4. The problem of phase transitions

4.1) The theorem of Lee and Yang

4.2) Existence of a first order phase transition at low temperatures

4.3) The theorem of Mermin and Wagner

4.4) Three additional statements without proof

This section follows Ruelle, Statistical mechanics, Chapter 5.

In this chapter we will take the point of view that that a phase transition manifests itself as a singularity of a thermodynamic function, which is otherwise real analytic.

4.1. The theorem of Lee and Yang

In 1952, Lee and Yang proposed a theory of phase transitions based on the idea that certain regions of the complex activity plane may remain free of zeros of the grand partition function when $\Lambda_\alpha \uparrow \mathbb{R}^d$. Lee and Yang exhibited a class of lattice systems for which all zeros lie on the circle with radius one. We will first establish the relevant mathematical results and afterwards discuss their application.

4.1.1. Proposition: Let $\{A_{ij}\}_{\substack{i,j=1\\ i\neq j}}^{n}$ be a family of real numbers such that $-1 \leq A_{ij} \leq 1$ and $A_{ij} = A_{ji}$ for $i,j=1\ldots n$. We define a polynomial P_n in n variables by

$$P_n(z_1 \ldots z_n) = \sum_{S} z^S \left(\prod_{i\in S} \prod_{j\in S'} A_{ij} \right), \qquad (1)$$

where the summation is over all subsets $S = (i_1 \ldots i_s)$ of $\{1,\ldots,n\}$, $z^S = z_{i_1} \ldots z_{i_s}$, and $S' = (j_1 \ldots j_{n-s})$ is the complement of S in $\{1,\ldots,n\}$. If S or S' is the empty set we define the coefficient in (1) to be 1. Then $P_n(z_1 \ldots z_n) = 0$ and $|z_1| \geq 1, \ldots, |z_{n-1}| \geq 1$ imply $|z_n| \leq 1$.

Proof: We can w.l.o.g. assume that $A_{ij} \neq 0, \pm 1$ for all $i,j = 1\ldots n$. If we can prove the Theorem with this assumption, then the general case follows by continuity.

We argue by induction and start by noting that

$$P_1(z_1) = 1 + z_1 . \tag{2}$$

If $P_1(z_1) = 0$ then $|z_1| = 1$, which proves the claim for $n = 1$. In case of $n = 2$ we have with $-1 < a < 1$

$$P_2(z_1, z_2) = 1 + a(z_1 + z_2) + z_1 z_2 = 0 \tag{3}$$

$$\Leftrightarrow \quad z_2 (a + z_1) = -1 - a z_1$$

$$\Leftrightarrow \quad |z_2|^2 = \frac{|1 + a z_1|^2}{|a + z_1|^2} = \frac{(1+ax)^2 + ay^2}{(a+x)^2 + y^2} \overset{!(*)}{\leq} 1$$

$z_1 = x + y$ with $x^2 + y^2 \geq 1$

$$\Leftrightarrow \quad \underbrace{(1+ax)^2}_{1 + 2ax + a^2x^2} \leq \underbrace{(a+x)^2}_{a^2 + 2ax + x^2} + (1-a) y^2$$

$$\Leftrightarrow \quad 1 + a^2 x^2 \leq a^2 + x^2 + \underbrace{(1-a)}_{\geq 0} y^2 .$$

$$\boxed{\begin{array}{l} 1 + a^2 x^2 \leq a^2 x^2 \\ \Leftrightarrow \underbrace{a^2}_{0 < \ldots < 1} \underbrace{(x^2 - 1)}_{\geq 0} \leq x^2 - 1 \quad \checkmark \\ \Rightarrow y^2\text{-term not needed} \\ \text{and } (*) \text{ holds.} \end{array}}$$

We will proceed by induction.

If $n \geq 3$ we assume that $|z_1| \geq 1, \dots, |z_{n-2}| \geq 1$, define $\Delta_n = \{1, \dots, n\}$ and write

$$P_n(z_1, \dots, z_n) = \sum_{S \subseteq \Delta_n} z^S \left(\prod_{i \in S} \prod_{j \in \Delta_n \setminus S} A_{ij} \right) \tag{4}$$

$$= A + B z_{n-1} + C z_n + D z_{n-1} z_n ,$$

where

$$D = \sum_{S \subseteq \Delta_{n-2}} z^S \left(\prod_{i \in S \cup \{n-1, n\}} \prod_{j \in \Delta_{n-2} \setminus S} A_{ij} \right). \tag{5}$$

We also define $\zeta_i = (A_{n-1,i} A_{ni})^{-1} z_i$ and note that

$$D = \left[\prod_{i=1}^{n-2} A_{n-1,i} A_{ni} \right] P_{n-2}(\zeta_1, \dots, \zeta_{n-2}). \tag{6}$$

By assumption, we have $|\zeta_1| > 1, \dots, |\zeta_{n-2}| > 1$, and the induction hypothesis $\left[|z_1|, \dots, |z_{n-3}| \geq 1 \text{ and } P_{n-2}(z_1 \dots z_{n-2}) = 0 \Rightarrow |z_{n-2}| \leq 1 \right] \Leftrightarrow \left[|z_{n-2}| > 1 \Rightarrow \text{either } |z_i| < 1 \text{ for one } i = 1 \dots n-2 \text{ or } P_{n-2}(z_1 \dots z_{n-2}) \neq 0 \right]$

implies $P_{n-2}(\xi_1 \dots \xi_{n-2}) \neq 0 \iff D \neq 0$.

The equation $P_n(z_1 \dots z_n) = 0$ can equivalently be expressed as

$$z_{n-1} = - \frac{A + C z_n}{B + D z_n} =: \omega(z_n) \tag{7}$$

which defines a fractional linear transformation $z_n \mapsto \omega(z_n)$.
We have $\lim_{z_n \to +\infty} \omega(z_n) = z_{n-1} = -\frac{C}{D}$. Using

$$C = \sum_{\substack{S \subset \Delta_{n-1} \\ z_{n-1} \notin S}} z^S \left(\prod_{i \in S \cup \{n\}} \prod_{j \in \Delta_{n-1} \setminus S} A_{ij} \right) \tag{8}$$

and

$$D z_{n-1} = \sum_{\substack{S \subset \Delta_{n-1} \\ z_{n-1} \in S}} z^S \left(\prod_{i \in S \cup \{n\}} \prod_{j \in \Delta_{n-1} \setminus S} A_{ij} \right), \tag{9}$$

we see that we thus have

$$0 = C + D z_{n-1} = \sum_{S \subset \Delta_{n-1}} z^{s} \left(\prod_{i \in S \cup \{n\}} \prod_{j \in \Delta_{n-1} \setminus S} A_{ij} \right). \tag{10}$$

We define $\zeta_i = A_{ni}^{-1}$ and note that (10) is equivalent to

$$\left[\prod_{i=1}^{n-1} A_{ni} \right] P_{n-1}(\zeta_1, \dots, \zeta_{n-1}) = 0. \tag{11}$$

By assumption, we have $|\zeta_1| > 1, \dots, |\zeta_{n-2}| > 1$. In combination with (11) and the induction hypothesis $[|z_1| \geq 1, \dots, |z_{n-2}| \geq 1$ and $P_{n-1}(z_1, \dots, z_{n-1}) = 0 \Rightarrow |z_{n-1}| \leq 1]$ this implies

$$|z_{n-1}| < |\zeta_{n-1}| \leq 1. \tag{12}$$

Suppose now that our proposition does not hold. Then there exist z'_{n-1} and z'_n such that

$$P_n(z_1, \dots, z_{n-2}, z'_{n-1}, z'_n) = 0 \tag{13}$$

and $|z'_{n-1}| \geq 1$, $|z'_n| > 1$. The points z'_{n-1} and z'_n obey $w(z'_n) = z'_{n-1}$.

We also know that $\lim_{z \to +\infty} w(z) = \tilde{z}$ with $|\tilde{z}| < 1$. By continuity we can find z''_n s.t. $|z''_n| > 1$ and

$$w(z''_n) = z''_{n-1} \tag{14}$$

with $|z''_{n-1}| = 1$. Since z''_n and z''_{n-1} are related via (14) we have $P_n(z_1, \dots, z_{n-2}, z''_{n-1}, z''_n) = 0$. We can now repeat the whole argument from the beginning with z_{n-1} replaced by z_i for $i = 1, 2, \dots, n-2$ and thereby obtain $\tilde{z}_1, \dots, \tilde{z}_n$ with

$$P_n(\tilde{z}_1, \dots, \tilde{z}_n) = 0 \tag{15}$$

and $|\tilde{z}_1| = 1, \dots, |\tilde{z}_{n-1}| = 1$, $|\tilde{z}_n| > 1$.

Using $A_{ij} = A_{ji}$ we check the identity

$$P_n(z_1^{-1}, \dots, z_n^{-1}) = z_1^{-1} \cdot \ldots \cdot z_n^{-1}\, P_n(z_1, \dots, z_n). \qquad (16)$$

complex conjugation

In combination with $\overline{\tilde{z}_i} = \tilde{z}_i^{-1}$ for $i = 1 \dots n-1$, this implies

$$P_n(\tilde{z}_1, \dots, \tilde{z}_{n-1}, \overline{\tilde{z}_n^{-1}}) = P_n\left(\overline{\tilde{z}_1^{-1}}, \dots, \overline{\tilde{z}_{n-1}^{-1}}, \overline{\tilde{z}_n^{-1}}\right) \qquad (17)$$

$$\overset{(16)}{=} \overline{\tilde{z}_n^{-1}} \cdot \ldots \cdot \overline{\tilde{z}_n^{-1}}\, \underbrace{P_n(\bar{\tilde{z}}_1, \dots, \bar{\tilde{z}}_n)} = 0$$

$$= \overline{P_n(\tilde{z}_1, \dots, \tilde{z}_n)} \overset{(15)}{=} 0$$

The l.h.s. is a polynomial of degree 1 in $\overline{\tilde{z}_n^{-1}}$, and therefore vanishes only if $\overline{\tilde{z}_n^{-1}} = \tilde{z}_n$, which contradicts $|\tilde{z}_n| > 1$. In this argument we need that the coefficient of z_n in P_n on the r.h.s. of (10) or the l.h.s. of (11) does not equal zero if $|z_1| \geq 1, \dots, |z_{n-1}| \geq 1$. This statement follows from the induction hypothesis. $\blacksquare$

4.1.2 Theorem (Lee, Yang): With the notation and the assumptions of Proposition 5.1.1 we define a polynomial $\mathcal{P}^n$ of degree n in z by

$$\mathcal{P}^n(z) = \mathcal{P}_n(z_1, \dots, z) = \sum_S z^{|S|} \left(\prod_{i \in S} \prod_{j \in S'} A_{ij} \right), \tag{18}$$

where $|S|$ is the number of elements in S. Then the zeros of $\mathcal{P}^n$ all lie on the circle $\{ z \in \mathbb{C} \mid |z| = 1 \}$.

