Cluster Swapping Techniques and Properties of Surface Tension

Chang Liu, Jianliang Ye

December 2023

Abstract

In this summary, we dissect the central theme of Chapter 8 of Scott Sheffield's thesis [She06] on discretized random surfaces: using a geometric technique called cluster swapping to deduce surface tension strict convexity, and ergodic gradient Gibbs phase uniqueness.

Contents

1 Notations

Configuration space and σ-algebras

- $(E, \mathcal{E})$ is a measurable space with Borel σ-algebra $\mathcal{E}$ on E. In this summary, E could be $\mathbb{R}^n$ or $\mathbb{Z}^n$ and the measure on $(E, \mathcal{E})$ is Lebesgue measure or counting measure respectively.
- Ω is the set of functions ϕ from $\mathbb{Z}^d$ to E.
- $\theta_x(\phi)$ is the translation of $\phi \in \Omega$ by $x \in \mathbb{Z}^d$, i.e. $(\theta_x \phi)(y) = \phi(x + y)$
- $\mathcal{F}$ is the Borel σ-algebra of the product topology on Ω.

- Let $\Lambda \subset \mathbb{Z}^d$, $\mathcal{F}_\Lambda$ is the smallest σ-algebra such that $\phi(x)$ is the measurable for all $x \in \Lambda$.
- $\mathcal{T}_\Lambda := \mathcal{F}_{\mathbb{Z}^d - \Lambda}$.
- $\mathcal{T} := \bigcap_{\Lambda \subset\subset \mathbb{Z}^d} \mathcal{T}_\Lambda$ and sets in $\mathcal{T}$ are called tail-measurable sets.

Gibbs Potentials and Hamiltonians

- We denote the Hamiltonian on Λ by H_Λ.
- γ_Λ^Φ is a probability kernel from $(\Omega, \mathcal{T}_\Lambda)$ to $(\Omega, \mathcal{F})$ where Φ is the potential in Hamiltonian. For any fixed $A \in \mathcal{F}$, it is a $\mathcal{T}_\Lambda$-measurable function of ϕ:

$$\gamma_\Lambda^\Phi(A, \phi) = Z_\Lambda(\phi)^{-1} \int \prod_{x \in \Lambda} d\phi(x) \exp\left(-H_\Lambda(\phi)\right) \mathbf{1}_A(\phi)$$

  where $Z_\Lambda(\phi) = \int \prod_{x \in \Lambda} d\phi(x) \exp\left(-H_\Lambda(\phi)\right)$.
- We say ϕ is Φ-admissible if each $Z_\Lambda(\phi)$ is finite and non-zero.
- $\gamma_\Lambda^\Phi(\cdot|\phi^0) := \gamma_\Lambda^\Phi(\cdot, \phi = \phi^0)$ is the Gibbs measure conditional on $\phi^0 \in \mathcal{T}_\Lambda$.
- $\mathcal{F}^\tau$ is the subset of $\mathcal{F}$ containing those sets that are invariant under translations $\phi \mapsto \phi + z$ for $z \in E$ and we write $\mathcal{T}_\Lambda^\tau = \mathcal{T}_\Lambda \cap \mathcal{F}^\tau$ and $\mathcal{F}_\Lambda^\tau = \mathcal{F}_\Lambda \cap \mathcal{F}^\tau$.

Spaces of probability measures on configuration space

- $\mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\tau)$ is the set of $\mathcal{L}$-invariant probability measures on $(\Omega, \mathcal{F}^\tau)$ where $\mathcal{L}$ is a set of rank-d sublattice of $\mathbb{Z}^d$.
- $\mathcal{G}(\Omega, \mathcal{F}^\tau)$ (abbriviated as $\mathcal{G}^\tau$) is the set of gradient Gibbs measures on $(\Omega, \mathcal{F}^\tau)$, i.e., measures μ such that for all $\Lambda \subset\subset Z^d, 0 < Z_\Lambda(\phi) < \infty$ μ-a.s. and $\mu\gamma_\Lambda = \mu$
- $\mathcal{G}_\mathcal{L}(\Omega, \mathcal{F}^\tau)$ (abbriviated as $\mathcal{G}_\mathcal{L}^\tau$) is the set of $\mathcal{L}$-invariant gradient Gibbs measures on $(\Omega, \mathcal{F}^\tau)$.

