolev Inequality for Boltzmann Measures with Exterior Magnetic Field over Spheres; *Acta App. Math.*, Vol.116(3), pp.305-315, 2011.

Eigenständigkeitserklärung

Hiermit versichere ich, dass ich die vorliegende Masterarbeit selbstständig und nur mit den angegebenen Hilfsmitteln verfasst habe. Alle Passagen und Bilder, die ich wörtlich oder sinngemäß aus der Literatur übernommen habe, habe ich deutlich als Zitat mit Angabe der Quelle kenntlich gemacht. Darüber hinaus versichere ich, dass die eingereichte Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist.

Ort und Datum

Unterschrift