Proof: The statement follows immediately from Proposition 5.1.1. $\boxtimes$

Let us consider a lattice gas with interaction energy

$$\mathcal{U}(X) = N(X) \Phi^1 + \sum_{\{x,y\} \subseteq X} \Phi^2(\{x,y\}), \tag{19}$$

where $X \subseteq \Lambda$. The interpretation of the above energy is that $X = \{x_1, \dots, x_m\}$ is the set of occupied lattice sites (with a particle), Φ^1 equals minus the chemical potential and Φ^2 is a pair interaction. In this model we can have at most one particle per lattice site. As discussed in Section 2 the model also has a spin interpretation. If the spins at $x_1, \dots, x_m$ are up and the other points are down in the region Λ, the corresponding potential energy is

$$U_\Lambda^1(X) = N(X)\left(\Phi^1 + C_{\Phi^2}\right) - \frac{1}{2} \sum_{x \in X} \sum_{y \in \Lambda \setminus X} \Phi^2(\{x, y\}), \qquad (20)$$

where

$$C_{\Phi^2} = \frac{1}{2} \sum_{x \neq 0} \Phi^2(\{0, x\}). \qquad (21)$$

In Section 2 we have shown that under the assumption

$$\|\Phi^2\| = \frac{1}{2} \sum_{x \neq 0} \left|\Phi^2(\{0, x\})\right| < +\infty \qquad (22)$$

we have

$$\mathcal{P}(\Phi) = \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{X\subset\Lambda} \exp\Big[-N(X)\,\Phi^1 - \sum_{\{x,y\}\subseteq X} \Phi^2(\{x,y\})\Big]$$

$$= \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{X\subset\Lambda} \exp\Big[-N(X)\big(\Phi^1 + C_{\Phi^2}\big) + \frac{1}{2}\sum_{x\in X}\ \sum_{y\in\Lambda_\alpha\setminus X} \Phi^2(\{x,y\})\Big]. \qquad (23)$$

Let us define $z = \exp\big[-(\Phi^1 + C_{\Phi^2})\big]$ and $A_{xy} = \exp\big[\frac{1}{2}\Phi^2(\{x,y\})\big]$.

Then we have

$$\sum_{X\subset\Lambda} \exp\Big[-N(X)\big(\Phi^1 + C_{\Phi^2}\big) + \frac{1}{2}\sum_{x\in X}\ \sum_{y\in\Lambda\setminus X} \Phi^{(2)}(\{x,y\})\Big]$$

$$= \sum_{X\subset\Lambda} z^{N(X)} \Big(\prod_{x\in X}\ \prod_{y\in\Lambda\setminus X} A_{xy}\Big) = \mathcal{P}^{N(\Lambda)}(z) \qquad (24)$$

so that

$$\mathcal{P}(\Phi) = \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \mathcal{P}^{N(\Lambda_\alpha)}(z), \quad z>0. \qquad (25)$$

Before we continue, let us recall (or learn) some basic facts about complex analysis.

Intermezzo (complex analysis)

The first statement we recall is Cauchy's integral theorem.

Theorem (Cauchy integral theorem): Let $D \subseteq \mathbb{C}$ be open and simply connected, and let f be analytic on D. Then

$$\oint_\gamma f(z)\,dz = 0, \tag{26}$$

for all closed paths $\gamma \subset D$.

Corollary: Let $D \subseteq \mathbb{C}$ be open, connected and simply connected and let f be analytic in D. Then, there exists an analytic function $F: D \to \mathbb{C}$, such that $F' = f$.

Proof: Fix some point $z_0 \in D$ and define

$$F(z) = \int_{z_0}^{z} f(\xi)\,d\xi, \qquad (27)$$

where the integral is over some path connecting z_0 and z. Any point $z \in D$ can be reached like this and by the previous theorem F does not depend on the path that we choose. The function F is analytic and a short computation shows $F' = f$. $\boxtimes$

Theorem: Let f be an analytic function on a connected and simply connected open set $D \subseteq \mathbb{C}$, which has no zeros in D. Then, there exists an analytic function $g: D \to \mathbb{C}$, called a **branch of the logarithm of f on D**, such that $f = e^{g}$.

Proof: Our assumptions imply that $\frac{f'}{f}$ is analytic on D. The Corollary on the previous page thus implies the existence of a function F, analytic in D, such that $F' = \frac{f'}{f}$. In particular,

$$\frac{d}{dz}\left(f e^{-F}\right) = f' e^{-F} - f \underbrace{F'}_{\frac{f'}{f}} e^{-F} = 0. \qquad (28)$$

Accordingly, there exists a constant $c \in \mathbb{C}$ s.t. $f e^{-F} = e^{c}$
$\Leftrightarrow f = e^{F+c}$.

$\boxtimes$

Remarks: • Let g be a branch of the logarithm of f on D. Then we have

$$|f| = |e^g| = e^{\operatorname{Re} g}, \tag{29}$$

that is, $\operatorname{Re} g = \ln |f|$.

•] Let g_1 and g_2 be two branches of the logarithm of f on D. Since, for each $z \in D$,

$$e^{g_2(z) - g_1(z)} = \frac{e^{g_2(z)}}{e^{g_1(z)}} = \frac{f(z)}{f(z)} = 1, \tag{30}$$

we conclude that $g_2(z) = g_1(z) + 2ik(z)\pi$ for some $k(z) \in \mathbb{Z}$. However, $z \mapsto k(z) = (g_2(z) - g_1(z))/(2\pi i)$ is continuous and integer-valued, that is, constant on D. This implies $g_2 = g_1 + 2ik\pi$ for some $k \in \mathbb{Z}$.

•] Assume that the domain D in the previous theorem is such that $D \cap \mathbb{R}$ is connected. Suppose also that $f(z) \in (0, \infty)$ for $z \in D \cap \mathbb{R}$. Then there is a branch g of the logarithm of f on D s.t. $g(z) \in \mathbb{R}$ for all $z \in D \cap \mathbb{R}$. In particular, g coincides with the usual logarithm of f (seen as a real function) on $D \cap \mathbb{R}$. To see that this is true, it suffices to observe that the function F in the proof can be constructed by starting from a point $z_0 \in D \cap \mathbb{R}$ at which one can fix $g(z_0) = \ln|f(z_0)|$. Since $F(z) = \int_{z_0}^{z} f'(x)/f(x)\, dx$ for $z \in D \cap \mathbb{R}$ this proves the claim.

We recall that a family $\mathcal{F}$ of functions on $\mathbb{C}$ is called **locally uniformly bounded** on a set $D \subseteq \mathbb{C}$ if, for each $z \in D$, there exists a real number M and a neighborhood U of z s.t. $|f(w)| \leq M$ for all $w \in U$ and all $f \in \mathcal{F}$.

Theorem (Vitali convergence Theorem): Let $D \subseteq \mathbb{C}$ be open and connected, and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of analytic functions on D, which are locally uniformly bounded

and converge on a set having a cluster point in D.
Then the sequence $\{f_n\}_{n=1}^{\infty}$ converges locally uniformly on D to an analytic function.

Theorem (Hurwitz theorem): Let D be an open subset of $\mathbb{C}$ and $\{f_n\}_{n=1}^{\infty}$ be a sequence of analytic functions, which converge, locally uniformly, on D to an analytic function f. If $f_n(z) \neq 0$ for all $z \in D$ and for all $n \geq 1$, then either f vanishes identically, or f is never zero on D.

The last statement we need is the following theorem.

Theorem: Let $D \subseteq \mathbb{C}$ be open and let $\{f_n\}_{n=1}^{\infty}$ be a sequence of analytic functions on D, which converges locally uniformly to f. Then f is analytic and f_n' converges locally uniformly to f'.

This concludes our intermezzo about complex analysis. We continue with the analysis of our lattice system.

If $\Phi^2 \leq 0$ then we have $0 < A_{xy} \leq 1$ and Theorem 5.1.2. applies. It shows that all zeros of $P^{N(\Lambda)}(z)$ obey $|z| = 1$. Accordingly, $P^{N(\Lambda_\alpha)}(z)$ is an analytic function on $D = \{z \in \mathbb{C} \mid |z| < 1\}$ with no zeros and the second theorem in our intermezzo allows us to define $\left[P^{N(\Lambda_\alpha)}(z)\right]^{\frac{1}{N(\Lambda_\alpha)}}$. To define these functions we use a branch of the logarithm s.t. $\left[P^{N(\Lambda_\alpha)}(z)\right]^{\frac{1}{N(\Lambda_\alpha)}}$ is real for $z \in D \cap \mathbb{R}$ and coincides with the usual roots viewed as real functions (see third point of remark in intermezzo). We have

$$\left| \exp\left[\frac{1}{N(\Lambda_\alpha)} \ln\left(P^{N(\Lambda_\alpha)}(z)\right)\right] \right| \tag{31}$$

$$\leq \exp\left[\frac{1}{N(\Lambda_\alpha)} \left|\underbrace{\mathrm{Re} \ln\left(P^{N(\Lambda_\alpha)}(z)\right)}_{\substack{= \ln\left(\left|P^{N(\Lambda_\alpha)}(z)\right|\right) \leq \ln\left(P^{N(\Lambda_\alpha)}(|z|)\right). \\ \uparrow \\ \text{first part remark}}}\right| \right]$$

That is, we have

$$\left| \left[\varphi^{N(\lambda_\alpha)}(z) \right]^{\frac{1}{N(\lambda_\alpha)}} \right| \le \left[\varphi^{N(\lambda_\alpha)}(|z|) \right]^{\frac{1}{N(\lambda_\alpha)}}$$

$$< \left[\varphi^{N(\lambda_\alpha)}(1) \right]^{\frac{1}{N(\lambda_\alpha)}} \qquad (32)$$

for all $z \in D$. We conclude that $\left[\varphi^{N(\lambda_\alpha)}(z) \right]^{\frac{1}{N(\lambda_\alpha)}}$ is a family of analytic function in D that is uniformly bounded by (32) and the convergence in (25) (we recall that the results in Section 2 show that the limit in (25) is finite). Moreover, by (25) it converges for $z \in D \cap (0,1)$. Vitali's theorem thus implies that the sequence $\left[\varphi^{N(\lambda_\alpha)}(z) \right]^{\frac{1}{N(\lambda_\alpha)}}$ converges uniformly to an analytic function $\lambda(z)$ on any disc $\{ z \in \mathbb{C} \mid |z| \le K \}$ with $0 < K < 1$ as $\alpha \to \infty$. Since each element of the sequence has no zeros in D we can apply Hurwitz theorem and conclude that also $\lambda(z)$ has no zeros in

D. The function $\lambda(z)$ cannot vanish identically because $\mathcal{Z}$ is strictly positive if $z \in D \cap (0,1)$. This shows that $P(\Phi)$ may be continued from the interval $0<z<1$ to an analytic function $p(z) = \ln \lambda(z)$ with $z \in D$.

With the identity

$$\mathcal{P}^{N(\Lambda_a)}(z^{-1}) = z^{-N(\Lambda_a)}\, \mathcal{P}^{N(\Lambda_a)}(z), \tag{33}$$

which follows from (16), we see that $P(\Phi)$ may also be continued from the interval $1<z<+\infty$ to the analytic function $p(z^{-1}) + \ln z$ with $z \in \tilde{D} = \{z \in \mathbb{C} \mid |z|>1\}$. We summarize our findings in the following theorem.

4.1.3. Theorem: If a lattice gas has a pair interaction as in (19) with fixed negative pair potential $\Phi^2 \leq 0$, $\|\Phi^2\| < +\infty$, and if we put

$$z = \exp\left[-\Phi^1 - \frac{1}{2}\sum_{x\neq 0} \Phi^2(\{0,x\})\right] \qquad (34)$$

then the thermodynamic function defined by

$$P(\Phi) = \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{X\subset\Lambda_\alpha} \exp[-U(X)] \qquad (35)$$

with $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of van Hove extends from the interval $0<z<1$ to a function $p(z)$ analytic in $\{z\in\mathbb{C} \mid |z|<1\}$, and from the interval $1<z<+\infty$ to $p(z^{-1}) + \ln(z)$, which is analytic in $\{z\in\mathbb{C} \mid |z|>1\}$.