Slopes and surface tension

- $S(\mu)$ is the slope of μ, $\mu \in \mathcal{P}_\mathcal{L}(\Omega, \mathcal{F}^\tau)$.
- u is slope variable (a linear function from $\mathbb{R}^d$ to $\mathbb{R}^m$)
- σ^Φ, σ is the surface tension

$$\sigma(u) = \inf\left\{SFE^\Phi(\mu) : \mu \in \mathcal{P}_\mathcal{L}\left(\Omega, \mathcal{F}^\tau\right), S(\mu) = u\right\}$$

  where SFE is the specific free energy that will be (re-)defined later (§3).
- U_Φ is the interior of set of slopes u with $\sigma(u) < \infty$.

2 The cluster swapping technique

We define $\mathbb{E}^d$ be the set of edges of the lattice $\mathbb{Z}^d$. Σ is the set of functions from $\mathbb{E}^d$ to $[0,\infty)$. Let π be the measure on Σ describing an independent product of parameter one exponential. We write $\bar{\Omega} = \Omega \times \Omega \times \Sigma$ and let $\bar{\mathcal{F}}$ be the product σ-algebra on $\bar{\Omega}$. We will use (ξ, ζ, t) to represent the triplet $(\phi_1, \phi_2, r) \in \bar{\Omega}$, where $\xi = (\min\{\phi_1(x), \phi_2(x)\}, \max\{\phi_1(x), \phi_2(x)\})$, $\zeta = \mathbf{1}_{\phi_1 > \phi_2} - \mathbf{1}_{\phi_1 < \phi_2}$, and t is the total energy. Note that the coordinate change is injective so the representation is unique. We define the operator T such that $T(\phi_1, \phi_2, r) = (\xi, \zeta, t)$.

For the Ising model with coupling constants $K_e \in \mathbb{R}$ on the edge of Λ, we define $\mathcal{S}$ be the set of edges of Λ at which $t(e) \geq |K_e|$. An edge is open if it lies in the complement of $\mathcal{S}$. Note that an isolated point that all its edges are in $\mathcal{S}$ is itself an open cluster.

Lemma 2.1. *Let $(\phi_1, \phi_2, r) \in \bar{\Omega}$ be sampled from the law $\gamma_\Lambda\left(\cdot|\phi_1^0\right) \otimes \gamma_\Lambda\left(\cdot|\phi_2^0\right) \otimes \pi$, and define (ξ, ζ, t) from (ϕ_1, ϕ_2, r) as above. Then conditioned on ξ and the total energy t (which determine the K_e and $\mathcal{S}$), the conditional law of ζ is as follows: throughout each component of the complement of $\mathcal{S}$ containing a vertex outside of Λ, ξ is equal to its value at that vertex. On each component of the complement of $\mathcal{S}$ that is strictly contained in Λ, ζ is a.s. constant, and the law of the values on these components is given by an independent fair coin toss on each component.*

Definition 2.1 (Stochastic domination)**.** Let μ and ν be probability measures on an arbitrary measure space $(X, \mathcal{X})$. We say that $\mu \prec \nu$ or ν stochastically dominates μ if there exists a measure ρ on $X \times X$ with the product σ-algebra such that for ρ-a.s. $(a, b) \in (X, X)$ we have $a \leq b$ and the first and second marginals of ρ are respectively μ and ν.

Theorem 2.2. *Suppose that $\phi_1^0, \phi_2^0 \in \Omega$ are admissible and $\phi_1^0 \leq \phi_2^0$. Then for any $\Lambda \subset\subset \mathbb{Z}^d$, we have $\gamma_\Lambda\left(\cdot \mid \phi_1^0\right) \prec \gamma_\Lambda\left(\cdot \mid \phi_2^0\right)$.*

Proof. We consider the triplet (ϕ_1, ϕ_2, r) with the law $\gamma_\Lambda\left(\cdot \mid \phi_1^0\right) \otimes \gamma_\Lambda\left(\cdot \mid \phi_2^0\right) \otimes \pi$ and we do the coordinate change to (ξ, ζ, t). Since $\phi_1^0 \leq \phi_2^0$, we have $\zeta \leq 0$ in $\mathbb{Z}^d - \Lambda$. By the preceding lemma, we know for open cluster not strictly contained in Λ, $\zeta \leq 0$. For other open cluster that strictly contained in Λ and $\zeta \neq 0$, the sign of ζ is determined by fair toss. That is, with $1/2$ probability that $\phi_1 = \xi_1$ and $\phi_2 = \xi_2$ and $1/2$ probability that $\phi_1 = \xi_2$ and $\phi_2 = \xi_1$. For these clusters, we can take $\phi_1 = \phi_2 = \xi_1$ or $\phi_1 = \phi_2 = \xi_2$, each with $1/2$ probability. Note that this change doesn't affect the marginal law of ϕ_1 and ϕ_2 and we guarantee that $\zeta \leq 0$ (equivalently $\phi_1 \leq \phi_2$). □