Therefore a phase transition can occur at most for $z=1$.

Remark: •] We note that a factor β has been absorbed in Φ^1 and Φ^2 and that z is the activity up to the fixed factor $\exp\left[-\frac{1}{2}\sum_{x\neq 0} \Phi^2(\{0,x\})\right]$.

•] In the last sentence of the theorem we disregard the dependence of $p(z)$ on β.

4.2. Existence of a first-order phase transition at low temperatures

In this section we consider a lattice gas with a pair interaction. Theorem 4.1.3. shows that, if the pair potential Φ^2 is negative, a phase transition can occur only at $z=1$. We shall prove that for a suitable class of pair potentials a phase transition does occur at $z=1$ if the temperature β^{-1} is low enough.

A phase transition is said to be of n-th order if the n-th derivative of the thermodynamic function is discontinuous at the transition. In our case the density, which can be written in terms of the first derivative of the pressure, has a jump. Accordingly, the phase transition is of first order. In the lattice gas interpretation of the model we consider, the phase transition can be interpreted as a liquid-vapor phase transition (water has a larger density than steam). The jump in the density occurs when we vary the chemical potential, which is equivalent to saying that we vary the activity z. In the spin interpretation of the model the density equals the fraction of up and down spins. This fraction has a jump as a function of the external magnetic field. Here one should think of a magnetized sample (non-zero net magnetization).

If one varies the magnetic field from $-\varepsilon$ to $+\varepsilon$ for some $\varepsilon > 0$ then the net magnetization will jump from $m(-\varepsilon)$ to $m(\varepsilon)$, where $m(-\varepsilon) = -m(\varepsilon)$.

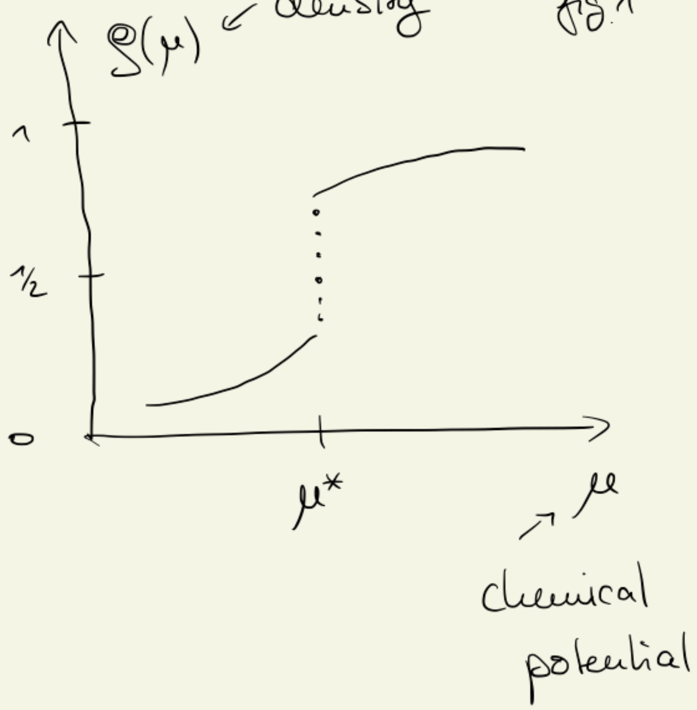


fig.1

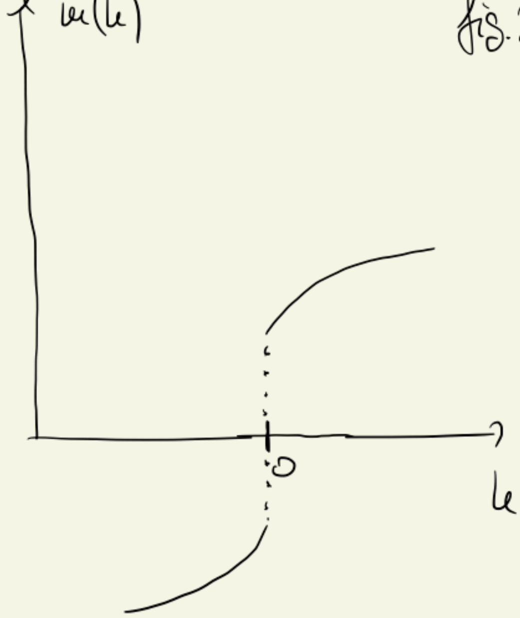


fig.2

A much more precise description of these two phase transitions can be found in Sections 3 and 4 in the book by Friedli and Velenik [2]. Our concern here will be to prove the following theorem.

4.2.1. Theorem: We consider a lattice system with pair interactions, i.e.,

$$\mathcal{U}(X) = N(X)\,\Phi^1 + \sum_{\{x,y\}\subseteq X} \Phi^2(\{x,y\}) \tag{36}$$

and we write

$$z = \exp\left[-\beta\Phi^1 - \frac{\beta}{2}\sum_{x\neq 0}\Phi^2(\{0,x\})\right] \tag{37}$$

$$p(z,\beta) = \frac{1}{\beta}\lim_{\alpha\to\infty}\frac{1}{N(\Lambda_\alpha)}\ln\sum_{X\subset\Lambda_\alpha}\exp\left[-\beta\,\mathcal{U}(X)\right], \tag{38}$$

where $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove. Let

$$\Phi_i = \Phi^2(\{0, e_i\}), \qquad 1 \le i \le d \tag{39}$$

$$e_i = (0,\dots,0,\overset{\text{i-th position}}{1},0,\dots,0)$$

as well as

$$D_i = \frac{1}{2}\sum_x{}' \,|x^i|\cdot\left|\Phi^2(\{0,x\})\right|, \tag{40}$$

where $\sum'$ extends over all $x \in \mathbb{Z}^d$ except 0 and the $2d$ nearest neighbors of 0.

If we have $d \geq 2$ and

$$\Phi_i + D_i < 0 \quad \text{for} \quad i = 1 \ldots d, \tag{41}$$

then a first-order phase transition occurs at $z = 1$ for sufficiently small β^{-1}.

The remaining part of this section is dedicated to the proof of the above Theorem. Let us define the canonical partition function

$$Z(\Lambda, n) = \sum_{\substack{S \subseteq \Lambda \\ N(S) = n}} \exp\left[\frac{\beta}{2} \sum_{x \in S} \sum_{y \in \Lambda \setminus S} \Phi^2(\{x,y\})\right]. \tag{42}$$

The corresponding grand canonical partition function is given by

$$\mathcal{P}^{N(\Lambda)}(z) = \sum_{n=0}^{N(\Lambda)} z^n \, Z(\Lambda, n). \qquad (43)$$

The thermodynamic potential related to $\mathcal{P}^{N(\Lambda)}(z)$ appears in the above Theorem and the one related to $Z(\Lambda, n)$ will appear in our proof. Their existence and some of their properties are provided by the following Proposition.

4.2.2. Proposition: Assume that $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove.

(a) If $\frac{n_\alpha}{N(\Lambda_\alpha)} \to \rho$ with $0 \le \rho \le 1$, then the limit

$$g(\rho) = \frac{1}{\beta} \lim_{\alpha \to \infty} \frac{1}{N(\Lambda_\alpha)} \ln Z(\Lambda_\alpha, n_\alpha) \qquad (44)$$

exists and does not depend on the sequence $\{\Lambda_\alpha\}_{\alpha=1}^{\infty}$.

The function g is positive and concave (and therefore continuous) and satisfies

$$g(1-\rho) = g(\rho). \tag{45}$$

(b) If $z>0$ then $p(z,\beta)$ satisfies

$$p(z,\beta) = \frac{1}{\beta} \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \Xi^{N(\Lambda_\alpha)}(z) \tag{46}$$

$$= \max_{0\le\rho\le 1} \left\{ \beta^{-1}\rho \ln(z) + g(\rho) \right\}$$

as well as

$$p(z,\beta) = \beta^{-1} \ln(z) + p(z^{-1},\beta). \tag{47}$$

Remarks: •] From a physics point of view g equals minus the free energy of the system.

•] The symmetry $g(\rho) = g(1-\rho)$ is not a typical feature of lattice gases and rather related to the spin interpretation of the model.

Proof: The existence of the functions g and ρ and the fact that they are related via a Legendre transformation can be proved with the techniques we discussed in Section 2 and 3. We therefore refrain from presenting the details.

The density is naturally between 0 and 1 because we are in a discrete setting and the extremes are the empty and the fully occupied lattice.

To see that (45) is true we choose a sequence $\{\Lambda_\alpha\}_{\alpha=1}^{\infty}$ with $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove and let $n_\alpha = \rho |\Lambda_\alpha|$ with $0 \leq \rho \leq 1$. When we consider $Z(n_\alpha, \Lambda_\alpha)$ we see that its value for ρ and $(1-\rho)$ are equal

In fact, for each α, changing ρ to $(1-\rho)$ amounts to changing the roles of the sums over x and y (Note that $\Phi^2(\{x,y\})$ does not depend on the order of x and y.). Eq. (47) follows from the identity for $p^{N(\lambda)}(z)$ in (33). ∎

When we choose $u_\alpha = 0$ for $\alpha \geq 0$, we have $Z(\lambda_\alpha, u_\alpha) = 1$ for all $\alpha \geq 0$, and hence

$$g(0) = 0. \tag{48}$$

Typical graphs of $g(\rho)$ thus like like this:

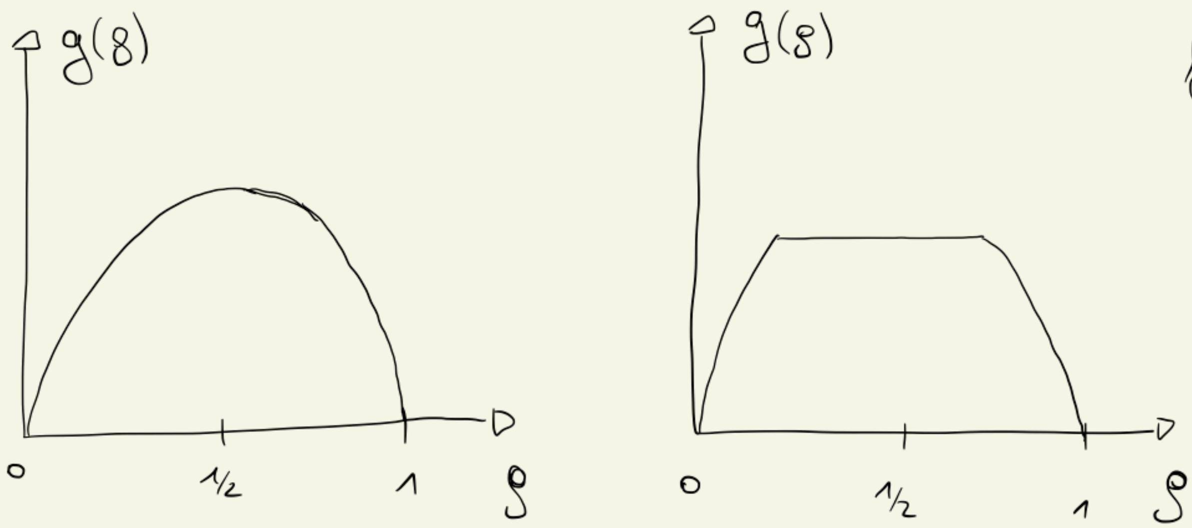


fig. 3

To prove Theorem 4.2.1. we will show that, for sufficiently small β^{-1}, a horizontal segment is present in the graph of $g(\varrho)$. As we will argue now, this implies the existence of a first order phase transition at $z=1$. First, we note that the derivative of $p(z,\beta)$ w.r.t. z is related to the density of the system. To see this, we invoke the second theorem on p. 17 and compute ($z \neq 1$):

$$\beta z \frac{\partial p(z,\beta)}{\partial z} = z \frac{\partial}{\partial z} \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{n=0}^{N(\Lambda_\alpha)} z^n \, Z(\Lambda_\alpha, n) \qquad (49)$$

$$= \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \underbrace{\frac{\sum_{n=0}^{N(\Lambda_\alpha)} n z^n \, Z(\Lambda_\alpha, n)}{\sum_{n=0}^{N(\Lambda_\alpha)} z^n \, Z(\Lambda_\alpha, n)}}_{\text{expectation of the particle number in the grand canonical ensemble}} =: \bar{\varrho}(z,\beta)$$

with the expected density $\bar{\rho}(z,\beta)$ of the grand canonical ensemble.

If ρ^* is the point, where the max in (46) is attained then we have $(z \neq 1)$

without a formal justification we assume that this derivative exists. Note that it exists p.w. a.e. because g is concave

$$\beta^{-1} \ln(z) = -g'(\rho^*) \tag{50}$$

generalized inverse function in the sense of Section 3, p. 29

$$\Leftrightarrow \quad \rho^* = \left(-g'\right)^{-1}\left(\beta^{-1}\ln(z)\right)$$

as well as

$$p(z,\beta) = \beta^{-1}\ln(z)\left(-g'\right)^{-1}\left(\beta^{-1}\ln(z)\right) + g\left(\left(-g'\right)^{-1}\left(\beta^{-1}\ln(z)\right)\right). \tag{51}$$

The derivative of p w.r.t. $\beta^{-1}\ln(z)$ is therefore given by

Also here we do not discuss why we are allowed to take the derivatives

$$\frac{\partial p(z,\beta)}{\partial(\beta^{-1}\ln(z))} = (-g')^{-1}\left(\beta^{-1}\ln(z)\right) + \beta^{-1}\ln(z)\left((-g')^{-1}\right)'\left(\beta^{-1}\ln(z)\right) \tag{52}$$
$$+ g'\left((-g')^{-1}\left(\beta^{-1}\ln(z)\right)\right)\left((-g')^{-1}\right)'\left(\beta^{-1}\ln(z)\right)$$

$$\overset{(50)}{=} (-g')^{-1}\left(\beta^{-1}\ln(z)\right).$$

In particular,

$$\frac{\partial p(z,\beta)}{\partial z} = \frac{\partial p(z,\beta)}{\partial(\beta^{-1}\ln(z))}\,\frac{\partial(\beta^{-1}\ln(z))}{\partial z} = (-g')^{-1}\left(\beta^{-1}\ln(z)\right)\frac{1}{\beta z}. \tag{53}$$

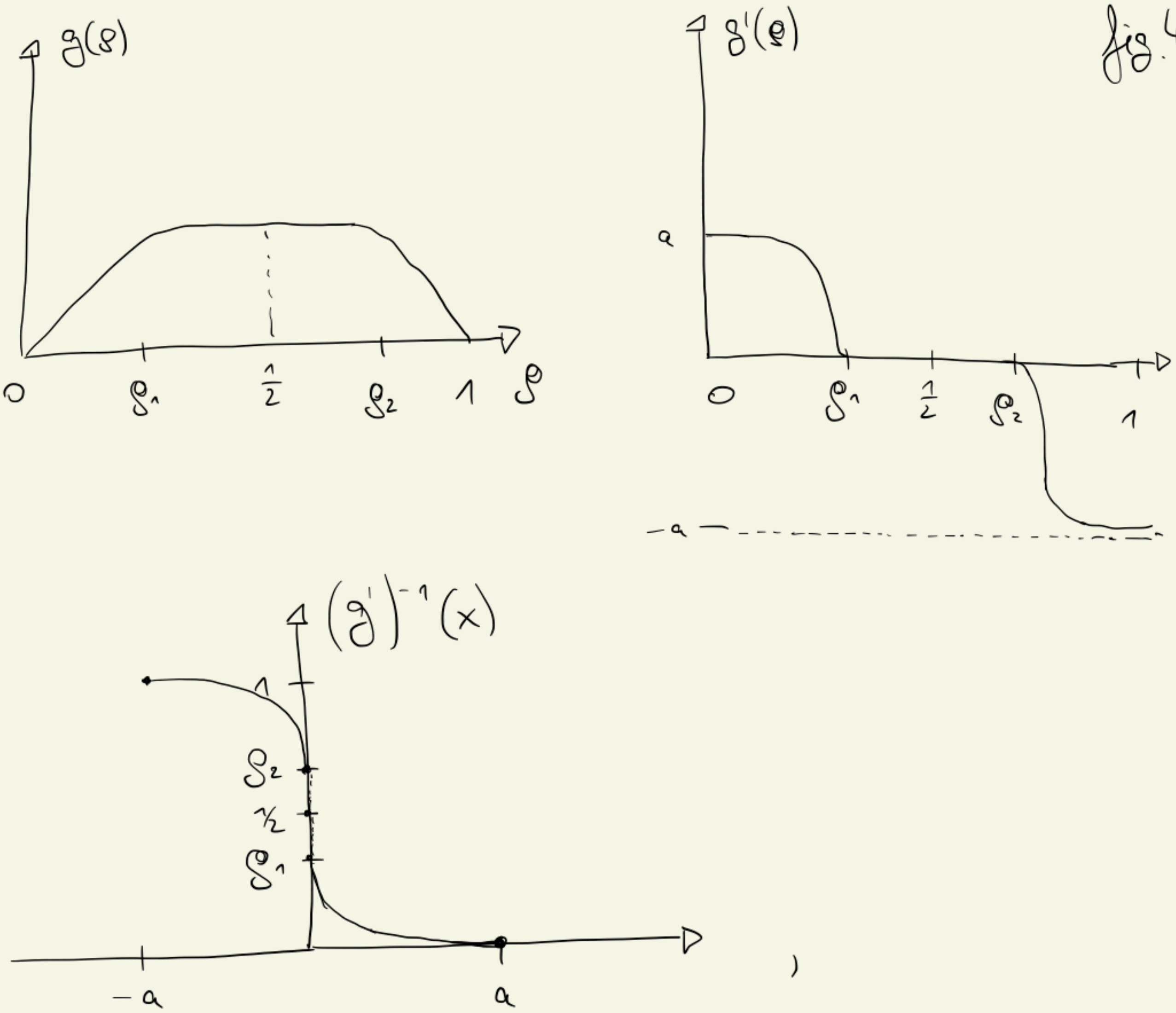
In combination with (49), this implies

$$\bar{\rho}(z,\beta) = \beta z\,\frac{\partial p(z,\beta)}{\partial z} = (-g')^{-1}\left(\beta^{-1}\ln(z)\right). \tag{54}$$

It remains to argue that $\bar{\rho}(z,\beta)$ has a jump at $z=1$ if $g(\rho)$ has a plateau.

In this case we have:

fig. 4



That is, $(g')^{-1}$ has a jump at $x=0$, and hence

$\bar{g}(z,\beta)$ has a jump at $\ln(z)=0 \iff z=1$.

The first step in the proof of Theorem 5.3.1 is the following Proposition.

4.2.3. Proposition: Assume that $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove. For each Λ_α, we choose a set $\mathcal{M}_\alpha$ of subsets of Λ_α and define

$$\tilde{Z}(\Lambda_\alpha, n) = \sum_{S\in\mathcal{M}_\alpha,\ N(S)=n} \exp\left[\frac{\beta}{2} \sum_{x\in S} \sum_{y\in\Lambda_\alpha\setminus S} \Phi^2(\{x,y\})\right]. \tag{55}$$

If

$$\beta^{-1} \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n) = p(\lambda, \beta) \tag{56}$$

and

$$\liminf_{\alpha\to\infty} \frac{\frac{1}{N(\Lambda_\alpha)} \sum_{n=0}^{\infty} n\,\tilde{Z}(\Lambda_\alpha, n)}{\sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n)} \le \varrho_0 < \frac{1}{2}, \tag{57}$$

then g reduces to a constant on the interval $[\varrho_0, (1-\varrho_0)]$.

Proof: For given $\varepsilon > 0$ we may choose a sequence $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove s.t.

$$\frac{\frac{1}{N(\Lambda_\alpha)} \sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n)\, n}{\sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n)} \leq \varrho_0 + \varepsilon/2 \tag{58}$$

holds for $\alpha \geq \alpha_0$. Then we have

$$\frac{\sum_{n \geq N(\Lambda_\alpha)(\varrho_0+\varepsilon)} \tilde{Z}(\Lambda_\alpha, n)\,(\varrho_0+\varepsilon)}{\sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n)} \leq \varrho_0 + \varepsilon/2 \tag{59}$$

as well as

$$\frac{\varepsilon/2}{\varrho_0+\varepsilon} \leq \frac{\sum_{n<N(\lambda_\alpha)(\varrho_0+\varepsilon)} \tilde{Z}(\lambda_\alpha,n)}{\sum_{n=0}^{\infty} \tilde{Z}(\lambda_\alpha,n)} \leq 1. \tag{60}$$

In particular,

$$\lim_{\alpha\to\infty} \frac{1}{N(\lambda_\alpha)} \ln \sum_{n<N(\lambda_\alpha)(\varrho_0+\varepsilon)} \tilde{Z}(\lambda_\alpha,n) \tag{61}$$

$$\geq \lim_{\alpha\to\infty} \frac{1}{N(\lambda_\alpha)} \ln\left(\frac{\varepsilon/2}{\varrho_0+\varepsilon} \sum_{n=0}^{\infty} \tilde{Z}(\lambda_\alpha,n)\right)$$

$$= \lim_{\alpha\to\infty} \frac{1}{N(\lambda_\alpha)} \ln \sum_{n=0}^{\infty} \tilde{Z}(\lambda_\alpha,n) = \beta p(1,\beta).$$

In combination with the trivial bound

$$\lim_{\alpha\to\infty} \frac{1}{N(\lambda_\alpha)} \ln \sum_{n<N(\lambda_\alpha)(\varrho_0+\varepsilon)} \tilde{Z}(\lambda_\alpha,n) \leq \lim_{\alpha\to\infty} \frac{1}{N(\lambda_\alpha)} \ln \sum_{n=0}^{\infty} \tilde{Z}(\lambda_\alpha,n) \tag{62}$$

$$= \beta p(1,\beta),$$

This proves

$$\lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{n<N(\Lambda_\alpha)(\rho_0+\varepsilon)} \tilde{Z}(\Lambda_\alpha, n) = \beta p(1,\beta). \tag{63}$$

We claim that if $n_\alpha < N(\Lambda_\alpha)(\rho_0+\varepsilon)$ is chosen s.t.

$$\tilde{Z}(\Lambda_\alpha, n_\alpha) = \max_{n<N(\Lambda_\alpha)(\rho_0+\varepsilon)} \tilde{Z}(\Lambda_\alpha, n), \tag{64}$$

then

$$\lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \tilde{Z}(\Lambda_\alpha, n_\alpha) = \beta p(1,\beta). \tag{65}$$

Indeed, (63) and (64) imply

$$\begin{aligned}\beta p(1,\beta) &\leq \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \left[\tilde{Z}(\Lambda_\alpha, n_\alpha) \left(\sum_{n<N(\Lambda_\alpha)(\rho_0+\varepsilon)} 1\right)\right] \\ &= \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \tilde{Z}(\Lambda_\alpha, n_\alpha) \qquad (66) \\ &\quad + \underbrace{\lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln\big(N(\Lambda_\alpha)(\rho_0+\varepsilon)\big)}_{=0}.\end{aligned}$$

But we also have the trivial bound

$$\lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \underbrace{\tilde{Z}(\Lambda_\alpha, n_\alpha)}_{\le \sum_{n=0}^{\infty} \tilde{Z}(\Lambda_\alpha, n)} \le \beta\, p(1,\beta). \tag{67}$$

In combination, (66) and (67) prove (65).