Corollary 2.3. *Let ϕ_c be the function which is equal to an admissible function ϕ_0 everywhere except at one vertex $x \in \mathbb{Z}^d \backslash \Lambda$ where it is equal to c; when c is in the interval for which ϕ_c is admissible, let $F(c)$ be the $\gamma_\Lambda(\cdot \mid \phi_c)$-expected value of $\phi(y)$, where $y \in \Lambda$. Then $F(c)$ is monotone increasing and $F(c_2) - F(c_1) \leq c_2 - c_1$ for all $c_1, c_2 \in \mathbb{E}$. In particular, if c is chosen from a distribution ν on*

E (supported on c for which ϕ_c is admissible), then the variance of $F(c)$ is less than or equal to the variance of c.

Proof. Note that $F(c) - c$ is the expected value of ϕ under the law $\gamma(\cdot|\phi_c - c)$. Then the first result follows from the theorem 2.2. To prove the second result, we sample a_1 and a_2 independently from ν. By $(F(a_1) - F(a_2))^2 \leq (a_1 - a_2)^2$ for all a_1, a_2, we have $\mathrm{Var}(F(a_1) - F(a_2)) \leq \mathrm{Var}(a_1 - a_2)$. The result follows by independence. □

Another application of the coordinates change (ξ, ζ, t) and the conditional law of ζ given by lemma 2.1 is to show the log-concavity of the law of ϕ. A probability distribution on E is log concave if the log of its Radon-Nikodym derivative f with respect to E is concave. On $\mathbb{Z}$, this is equivalent to $f(a)^2 \geq f(a+1)f(a-1)$ for all $a \in \mathbb{Z}$.

Lemma 2.4. *Suppose $\Lambda \subset\subset \mathbb{Z}^d, x_0 \in \Lambda$, and $\phi_0 \in \Omega$ is admissible. If ϕ is a random function chosen from $\gamma_\Lambda(\cdot \mid \phi_0)$, then the law of $\phi(x_0)$ is log concave.*

Proof. We give the proof for the case that $E = \mathbb{Z}$. A similiar proof using regular conditional probabilities applies to the case that $E = \mathbb{R}$. We consider $\phi_1^0 = \phi_0$ and $\phi_2^0 = \phi_0 + c$ for some $c \in \mathbb{Z}$ greater than 0. We sample the triplet (ϕ_1, ϕ_2, r) from the law $\gamma_\Lambda\left(\cdot \mid \phi_1^0\right) \otimes \gamma_\Lambda\left(\cdot \mid \phi_2^0\right) \otimes \Sigma$. Conditioned on $\xi(x_0) = (a, a+c)$ and t, by lemma 2.1, we have $\phi_2(x_0) \geq \phi_1(x_0)$ with probability $\frac{1}{2}$ if x is in an open cluster that strictly contained in Λ and with probability 1 otherwise. Hence we have

$$\mathbb{P}\left(\phi_2(x_0) \geq \phi_1(x_0) \mid \xi(x_0) = (a, a+c)\right) \geq \frac{1}{2}.$$

This implies that

$$\begin{aligned}\mathbb{P}(\phi_1(x_0) = a, \phi_2(x_0) = a+c) &\geq \frac{1}{2}\mathbb{P}(\xi(x_0) = (a, a+c)) \\ &= \frac{1}{2}\mathbb{P}(\phi_1(x_0) = a, \phi_2(x_0) = a+c) \\ &+ \frac{1}{2}\mathbb{P}(\phi_1(x_0) = a+c, \phi_2(x_0) = a)\end{aligned}$$

Then

$$\mathbb{P}(\phi_1(x_0) = a, \phi_2(x_0) = a+c) \geq \mathbb{P}(\phi_1(x_0) = a+c, \phi_2(x_0) = a).$$