We know that $0 \le n_\alpha < N(\Lambda_\alpha)(\rho_0+\varepsilon)$. By going to a subsequence, we can assume that

$$\frac{n_\alpha}{N(\Lambda_\alpha)} \xrightarrow{\alpha\to\infty} \rho_1 \le \rho_0 + \varepsilon . \tag{68}$$

In combination with (44), (46) and (65), this implies

$$\underset{}{p(1,\beta)} \underset{(44),(65)}{=} g(\rho_1) \underset{(68)}{\le} g(\rho_0+\varepsilon) \underset{(57)}{\le} g(\tfrac{1}{2}) \underset{(46)}{=} p(1,\beta) \tag{69}$$

and proves the claim.

$\boxtimes$

In the next step in our proof of Theorem 4.2.1 we make a choice for the sets $\mathcal{U}_\alpha$. We are also free to choose the sequence Λ_α as long as $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove. We define

$$\Lambda_\alpha = \left\{ x \in \mathbb{Z}^d \mid 0 \le x_i < \alpha \text{ for } i = 1 \dots d \right\} \tag{70}$$

as well as

$$\Lambda'_\alpha = \left\{ x \in \mathbb{Z}^d \mid 1 \le x_i < \alpha - 1 \text{ for } i = 1 \dots d \right\}. \tag{71}$$

The sets $\mathcal{U}_\alpha$ are chosen as the sets of all subsets of Λ'_α. With this choice, Proposition 4.2.3 can be interpreted as follows. We modify the grand canonical ensemble (at $z = 1$) by introducing a boundary effect (i.e. requiring that the sites on the boundary of Λ_2 be empty). If this tiny modification produces a change in the density (with ½ being replaced by a value less than or equal to $\rho_0 < 1/2$), then a first order phase transition occurs.

Let us explain the intuition behind this in some more detail. Without the restriction to the sets $\mathcal{U}_\alpha$, the partition function in (55) has a symmetry in the sense that the sums over x and y in the exponent appear in a symmetric way. In the spin interpretation this is related to the fact that there is zero external magnetic field if $z=1$ and that the model has a spin flip symmetry (the energy of a configuration stays the same if we flip all spins). If we introduce a symmetry breaking term (as above) this picture changes. Above a certain critical temperature the model will still have zero magnetization (aka density ½) in the thermodynamic limit. If the temperature is sufficiently low the symmetry breaking terms will lead to a macroscopic magnetization (density ≠ ½), which indicates a phase transition.

In the lattice gas interpretation one should interpret the plateau in fig. 3 as a signature for the co-existence of a gas and a liquid phase.

That (56) holds for our choice of λ_α and u_α can easily be proved with the techniques we introduced in Section 2. We omit the details. We shall thus have proved Theorem 4.2.1 if we can verify (57). In order to do this, we introduce some new terminology and notation.

We embed the set $\mathbb{Z}^d$ in the natural way in $\mathbb{R}^d$. Given $S \in \mathfrak{M}_\alpha \Leftrightarrow S \subseteq \Lambda'_\alpha$, we draw around each $x \in S$ the $2d$ faces of the unit cube centered at x and suppress the faces, which occur twice. The closed polyhedron obtained in this way is called $\Gamma(S)$. Each face of $\Gamma(S)$ separates a point $x \in S$ and a point $y \notin S$ in $\mathbb{Z}^d$. Along a $(d-2)$-dimensional edge of $\Gamma(S)$ either two or four faces meet. In the case of four faces we slightly deform the polyhedron, "chopping off" the edge from the cubes containing a point of S, see fig. 5. When this is done $\Gamma(S)$ splits into connected components $\gamma_1, \dots, \gamma_r$, which we call **cycles**. We have $\gamma_i \cap \gamma_j = \emptyset$ if $i \neq j$ and $\Gamma(S) = \bigcup_{j=1}^{r} \gamma_j$.

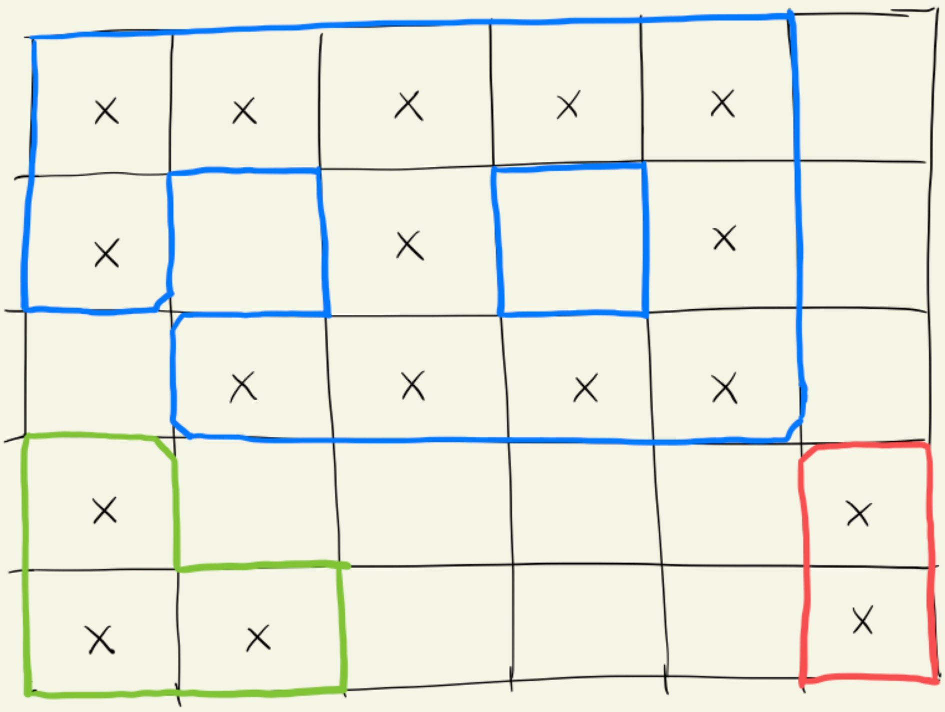


fig. 5

x = occupied sites

$\gamma_1, \gamma_2, \gamma_3$

Given a cycle γ, we denote by $N(\gamma)$ the number of points of $\mathbb{Z}^d$ inside of γ and by $|\gamma|_i$ the number of faces of γ orthogonal to the i-th coordinate axis. By construction we have $N(S) = \sum_{j=1}^{r} N(\gamma_j)$ if $\gamma_1, \dots, \gamma_r$ are the cycles that constitute S.

We call the origin of a cycle the point that is closest to the point $O \notin \Lambda'_L$. If an ambiguity occurs we take the point with the smallest d-component, smallest $(d-1)$-component and so on.

In the next step we prove three lemmas involving the above objects that we will afterwards use to prove our theorem.

Lemma 4.2.5.: The number of cycles $\gamma \subseteq \Lambda_\alpha'$ with prescribed $|\gamma|_1, \dots, |\gamma|_d$ is smaller than $D(\Lambda_\alpha)$ $\prod_{i=1}^{3} 3^{|\gamma|_i - 1}$.

Proof: We start by fixing a point $x \in \Lambda_\alpha'$ and only consider those cycles, whose origin is given by x. We think of building up the cycle face by face and realize that there are three ways to attach a new face, see fig. 6. This is obvious in dimension two and three but also not difficult to see for $d \geq 4$. This explains the basis 3 in the claimed bound.

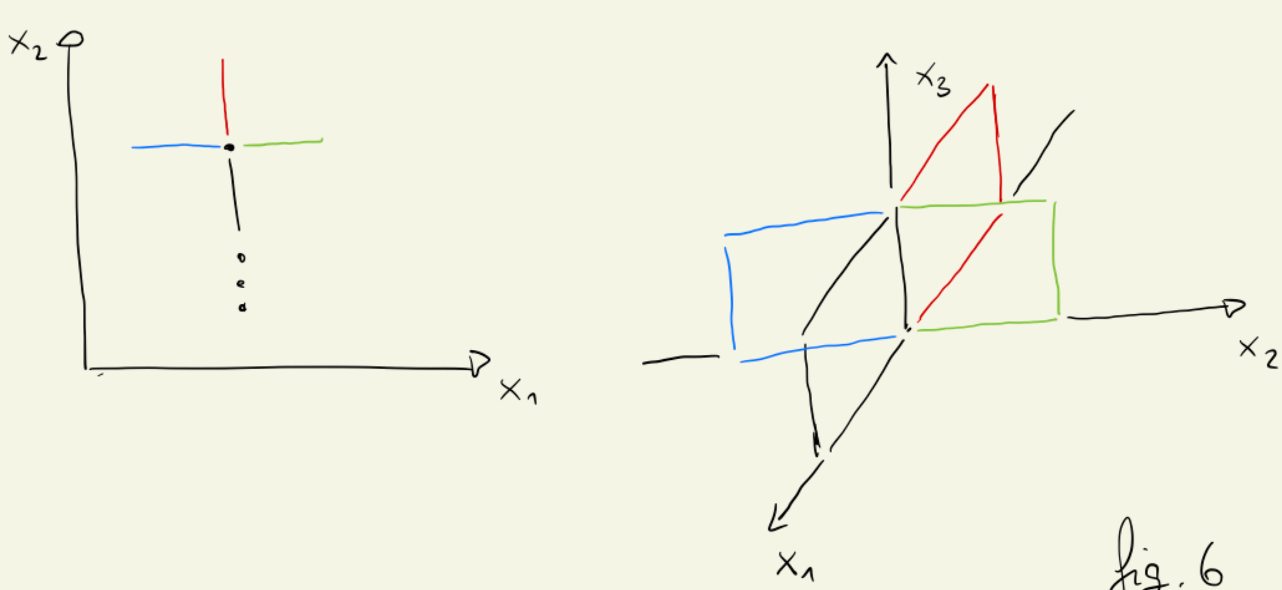


fig. 6

Option 1, Option 2, Option 3.

That $|\gamma|_i - 1$ appears in the exponent rather than $|\gamma|_i$ is because the origin of γ is fixed. The number $N(\Lambda_\alpha)$ comes about because $x \in \Lambda_\alpha' \subset \Lambda_\alpha$. This proves the claim.

Lemma 4.2.6.: Let $\gamma \in \mathcal{M}_\alpha$ be a cycle. We have

$$N(\gamma) \leq \prod_{i=1}^{d} \left(\frac{1}{2}|\gamma|_i\right)^{\frac{1}{d-1}}. \qquad (72)$$

Proof: The number $N(\gamma)$ equals (up to the chopped off edges that we put back in for this proof) the volume surrounded by γ. This volume is bounded by the volume of the rectangle, whose faces have areas (volume) $\frac{1}{2}|\gamma|_1, \dots, \frac{1}{2}|\gamma|_d$, see fig. 7.

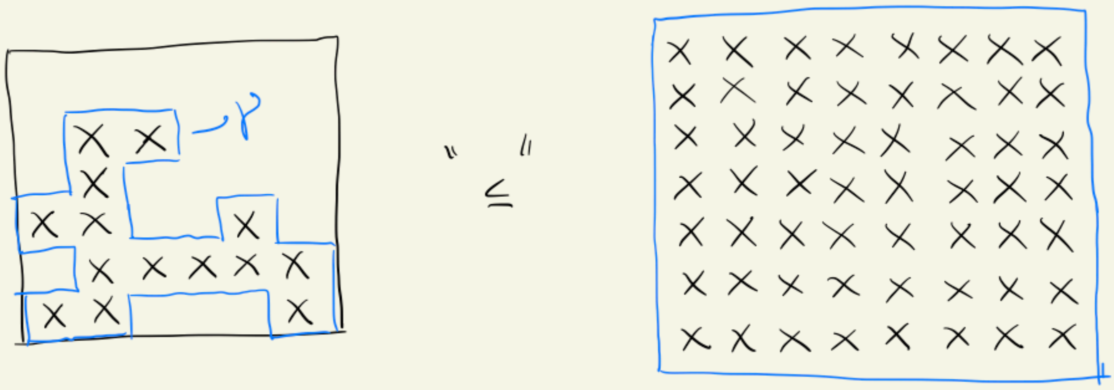


fig. 7

$N(\gamma) = 14$ $\qquad$ $N(\cdots) = 8\cdot 7 = 56$

$|\gamma|_1 = 16, \ |\gamma|_2 = 14$

As in the previous proof, this is obvious in $d=2,3$ and it is not difficult to prove for $d \geq 4$. $\blacksquare$

4.2.7. Lemma: Let $S \in \mathcal{M}_\alpha$ and define

$$Z(S) = \exp\Big(\frac{\beta}{2} \sum_{x \in S} \sum_{y \in \Lambda \setminus S} \Phi^2(\{x,y\})\Big). \tag{73}$$

Let $\Gamma(S)$ consist of the cycles $\gamma_1, \dots, \gamma_r$ and $\Gamma(S')$ of $\gamma_1 \dots \gamma_{r-1}$, then

$$\frac{Z(S)}{Z(S')} \leq \exp\Big[\frac{\beta}{2} \sum_{i=1}^{d} |\gamma_r|_i \big(\Phi_i + D_i\big)\Big]. \tag{74}$$

Proof: We have $\Gamma(S) = \Gamma(S') \cup \gamma_r$ (where we understand the union in such a way that we disregard faces that appear twice). Two points $x, y \in \Lambda_\alpha$, which are both inside or both outside of γ_r, yield the same contribution to $Z(S)$ and $Z(S')$. We therefore only need to consider interactions between points inside γ and points outside of γ. Each face of γ_r separates two nearest neighbor points $x, y \in \Lambda_L$ s.t. $x \in S$ and $y \notin S$. The following two situations can occur: Either $x \in S'$ and $y \notin S'$ or $x, y \notin S'$. Accordingly, we obtain the contribution $\exp\left[\frac{\beta}{2} \sum_{i=1}^{d} |\gamma_r|_i \, \Phi_i\right]$ (this comes from the second option, the first does not contribute).

It remains to consider the contribution from the interaction between points that are not nearest neighbors. The number of ways, in which a vector $\xi = (\xi_1 \dots \xi_d)$ may occur as the difference $\xi = x - y$ or $\xi = y - x$ with x contained in γ and y not contained in γ is bounded by

$$\sum_{i=1}^{d} |\gamma|_i \, |\xi_i| . \tag{75}$$

The reasoning behind this bound is the following. The straight line between x and y crosses a face of γ. If this face is parallel to the i-th coordinate axis, then there are not more than $|\gamma|_i \, |\xi_i|$ ways, in which this may happen (we need to cross one of the $|\gamma|_i$ faces and $|\xi_i|$ is bounding the number of ways how this can be done). Therefore, the interaction between pairs of particles that are not nearest neighbors yields a contribution that is bounded by

$$\exp\left[\frac{\beta}{2} \sum_{\xi}{}' \, |\Phi(\xi)| \sum_{i=1}^{d} |\gamma_r|_i \, |\xi_i| \right] \tag{76}$$

sum excluding 0 and the d nearest neighbors of 0.

$$\overset{(40)}{=} \exp\left[\frac{\beta}{2} \sum_{i=1}^{d} |\gamma_r|_i \, D_i \right].$$

This proves the claim. $\blacksquare$

We are now prepared for the proof of (57) given our choice for $\mathfrak{M}_\alpha$. Let $\gamma \subseteq \Lambda'_\alpha$ be a cycle and define

$$Z_\alpha(\gamma) = \sum_{\substack{S \in \mathfrak{M}_\alpha, \\ \gamma \subseteq \Gamma(S)}} Z(S) \tag{77}$$

with $Z(S)$ in (73). We have

$$\frac{1}{N(\Lambda_\alpha)} \sum_{S \in \mathfrak{M}_\alpha} Z(S)\, N(S) = \frac{1}{N(\Lambda_\alpha)} \sum_{S \in \mathfrak{M}_\alpha} Z(S) \sum_{j=1}^{r} N(\gamma_j)$$

$$\Gamma(S) = \dot{\bigcup_{j=1}^{r}} \gamma_j \quad \text{disjoint union}$$

Two cycles γ_1 and γ_2 in this sum need not obey $\gamma_1 \cap \gamma_2 = \phi$.

$$\leq \frac{1}{N(\Lambda_\alpha)} \sum_{S \in \mathfrak{M}_\alpha} Z(S) \sum_{\gamma \subseteq \Gamma(S)} N(\gamma)$$

$$= \frac{1}{N(\Lambda_\alpha)} \sum_{\gamma \subseteq \Lambda'_\alpha} N(\gamma)\, Z_\alpha(\gamma). \tag{78}$$

We also claim that

$$\frac{Z_\alpha(\gamma)}{\sum_{S' \in \mathcal{M}_\alpha} Z(S')} \le \exp\left[\frac{\beta}{2}\sum_{i=1}^{d} |\gamma|_i \left(\Phi_i + D_i\right)\right]. \tag{79}$$

To see that this is true, we note that

$$\sum_{S' \in \mathcal{M}_\alpha} Z(S') \geq \sum_{\substack{S' \in \mathcal{M}_\alpha : \\ \Gamma(S') \cap \gamma = \emptyset}} Z(S'). \tag{80}$$

When we apply Lemma 4.2.7. in the numerator of the left side of (79) and (80) in the denominator, this proves (79).

When we combine (78) and (79) we have

$$\frac{\frac{1}{N(\Lambda_\alpha)} \sum_{S \in \mathcal{M}_\alpha} Z(S)\, N(S)}{\sum_{S \in \mathcal{M}_\alpha} Z(S)} \overset{(78),\,(79)}{\leq} \frac{1}{N(\Lambda_\alpha)} \sum_{\gamma \subseteq \Lambda'_\alpha} N(\gamma) \exp\left[\frac{\beta}{2} \sum_{i=1}^{d} |\gamma|_i \left(\Phi_i + D_i \right) \right]$$

$$\underset{\text{Lemma 4.2.6}}{\leq} \frac{1}{N(\Lambda_\alpha)} \sum_{\gamma \subseteq \Lambda'_\alpha} \prod_{i=1}^{d} \left(\frac{1}{2} |\gamma|_i \right)^{\frac{1}{d-1}} \exp\left[\frac{\beta}{2} |\gamma|_i \left(\Phi_i + D_i \right) \right]. \tag{81}$$

To bound the r.h.s. we write the sum over $\gamma \subseteq \Lambda'_\alpha$ as

$$\sum_{l_1=1}^{\infty} \cdots \sum_{l_d=1}^{\infty} \sum_{\substack{\gamma \subseteq \Lambda'_\alpha : \\ |\gamma|_i = 2 l_i \text{ for } i=1 \ldots d}} \prod_{i=1}^{d} \left(\frac{1}{2} |\gamma|_i \right)^{\frac{1}{d-1}} \exp\left[\frac{\beta}{2} |\gamma|_i \left(\Phi_i + D_i \right) \right]$$

$$\underset{\text{Lemma 4.2.5}}{\leq} N(\Lambda_\alpha) \sum_{l_1=1}^{\infty} \cdots \sum_{l_d=1}^{\infty} \prod_{i=1}^{d} l_i^{\frac{1}{d-1}}\, 3^{l_i - 1} \exp\left[\beta l_i \left(\Phi_i + D_i \right) \right]$$

$$= N(\Lambda_d) \prod_{i=1}^{d} \left\{ \sum_{\ell=1}^{\infty} \exp\left[\frac{1}{d-1} \ln(\ell) + (\ell-1)\ln(3) + \beta \ell \left(\underbrace{\Phi_i + D_i} \right) \right] \right\}. \tag{82}$$

$< 0 \ \forall\, i = 1 \ldots d$ by assumption $\Rightarrow$ Sums converge for β large enough.

We have thus shown:

$$\frac{\frac{1}{N(\Lambda_d)} \sum_{n=0}^{\infty} \tilde{Z}(\Lambda_d, n)\, n}{\sum_{n=0}^{\infty} \tilde{Z}(\Lambda_d, n)} = \frac{\frac{1}{N(\Lambda_d)} \sum_{S \in \mathcal{U}_d} Z(S)\, N(S)}{\sum_{S \in \mathcal{U}_d} Z(S)}$$

$$\leq \prod_{i=1}^{d} \left\{ \sum_{\ell=1}^{\infty} \exp\left[\beta \overbrace{\left(\Phi_i + D_i \right)}^{<0} + \left(1 - \frac{1}{\ell}\right) \ln(3) + \frac{1}{d-1} \frac{\ln(\ell)}{\ell} \right]^{\ell} \right\} < \frac{1}{2} \tag{83}$$

if β is chosen large enough. In particular, Proposition 4.2.3 is applicable and f reduces to a constant on the interval $[\rho_0, 1-\rho_0]$. This concludes the proof of Theorem 4.2.1.

4.3. The theorem of Mermin and Wagner

In the previous section we proved that classical spin systems on $\mathbb{Z}^d$, $d \geq 2$, whose spins may take two values, have a phase transition at sufficiently low temperature if the activity is varied (this is equivalent to saying that we vary the external magnetic field). We could show that the magnetization of the system (the density in the lattice gas interpretation) has a jump at $z=1$ ($\Leftrightarrow$ magnetic field $h=0$). In particular, there is spontaneous magnetization in the system (density unequal to $1/2$ in the lattice gas interpretation).