Since ϕ_1 and ϕ_2 are sampled independently under the law $\gamma_\Lambda\left(\cdot \mid \phi_1^0\right) \otimes \gamma_\Lambda\left(\cdot \mid \phi_2^0\right) \otimes \Sigma$, $\mathbb{P}(\phi_1(x_0) = m, \phi_2(x_0) = n) = \mathbb{P}(\phi_1(x_0) = m)\mathbb{P}(\phi_2(x_0) = n)$. Note that we also have $\mathbb{P}(\phi_1(x_0) = m) = \mathbb{P}(\phi_2(x_0) = m + c)$. Therefore,

$$\mathbb{P}(\phi_1(x_0) = a)^2 \geq \mathbb{P}(\phi_1(x_0) = a+c)\mathbb{P}(\phi_1(x_0) = a-c),$$

which is exactly $f(a)^2 \geq f(a+c)f(a-c)$ under the case $E = \mathbb{Z}$. □

Definition 2.2 (cluster swap)**.** A *cluster swap* is an operator R_Λ^x from $\bar{\Omega}$ to $\bar{\Omega}$ that sends (ϕ_1, ϕ_2, r) to (ϕ_1', ϕ_2', r'). Let $T(\phi_1, \phi_2, r) = (\xi, \zeta, t)$ and $T(\phi_1', \phi_2', r') = (\xi', \zeta', t')$. We have $\xi' = \xi, t' = t$, and $\zeta' = \zeta$ unless the vertices in the open cluster containing x are all contained within Λ, in which case $\zeta' = -\zeta$ on that cluster and $\zeta' = \zeta$ everywhere else. We write $R^x = R_{\mathbb{Z}^d}^x$.

Corollary 2.3 and Lemma 2.4 along with other established results in [She06, §4.2, §6.1] can be applied to prove the existence of gradient Gibbs measures of a given slope and it can be shown that is minimal:

Lemma 2.5. *Let Ω be a space with a σ-algebra $\mathcal{F}^\tau$, and let Φ be a given set of potentials defining a gradient Gibbs measure. For every slope u in the admissible set U_Φ, there exists an ergodic gradient phase μ_u in the space $\mathcal{G}_\mathcal{L}^\tau(\Omega, \mathcal{F}^\tau)$. Furthermore, this ergodic gradient phase μ_u is a minimal gradient phase. That is, $SFE(\mu_u) = \sigma(u)$.*

3 Measures on triplets and definitions

Given the space $\bar{\Omega} = \Omega \times \Omega \times \Sigma$ of infinite triplets, let $\bar{\mathcal{F}}$ be the product σ-algebra on $\bar{\Omega}$ and let $\bar{\mathcal{F}}^\tau$ be the σ-algebra generated by $\mathcal{F}^\tau \times \mathcal{F}^\tau$ times the product topology on Σ.

Many concepts defined in Ω can now be extended analogously to those in $\bar{\Omega}$. Let Φ be extended as

$$\Phi_\Lambda(\phi_1, \phi_2, r) = \Phi_\Lambda(\phi_1) + \Phi_\Lambda(\phi_2) + \sum_e r(e)$$

where the last sum is over all edges e which contain at least one vertex of Λ.

We now write

$$\gamma_\Lambda^\Phi(A, (\phi_1, \phi_2, r)) = Z_\Lambda(\phi_1)^{-1} Z_\Lambda(\phi_2)^{-1}$$
$$\int \prod_{x \in \Lambda} d\phi_1(x) \prod_{x \in \Lambda} d\phi_2(x) \prod_e dr(e) \exp\left[-H_\Lambda(\phi_1, \phi_2, r)\right] \mathbf{1}_A(\phi_1, \phi_2, r)$$

where the products over e are taken over edges with at least one vertex of Λ.

A Gibbs measure on $(\bar{\Omega}, \bar{\mathcal{F}})$ is a measure on $(\bar{\Omega}, \bar{\mathcal{F}})$ which is preserved by these kernels. A gradient Gibbs measures is defined accordingly, replacing $\bar{\mathcal{F}}$ with $\bar{\mathcal{F}}^\tau$. Then in any Gibbs measure or gradient Gibbs measure on $(\bar{\Omega}, \bar{\mathcal{F}})$, the random variables $r(e)$ are independent of ϕ_1, ϕ_2, and independently identically distributed according to an $\exp(1)$ distribution on $[0, \infty)$. We denote this last measure on Σ by π. A Gibbs measure is extremal if it is an extreme point of the convex set of all Gibbs measures or, equivalently, if it satisfies a zero-one law on tail events.

We say that a (gradient) Gibbs measure on triplets is $\mathcal{L}$-invariant if it is invariant under the shifts $\theta_v, v \in \mathcal{L}$ which move the three components ψ_1, ψ_2, r in tandem;

it is $\mathcal{L}$-ergodic if it is extremal in the set of $\mathcal{L}$-invariant measures and extremal if it is extremal in the set of Gibbs measures (gradient Gibbs measures) on $(\bar{\Omega}, \bar{\mathcal{F}})$ (resp., $(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$).

We define Gibbs measures, free energy, and specific free energy (SFE) by first replacing H_Λ, the usual Hamiltonian in Λ, with

$$\bar{H}_\Lambda(\phi_1, \phi_2, r) = H_\Lambda(\phi_1) + H_\Lambda(\phi_2) + \sum_e r(e)$$

In particular, if $\mu \in \mathcal{P}_\mathcal{L}(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$, then we write

$$SFE(\mu) = \lim_{n\to\infty} |\Lambda_n|^{-1} \mathcal{H}\left(\mu_{\Lambda_n}, e^{-\bar{H}^o_{\Lambda_n}} \lambda^{|\Lambda_n - 1|} \otimes \lambda^{|\Lambda_n - 1|} \otimes [0,\infty)^{|\Sigma_n|}\right)$$

where $\lambda^{|\Lambda_n - 1|}$ are the measure ν on $\left(\Omega, \mathcal{F}^\tau_{\Lambda_n}\right)$ such that for any measurable $A \subset E^{|\Lambda_n|-1}$, the value

$$\nu\left(\left\{\phi \mid \left(\phi(v_1) - \phi(v_0), \phi(v_2) - \phi(v_0), \ldots, \phi_{|\Lambda_n|-1} - \phi(v_0)\right) \in A\right\}\right)$$

is equal to the measure of A in the product measure $\lambda^{|\Lambda_n|-1}$, and $\bar{H}^o_\Lambda$ is the sum of the energy contributions from edges strictly contained in Λ, and Σ_n is the set of edges with both endpoints in Λ_n. Note that $SFE(\mu)$ affine, so that by definition the surface tension σ is convex.

We also define the slope $S(\mu) = (u, v)$ to be the two slopes of the marginal distributions of μ. We write $S_a(\mu) = (u+v)/2$ for the average slope of μ.

We state but not prove the following analog for triplets of the first half of *variational principle*:

Lemma 3.1. *If $\mu \in \mathcal{P}_\mathcal{L}(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$ has minimal specific free energy among $\mathcal{L}$-invariant measures with slope (u, v), then μ is a gradient Gibbs measure.*

We will also need the following lemma, which follows by Lemma 2.5.

Lemma 3.2. *The minimal specific free energy among measures of slope (u, v) is equal to $\sigma(u) + \sigma(v)$ and is obtained by an independent product $\mu_u \otimes \mu_v \otimes \nu$ where ν is an independent product of parameter one exponentials.*

4 More definitions on infinite clusters

Given $(\phi_1, \phi_2, r) \in \bar{\Omega}$, we define the following variables:

- $\mathcal{S}_c$ is the set of all edges for which $(\phi_1 + c, \phi_2, r)$ is swappable.
- T^+_c is the union of all vertices v in infinite clusters of $\mathbb{E}^d \backslash \mathcal{S}_c$ for which $\phi_2(v) > \phi_1(v) + c$ throughout the cluster. Similarly, T^-_c contains the vertices v in infinite clusters of $\mathbb{E}^d \backslash \mathcal{S}_c$ for which $\phi_2(v) < \phi_1(v) + c$ throughout the cluster.

- $B^+ = \inf\{c : T_c^+ \text{ is empty}\}$. We say $B^+ = \infty$ if no such c exists, i.e., if T_c^+ fails to be empty for any finite c. Similarly, $B^- = \sup\{c : T_c^+ \text{ is empty}\}$ and $B^- = -\infty$ if no such c exists.

With these definitions, we state but not prove the following lemmas.

Lemma 4.1 (coupling)**.** *Suppose that μ_1 and μ_2 are extremal (non-gradient) Gibbs measures on Ω. Then there exist values B_0^+ and B_0^- such that $\mu_1 \otimes \mu_2 \otimes \pi$-almost surely, we have $B^+ = B_0^+$ and $B^- = B_0^-$. Moreover, $\mu_1 + B_0^- \prec \mu_2 \prec \mu_1 + B_0^+$. In particular, $B_0^- \leq B_0^+$ with equality if and only if, up to additive constant, $\mu_1 = \mu_2$ (i.e., the restrictions of μ_1 and μ_2 to $\mathcal{F}^\tau$ are equivalent).*

Corollary 4.2. *If μ_1 and μ_2 are distinct gradient phases, then $\mu_1 \otimes \mu_2 \otimes \pi$-almost surely, $B^- < B^+$.*