In dimensions $d \leq 2$ there is a famous theorem due to Mermin and Wagner, which says that the magnetization equals zero for all values of β if we consider quantum spin systems on $\mathbb{Z}^d$. The same statement is true if we consider classical systems, where the spins take values on the two-sphere. The intuition behind this phenomenon is that quantum or classical continuous spins can vary continuously between up and down. Accordingly, it is possible to have neighboring spins almost aligned without changing the energy much, and without having a coherent orientation of the spins over the whole lattice. The statement we are going to prove concerns the quantum Heisenberg model in $d \leq 2$, which we introduced on p.40 in Section 2. We prefer to present the statement for the quantum Heisenberg model instead of for classical spins because it is more relevant.

We start by introducing the quantum Heisenberg model for more general spins that just $1/2$. Let Λ be a finite subset of $\mathbb{Z}^d$. As discussed in Section 2.2 we have for each $x \in \Lambda$ a copy $\mathcal{H}_x$ of a finite-dimensional Hilbert space $\mathcal{H}$. We also assume that we have in $\mathcal{H}$ an irreducible unitary representation of the group $SU(2)$. This is equivalent to saying that we have a single- or double-valued representation of the three-dimensional rotation group $SO(3)$. By S_1, S_2, S_3 we denote the self-adjoint generators of this representation satisfying the $su(2)$ commutation relations

$$[S_1, S_2] = iS_3, \quad [S_2, S_3] = iS_1, \quad [S_3, S_1] = iS_2. \qquad (84)$$

We also define the operator

$$S_{\pm} = \frac{1}{\sqrt{2}}\left(S_1 \pm i S_2\right). \qquad (85)$$

It can be shown that (see [8] or almost any other book on quantum mechanics)

$$\sum_{i=1}^{3} S_i^2 = S_+S_- + S_-S_+ + S_3^2 = s(s+1)\,\mathbb{1}, \qquad (86)$$

where $2s+1$ is the dimension of $\mathcal{H}$. The number $s \in \{1/2, 1, 3/2, 2, \dots\}$ is interpreted as the length of the spin. We denote the eigenvector of S_3 by $\psi_{s,m}$, i.e., $S_3 \psi_{s,m} = m S_3$. It can be shown, see [8], that $m \in \{-s, -s+1, \dots, s-1, s\}$. The eigenvalue m is interpreted as the projection of the spin vector onto the z-axis. We highlight that, because of the commutation relations in (84), not all spin operators S_1, S_2, S_3 can be diagonalized in the same basis (otherwise this would be a classical spin model).

The operators $S_\pm$ act in the following way on the basis vectors $\Psi_{s,m}$:

$$S_\pm \Psi_{s,m} = \sqrt{s(s+1) - m(m+1)}\, \Psi_{s,m}. \qquad (87)$$

This is a consequence of the commutation relations

$$[S_+, S_-] = S_3, \qquad [S_3, S_\pm] = \pm S_\pm. \qquad (88)$$

Let $S_i(x)$ be the operator corresponding to S_i in $\mathcal{H}_x$. The Hamiltonian we consider in this section reads

$$H(\Lambda) = -\sum_{x,x' \in \Lambda} J(x-x') \sum_{i=1}^{3} S_i(x) S_i(x') - h \sum_{x\in\Lambda} S_3(x). \qquad (89)$$

Here h denotes the strength of the external magnetic field pointing in the e_3-direction. We assume that $J(0)=0$, $J(-x) = J(x)$, and

$$\mathcal{U} = \sum_{x \in \mathbb{Z}^d} x^2 \, |J(x)| < +\infty . \tag{90}$$

We also note that this model describes three-dimensional quantum spins (we consider representations of the rotation group $SO(3)$) that are arranged on a d-dimensional lattice.

Before we come to the main theorem of this section, we discuss an inequality that will be needed in its proof but is also of independent interest.

4.3.1. Lemma (Bogoliubov inequality):

Let $\mathcal{H}$ be a (complex) finite-dimensional Hilbert space, let H be a self-adjoint operator acting on $\mathcal{H}$, let $\beta > 0$, and write

$$\langle X \rangle = \frac{\mathrm{tr}\left[X e^{-\beta H}\right]}{\mathrm{tr}\left[e^{-\beta H}\right]} \tag{S1}$$

of X is an operator on $\mathcal{H}$. Then for all operators A, C, acting on $\mathcal{H}$, we have

$$\frac{1}{2}\beta \langle AA^* + A^*A \rangle \langle [[C,H], C^*] \rangle \geq |\langle [C,A] \rangle|^2. \tag{S2}$$

Here A^* is the adjoint of A and $[C,H] = CH - HC$ is the commutator.

Proof: We start by defining a positive semidefinite inner product in the space of operators acting on $\mathcal{H}$ by

$$(A,B) = \sum_{\varphi,\psi}{}' \langle \psi, A^* \varphi \rangle \langle \varphi, B\psi \rangle \, \frac{W_\varphi - W_\psi}{E_\psi - E_\varphi}, \tag{S3}$$

where the sum extends over the pairs (φ,ψ) with

φ, ψ elements of an orthonormal basis of eigenvectors of H, omitting pairs s.t. $E_\varphi = E_\psi$. Here and in (93) we used the notation

$$E_\varphi = \langle \varphi, H\varphi \rangle, \qquad w_\varphi = \frac{\exp[-\beta E_\varphi]}{\mathrm{tr}[e^{-\beta H}]}. \tag{94}$$

That (93) defines a positive semi-definite inner product can easily be checked.

If $E_\varphi < E_\psi$ then we have

$$\frac{\exp[-\beta E_\varphi] - \exp[-\beta E_\psi]}{\exp[-\beta E_\varphi] + \exp[-\beta E_\psi]} = \tanh\left[\frac{\beta(E_\psi - E_\varphi)}{2}\right]$$

$$< \tfrac{1}{2}\beta\left(E_\psi - E_\varphi\right), \tag{95}$$

and hence,

$$0 \leq \frac{w_\varphi - w_\psi}{E_\psi - E_\varphi} = \frac{1}{\operatorname{tr} e^{-\beta H}} \frac{\exp[-\beta E_\varphi] - \exp[-\beta E_\psi]}{E_\psi - E_\varphi}$$

$$< \frac{1}{\operatorname{tr} e^{-\beta H}} \frac{\beta}{2} \Big(\exp[-\beta E_\varphi] + \exp[-\beta E_\psi] \Big)$$

$$= \frac{\beta}{2} \Big(w_\varphi + w_\psi \Big). \tag{86}$$

Using the above, we estimate

$$(A, A) = \sum_{\varphi\psi}{}' \underbrace{\langle \psi, A^* \varphi \rangle \langle \varphi, A \psi \rangle}_{= |\langle \varphi, A\psi \rangle|^2 \geq 0} \frac{w_\varphi - w_\psi}{E_\psi - E_\varphi}$$

$$< \sum_{\varphi\psi}{}' |\langle \varphi, A\psi \rangle|^2 \frac{\beta}{2} \Big(w_\varphi + w_\psi \Big)$$

$$= \frac{\beta}{2} \sum_{\varphi\psi}{}' \langle \psi, A^* \varphi \rangle w_\varphi \langle \varphi, A\psi \rangle +$$

$$\frac{\beta}{2}\sum_{\varphi\psi}{}' \langle \varphi, A\psi\rangle\, W_\psi \langle \psi, A^*\varphi\rangle$$

$$= \frac{\beta}{2}\ \frac{\mathrm{tr}\left[A^* e^{-\beta H} A\right] + \mathrm{tr}\left[A e^{-\beta H} A^*\right]}{\mathrm{tr}\, e^{-\beta H}}$$

$$= \frac{\beta}{2}\ \langle AA^* + A^*A\rangle. \tag{97}$$

The Cauchy-Schwarz inequality reads

$$|(A,B)|^2 \leq (A,A)(B,B). \tag{98}$$

We choose $B = [C^*, H]$ and compute

$$(A,B) = \sum_{\varphi\psi}{}' \langle \psi, A^*\varphi\rangle \underbrace{\langle \varphi, [C^*,H]\psi\rangle}_{= \langle \varphi, C^*\psi\rangle (E_\psi - E_\varphi)} \frac{W_\varphi - W_\psi}{E_\psi - E_\varphi}$$

$$= \sum_{\varphi\psi}{}' \langle \psi, A^*\varphi\rangle\langle \varphi, C^*\psi\rangle (W_\varphi - W_\psi)$$

$$= \frac{\operatorname{tr}\left[A^* e^{-\beta H} C^*\right] - \operatorname{tr}\left[C^* e^{-\beta H} A^*\right]}{\operatorname{tr} e^{-\beta H}}$$

$$= \left\langle \left[C^*, A^*\right] \right\rangle \tag{99}$$

as well as

$$(B, B) = \left\langle \left[C^*, \left[H, C\right]\right] \right\rangle. \tag{100}$$

Putting everything together, we find

$$\left|(A, B)\right|^2 \underset{(99)}{=} \left|\left\langle \left[C, A\right] \right\rangle\right|^2 \underset{(98)}{\leq} (A, A)(B, B)$$

$$\underset{(97),(100)}{\leq} \frac{\beta}{2} \left\langle AA^* + A^*A \right\rangle \underbrace{\left\langle \left[C^*, \left[H, C\right]\right] \right\rangle}_{\left\langle \left[\left[C, H\right], C^*\right] \right\rangle} \tag{101}$$

This proves the claim. ∎

The main statement we are going to prove in this section is the following theorem.

4.3.2 Theorem (Mermin-Wagner): If we define

$$f(h) = \beta^{-1} \lim_{\alpha \to \infty} \frac{1}{N(\Lambda_\alpha)} \ln \operatorname{tr}_{\mathcal{H}(\Lambda_\alpha)} \exp\left[-\beta H(\Lambda_\alpha)\right], \quad (102)$$

$$\overline{\mathcal{L}}(h) = \frac{d}{dh} f(h), \quad (103)$$

where $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove and $H(\Lambda_\alpha)$ is given by (89) (with the conditions for J below (89) satisfied), then

$$\lim_{h \to 0} \overline{\mathcal{L}}(h) = 0 \quad (104)$$

if $d \leq 2$. That is, we have no spontaneous magnetization.

Proof: That the thermodynamic limit for $g(h)$ exists has been proved in Section 2. In the following we will also assume without proof that

$$\bar{\sigma}(h) = \frac{d}{dh} g(h) = \lim_{\alpha \to \infty} \frac{1}{|\Lambda_\alpha|} \frac{\operatorname{tr} \sum_{x \in \Lambda_\alpha} S_3(x)\, e^{-\beta H(\Lambda_\alpha)}}{\operatorname{tr} e^{-\beta H(\Lambda_\alpha)}} \qquad (105)$$

holds. That means we assume that the limit on the r.h.s. exists if $\Lambda_\alpha \uparrow \mathbb{Z}^d$ in the sense of Van Hove (and that it does not depend on the sequence $\{\Lambda_\alpha\}_{\alpha=1}^{\infty}$) and that it equals $\frac{d}{dh} g(h)$. Both statements require a proof, which can be given with techniques similar to the ones in Section 2. It should be noted that quantities like the one on the r.h.s. of (105) can depend in a delicate way on the choice of the boundary conditions. We understand (105) in the way that either open (as in Section 2) or periodic boundary conditions are applied, See the remark on p. 38 in Section 2.