A function $h : \Omega \mapsto \mathbb{E} \cup \{\infty\}$ is called a *height difference variable* for μ if the following are true:

1. $h(\phi_1 + c_1, \phi_2 + c_2, r) = h(\phi_1, \phi_2) + c_2 - c_1$ for all $(\phi_1, \phi_2, r) \in \bar{\Omega}$ and $c_1, c_2 \in E$.
2. h is tail-measurable on $\bar{\Omega}$.
3. h is μ-almost surely finite.
4. If $v \in \mathcal{L}$, then, μ-almost surely, $h(\phi_1, \phi_2, r) = h(\theta_v \phi_1, \theta_v \phi_2, \theta_v r)$.

Lemma 4.3 (uniqueness of infinite clusters)**.** *Let μ be an $\mathcal{L}$-ergodic gradient Gibbs measure on $(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$ with slope (u, v), where $u, v \in U_\Phi$; suppose that h is a height difference variable for μ. Then there exists no $c \in \mathbb{R}$ for which, with μ-positive probability, either T_{c-h}^+ or T_{c-h}^- consists of more than one infinite component.*

5 The main convexity and uniqueness results

The following theorem is central to the entire chapter.

Theorem 5.1. *Suppose that μ is a measure on $(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$ whose first two marginals μ_1 and μ_2 are minimal $\mathcal{L}$-ergodic gradient phases on $(\Omega, \mathcal{F}^\tau)$ with slopes in U_Φ. Suppose further that with μ positive probability, we have $B^+ > B_-$ (when $E = \mathbb{R}$) or $B^+ > B^- + 1$ (when $E = \mathbb{Z}$). Then for some appropriately defined "infinite cluster swapping map" R, we have $SFE(R(\mu)) \leq SFE(\mu)$ and $S_a(\mu) = S_a(R(\mu))$; moreover, $R(\mu)$ is an L-invariant gradient measure on $(\bar{\Omega}, \bar{\mathcal{F}}^\tau)$ which is not a gradient Gibbs measure.*

The above theorem can be proved by infinite cluster swaps, which we will show in the next section. By this theorem however, we can first deduce strict convexity of surface tension.

Theorem 5.2. *The surface tension σ is strictly convex in U_Φ.*

Proof. Pick distinct slopes u_1 and u_2 in U_Φ. By Lemma 2.5, there exist $\mathcal{L}$-ergodic Gibbs measures μ_1 and μ_2 of slopes u_1 and u_2 such that $SFE(\mu_1) = \sigma(u_1), SFE(\mu_2) = \sigma(u_2)$. Denote $\mu := \mu_1 \otimes \mu_2 \otimes \pi$. It can be shown by Lemma 4.1 that μ satisfies the requirements of Theorem 5.1, which implies that $R(\mu)$ is not a gradient Gibbs measure, and $S_a(\mu) = S_a(R(\mu))$. Then, by Lemma 3.1 and 3.2, we see that $SFE(R(\mu))$ is not minimal. Together with the facts that the swapping R cannot increase SFE and that σ is convex, we have

$$\begin{aligned}\sigma(u_1) + \sigma(u_2) &= SFE(\mu_1) + SFE(\mu_2)\\ &\geq SFE(R(\mu)) > \sigma(v_1) + \sigma(v_2)\\ &\geq 2\sigma\left(\frac{v_1+v_2}{2}\right) = 2\sigma\left(\frac{u_1+u_2}{2}\right)\end{aligned}$$

Therefore σ is strictly convex on U_Φ. □

Theorem 5.1 can also be used in a similar way to show the uniqueness of ergodic gradient Gibbs phase.

Theorem 5.3. *If $E = \mathbb{R}$ (the Ginzburg-Landau case), then for every $u \in U_\Phi$, there exists a unique minimal gradient phase μ_u of slope u.*

Proof. By Lemma 2.5, there exists at least one minimal gradient phase of slope u. Suppose that μ_1 and μ_2 are distinct minimal gradient phases of slope u; then $\mu = \mu_1 \otimes \mu_2 \otimes \pi$ satisfies the requirements of Theorem 5.1.

Let (v_1, v_2) be the slope of $R(\mu)$. Since $R(\mu)$ is not a gradient Gibbs measure, by Lemma 3.1 and 3.2, the convexity of σ, and that $S_a(\mu) = S_a(R(\mu))$, we have

$$2\sigma(u) \geq SFE(R(\mu)) > \sigma(v_1) + \sigma(v_2) \geq 2\sigma\left(\frac{v_1+v_2}{2}\right) = 2\sigma(u)$$

a contradiction. □

We omit the similar result applied on $E = \mathbb{Z}$ where we need the definition *quasi-equivalence.*

6 A sketch of the proof of Theorem 5.1

We first assume that there exists a hight difference variable for the measure μ. We define $B_0^- = B^- - h$ and $B_0^+ = B^+ - h$. Then, with positive probability, $B_0^- < B_0^+$, so there exists a $c \in (B_0^-, B_0^+)$ that $B^- < c + h < B^+$. Note that with this c, both T_{c+h}^+ and T_{c+h}^- are nonempty. For any c, let the coordinate change R_c maps the triplet (ϕ_1, ϕ_2, r) to (ψ_1, ψ_2, s) where

$$\psi_1(x) = \begin{cases} \phi_1(x) & x \in T^-_{c+h} \\ \phi_2(x) - c - h & \text{otherwise} \end{cases}$$
$$\psi_2(x) = \begin{cases} \phi_2(x) & x \in T^-_{c+h} \\ \phi_1(x) + c + h & \text{otherwise} \end{cases}$$

and s is determined by preserving the total energy of each edge. Now we define infinite cluster swapping map

$$R(\mu) = \begin{cases} R_c(\mu) & E = \mathbb{Z} \text{ and } c \in \mathbb{Z} \text{ s.t. } B_0^- < c < B_0^+; \\ \frac{1}{c_2 - c_1} \int_{c_1}^{c_2} R_c(\mu) dc & E = \mathbb{R} \text{ and } B_0^- < c_1 < c_2 < B_0^+. \end{cases}$$

$R(\mu)$ is $\mathcal{L}$-invariant by construction. Since both T^-_{c+h} and T^+_{c+h} are infinite cluster after swapping, $R(\mu)$ cannot be a gradient Gibbs measure by lemma 4.3. As the averages

$$\frac{[\phi_1(y) - \phi_1(x)] + [\phi_2(y) - \phi_2(x)]}{2}$$

are left unchanged by swapping maps, it is also clear that $S_a(\mu) = S_a(R(\mu))$. We now claim that

$$SFE(R(\mu)) \le SFE(\mu).$$

By the definition of specific free energy, this is

$$\lim_{n\to\infty} |\Lambda_n|^{-1} FE_{\Lambda_n}(R(\mu)) \le \lim_{n\to\infty} |\Lambda_n|^{-1} FE_{\Lambda_n}(\mu).$$

where $FE_{\Lambda_n}(\mu)$ is defined as

$$\mathcal{H}_{\overline{\mathcal{F}}^\tau_{\Lambda_n}} \left(\mu_\lambda, e^{-H^o_\Lambda(\phi_1,\phi_2,r)} \prod_{x \in \Lambda_n \setminus \{x_0\}} d\left[\phi_1(x) - \phi_1(x_0)\right] d\left[\phi_2(x) - \phi_2(x_0)\right] \prod_e dr(e) \right)$$

To prove it, we define $F(x) = 0$ if $x \in T^-_{c+h}$, and $F(x) = 1$ otherwise. $F_{\partial\Lambda_n}$ is the restriction of F to the boundary of Λ_n. Let μ'_{Λ_n} be the law of the five-tuple $(\phi_1 - \phi_1(x_0), \nabla\phi_2 - \phi_2(x_0), r, F_{\partial\Lambda_n}, c_0)$, where x_0 is a reference vertex defined on the boundary of Λ_n and c_0 is defined as follow. For given values of the triplet $(\phi_1 - \phi(x_0), \phi_2 - \phi_2(x_0), r)$, let b_1 and b_2 be the lower and upper bounds on the set of choices of c_0 for which $(\phi_1 - \phi_1(x_0), \phi_2 - \phi_2(x_0) + c_0, r)$ has any swappable edges on Λ_n. Let

$$B = \begin{cases} b_1 + c_1 & \phi_2(x_0) - \phi_1(x_0) - h \le b_1 + c_1 \\ b_2 + c_2 & \phi_2(x_0) - \phi_1(x_0) - h \ge b_2 + c_2 \\ \phi_2(x_0) - \phi_1(x_0) - h & \text{otherwise} \end{cases}$$

and

$$c_0 = \begin{cases} B - c & E = \mathbb{R} \text{ and } c \text{ is chosen uniformely in } [c_1, c_2] \\ B & E = \mathbb{Z}. \end{cases}$$

We can view R as a map on the five-tuple and it maps $(\phi_1 - \phi_1(x_0), \phi_2 - \phi_2(x_0), r, F_{\partial\Lambda_n}, c_0)$ to $(\psi_1, \psi_2, s, F_{\partial\Lambda_n}, c_0)$, which swaps the triplet on the swappable set and then transforms it by R and leaves $F_{\partial\Lambda_n}$ and c_0 fixed. Now, define $FE_{\Lambda_n}(\mu')$ to be the relative entropy of μ' with respect to $\nu_1 \otimes \nu_2 \otimes \nu_3$ where

$$\nu_1 = e^{-H^o_{\Lambda_n}(\phi_1,\phi_2,r)} \prod_{x\in\Lambda_n\setminus\{x_0\}} d\left[\phi_1(x) - \phi_1(x_0)\right] d\left[\phi_2(x) - \phi_2(x_0)\right] \prod_{e\in\Lambda_n} dr(e);$$

$\nu_2 = dc_0$ is Lebesgue measure; and $\nu_3 = dF(x)$ is counting measure. Then R by construction is invertible and measure preserving under $\nu_1 \otimes \nu_2 \otimes \nu_3$. Then $FE_{\Lambda_n}(\mu') = FE_{\Lambda_n}(R(\mu'))$. We have the decomposition

$$\begin{aligned} FE_{\Lambda_n}(\mu') &= FE_{\Lambda_n}(\mu) + \mu\mathcal{H}\left(\mu^{\phi_1,\phi_2,r}, \nu_2\right) + \mu'\mathcal{H}\left(\mu^{\phi_1,\phi_2,r,c_0}, \nu_3\right) \\ &\leq FE_{\Lambda_n}(\mu) + o\left(|\Lambda_n|\right) \end{aligned}$$

and similiarly

$$\begin{aligned} &FE_{\Lambda_n}(R(\mu')) \\ =& FE_{\Lambda_n}(R(\mu)) + R(\mu)\mathcal{H}\left(R(\mu)^{\phi_1,\phi_2,r} \mid \nu_2\right) + R(\mu')\,\mathcal{H}\left(R(\mu)^{\phi_1,\phi_2,r,c_0} \mid \nu_3\right). \end{aligned}$$

Since $SFE(\mu) < \infty$, the expected value of $b_2 - b_1$ is $O(|\Lambda_n|)$ and one can show that $R(\mu)\mathcal{H}\left(R(\mu)^{\phi_1,\phi_2,r} \mid \nu_2\right) \geq -O(\log|\Lambda_n|)$. Then we have

$$\begin{aligned} FE_{\Lambda_n}(R(\mu)) &\leq FE_{\Lambda_n}(R(\mu')) + O(\log|\Lambda_n|) \\ &= FE_{\Lambda_n}(\mu') + o(|\Lambda_n|) \leq FE_{\Lambda_n}(\mu) + o(|\Lambda_n|). \end{aligned}$$

Therefore,

$$\lim_{n\to\infty} |\Lambda_n|^{-1} FE_{\Lambda_n}(R(\mu)) \leq \lim_{n\to\infty} |\Lambda_n|^{-1} FE_{\Lambda_n}(\mu).$$

In the case that there is no height difference variable h for μ, if either B^+ or B^- were finite with positive probability, then this B^+ or B^- would itself be a height difference variable for the measure μ_0 equal to μ conditioned on this event and the previous method is applicable. Hence we can assume neither B^+ or B^- were finite, by Lemma 4.3, we have positive probability that $B^+ = \infty$ and $B^- = -\infty$. For any $x \in \mathbb{Z}^d$, we define $F(x)$ be the smallest value c for which $x \in T_c^-$. We write $R_c(\phi_1, \phi_2, r, F_{\partial\Lambda_n}) = (\psi_1, \psi_2, s, F_{\partial\Lambda_n})$, where $F_{\partial\Lambda_n}$ is left unchanged and $(\psi_1, \psi_2, s) = R_c(\phi_1, \phi_2, r)$ as we defined previously. For a positive value $M \in E$, we define $R = \prod_{i=k_1}^{k_2} R_{kM}$ where k_1 is any odd integer for which $k_1 M < \inf_{x\in\partial\Lambda_n} F(x)$ and k_2 is any integer for which $k_2 M > \sup_{x\in\partial\Lambda_n} F(x)$. The detailed proof for properties of this R described in Theorem 5.1 can be found in [She06] and is ommited here.

References

[She06] Scott Sheffield. Random surfaces, 2006. arXiv:math/0304049.