We first assume that J has finite range, that is, $\operatorname{supp} J = \{x \in \mathbb{Z}^d \mid J(x) \neq 0\}$ is a finite set, and later consider the general case. We also assume that the system is contained in a periodic box $\Lambda(a) = \{x \in \mathbb{Z}^d \mid 0 \leq x_i < a_i \text{ for } i = 1 \dots d\}$, that is, we "glue together" opposite faces of $\Lambda(a)$, see again the remark on p. 38 in Section 2. More formally we could say that we consider the torus $\mathbb{T}(a)$ that we obtain when we take the quotient of $\mathbb{Z}^d$ by the subgroup $\mathbb{Z}^d_a = \{na \mid n \in \mathbb{Z}^d\}$. If $\alpha \in \mathbb{T}(a)$, we would denote by $[\alpha]$ the representative in $\Lambda(a)$. In the following we will not use this notation to not artificially complicate things.

Let us define the set of dual (in the sense of Fourier discrete Fourier analysis) vectors

$$\Delta = \left\{ k \in \mathbb{R}^d \;\middle|\; k_i = \frac{2\pi}{a_i} n_i \text{ with } n_i \in \mathbb{Z} \text{ and } -\frac{a_i}{2} < n_i \leq \frac{a_i}{2} \text{ for } i = 1 \dots d \right\}. \tag{106}$$

We write $k \cdot x = \sum_{i=1}^{d} k_i x_i$ and introduce the Fourier transformed spin operators

$$\hat{S}_i(k) = \sum_{x \in \Lambda(a)} e^{-ik\cdot x} S_i(x), \qquad i = 1,2,3 \tag{107}$$

as well as

$$\hat{J}(k) = \sum_{x \in \Lambda(a)} e^{-ik\cdot x} J(x). \tag{108}$$

The inverse transformation reads

$$S_i(x) = \frac{1}{V(a)} \sum_{k \in \Delta} e^{ik\cdot x} \hat{S}_i(k) \tag{109}$$

and the same for $J(x)$. Here and in the following we use the short-hand notation $V(a) = N(\Lambda(a))$.

The quantity corresponding to $\overline{\sigma}(h)$ in our finite system is

$$\frac{d}{dh} \frac{1}{\beta V(a)} \ln \operatorname{tr} \exp\left[-\beta H(\Lambda)\right] = \left\langle \frac{1}{V(a)} \sum_{x \in \Lambda(a)} S_3(x) \right\rangle, \tag{110}$$

where $\langle \circ \rangle$ is defined by (91) with the Hamiltonian $H(\Lambda)$ of the quantum Heisenberg model in (89). The quantity in (110) is the expectation of the e_3-component of the magnetization per unit volume (in finite volume). It is interesting, for later purposes, to use instead of (109) the more general expression

$$\sigma(h) = \left\langle \frac{1}{V(a)} \sum_{x \in \Lambda(a)} S_3(x)\, e^{ik\cdot x} \right\rangle, \qquad (111)$$

where k is chosen such that $e^{ik\cdot x} \in \{-1,1\}$ for every $x \in \Lambda(a)$, and $H(\Lambda(a))$ is now defined by

$$H(\Lambda(a)) = -\sum_{x,x' \in \Lambda(a)} J(x-x') \sum_{i=1}^{3} S_i(x) S_i(x') - h \sum_{x \in \Lambda(a)} e^{-ik\cdot x} S_3(x). \qquad (112)$$

This slightly more general set-up will later allow us to discuss also the absence of antiferromagnetism in one- and two-dimensional Heisenberg models.

A system is called ferromagnetic if below a certain critical temperature the spins tend to point all in one direction such that we have a non-vanishing magnetisation per unit volume in the thermodynamic limit. That is, locally such a system looks e.g. like ↑↑↑↓↑↑↑↑↓↑↑ in $d=1$ and similarly in $d=2$. Another possible order of the spins, called antiferromagnetic order, that could occur is that neighboring spins tend to point in opposite directions. Locally such a system would look like ↑↓↑↓↓↑↓↑↓↑↓↓ in $d=1$ and similarly in $d=2$. A constant magnetic field pointing in e.g. the e_3-direction is the natural symmetry breaking term to check whether the system has a tendency to display ferromagnetic order. The natural symmetry breaking term to test for anti-ferromagnetic order is a magnetic field with constant absolute value pointing e.g. in the e_3-direction, whose e_3-component changes sign from lattice site to next neighboring lattice site. $\lrcorner$

We will obtain an inequality for $b(k)$ by applying Lemma 4.3.1 with the choice

$$A = \hat{S}_-(-k-K) \quad \left(\text{discrete Fourier transform of } S_\mp(x)\right),$$

$$C = \hat{S}_+(k) \quad \text{and} \quad H = H(\Lambda(a)). \tag{113}$$

We first evaluate the quantities, which appear in Lemma 4.3.1, and start by computing

$$\hat{S}_-(k) = \sum_{x \in \Lambda(a)} e^{-ik\cdot x} S_-(x)$$

$$\Rightarrow \left[\hat{S}_-(k)\right]^* = \sum_{x \in \Lambda(a)} e^{ik\cdot x} \underbrace{\left[S_-(x)\right]^*}_{= S_+(x)} = \hat{S}_+(-k) \tag{114}$$

and

$$\langle AA^* + A^*A \rangle = \langle \hat{S}_-(-k-K)\hat{S}_+(k+K) + \hat{S}_+(k+K)\hat{S}_-(-k-K) \rangle . \tag{115}$$

We also have

$$\langle [C, A] \rangle = \langle [\hat{S}_+(k), \hat{S}_-(-k-K)] \rangle \tag{116}$$

$$= \sum_{x,y \in \Lambda(a)} e^{-ik\cdot x} e^{i(k+K)\cdot y} \underbrace{\langle [S_+(x), S_-(y)] \rangle}_{\overset{(88)}{=} S_3(x)\,\delta_{x,y}}$$

$$= \sum_{x \in \Lambda(a)} e^{iK\cdot x} \langle S_3(x) \rangle = V(a)\, \sigma(h).$$

The next term we compute is

$$[C, H] = -\sum_{x,x',y \in \Lambda(a)} J(x-x')\, e^{-ik\cdot y} \sum_{i=1}^{3} \overset{(**)}{[S_+(y), S_i(x) S_i(x')]} \tag{117}$$

$$-h \sum_{x,y \in \Lambda(a)} e^{-iK\cdot x} e^{-ik\cdot y} [S_+(y), S_3(x)] = (*)$$

$$(**) = \sum_{i=1}^{3} \left\{ S_i(x) \left[S_+(y), S_i(x') \right] + \left[S_+(y), S_i(x) \right] S_i(x') \right\} \tag{118}$$

$$\left[[A, BC] = B[A, C] + [A, B]C \right]$$

- $[S_+, S_1] = \frac{1}{\sqrt{2}} [S_1 + iS_2, S_1] = \frac{i}{\sqrt{2}} \underbrace{[S_2, S_1]}_{\overset{(84)}{=} -iS_3}$

  $= \frac{1}{\sqrt{2}} S_3$

- $[S_+, S_2] = \frac{1}{\sqrt{2}} [S_1, S_2] \overset{(84)}{=} \frac{i}{\sqrt{2}} S_3$ (119)

- $[S_3, S_\pm] \overset{(88)}{=} \pm S_\pm$

$$= \left(\frac{1}{\sqrt{2}} S_1(x) S_3(y) + \frac{i}{\sqrt{2}} S_2(x) S_3(y) - S_3(x) S_+(y) \right) \delta_{x', y}$$

$$+ \left(\frac{1}{\sqrt{2}} S_3(y) S_1(x') + \frac{i}{\sqrt{2}} S_3(y) S_2(x') - S_+(y) S_3(x') \right) \delta_{x, y}$$

$$= \Big(S_+(x) S_3(y) - S_3(x) S_+(y) \Big) \delta_{x',y}$$

$$+ \Big(S_3(y) S_+(x') - S_+(y) S_3(x') \Big) \delta_{x,y} \, .$$

Accordingly, we have

$$(*) = - \sum_{x,x' \in \Lambda(a)} J(x-x') \Big\{ e^{-ik\cdot x'} \Big(S_+(x) S_3(x') - S_3(x) S_+(x') \Big) \qquad (120)$$

$$+ e^{-ik\cdot x} \Big(S_3(x) S_+(x') - S_+(x) S_3(x') \Big) \Big\}$$

$$+ h \sum_{x \in \Lambda(a)} e^{-i(K+k)\cdot x} S_+(x)$$

$$\overset{J(x)=J(-x)}{\underset{(88)}{=}} \; 2 \sum_{x,x' \in \Lambda(a)} J(x-x') \Big(e^{-ik\cdot x'} - e^{-ik\cdot x} \Big) S_3(x) S_+(x')$$

$$+ h \sum_{x \in \Lambda(a)} e^{-i(K+k)\cdot x} S_+(x)$$

$$\overset{\text{recall}}{=} \big[C, H(\Lambda(a)) \big] \, .$$

A similar computation yields (please check!)

$$\left\langle \left[[C, H], C^* \right] \right\rangle = 2 \sum_{x, x' \in \Lambda(a)} \left(1 - e^{-ik\cdot(x-x')} \right) J(x-x') \tag{121}$$

$$\left\langle S_-(x') S_+(x) + S_3(x') S_3(x) \right\rangle$$

$$+ h V(a) \sigma(h).$$

We highlight that (S2) implies that (121) is nonnegative for all $k \in \Delta$. When we add the same quantity with k replaced by $-k$ to the r.h.s., we thus obtain an upper bound for the l.h.s.:

$$\left\langle \left[[C, H], C^* \right] \right\rangle \leq 2 \sum_{x, x' \in \Lambda(a)} \left(1 - \cos\left(k \cdot (x-x') \right) \right) J(x-x') \tag{122}$$

$$\left\langle S_-(x') S_+(x) + \underbrace{S_+(x') S_-(x)} + 2 S_3(x') S_3(x) \right\rangle$$

$$+ 2 h V(a) \sigma(h).$$

Comes about if we go through the computation that led to (120) with k replaced by $-k$.

Next, we claim that

$$\left| \langle S_-(x')S_+(x) + S_+(x')S_-(x) + 2S_3(x')S_3(x) \rangle \right| \tag{123}$$

$$\leq \langle S_-(0)S_+(0) + S_+(0)S_-(0) + 2S_3(0)^2 \rangle$$

$$\leq 2s(s+1)$$

holds. To see this, we first note that

sum over ONB of EV's of H

notation from (84)

$$\left| \langle S_-(x')S_+(x) \rangle \right| = \left| \sum_{\Psi} \langle \Psi, S_-(x')S_+(x)\Psi \rangle \, \omega_\Psi \right| \tag{124}$$

$$\leq \sum_{\Psi} \underbrace{\left| \langle S_+(x')\Psi, S_+(x)\Psi \rangle \right|}_{\overset{CS}{\leq} \|S_+(x')\Psi\| \, \|S_+(x)\Psi\|} \underbrace{\omega_\Psi}_{\geq 0}$$

$$\leq \left(\sum_{\Psi} \langle \Psi, S_-(x')S_+(x')\Psi \rangle \, \omega_\Psi \right)^{1/2}$$

$$\left(\sum_{\Psi} \langle \Psi, S_-(x)S_+(x)\Psi \rangle \, \omega_\Psi \right)^{1/2}$$

$$= \left(\frac{\mathrm{tr}\, S_-(x')S_+(x')\, e^{-\beta H}}{\mathrm{tr}\, e^{-\beta H}} \right)^{1/2} \left(\frac{\mathrm{tr}\, S_-(x)S_+(x)\, e^{-\beta H}}{\mathrm{tr}\, e^{-\beta H}} \right)^{1/2}.$$

Let us define for $v \in \Lambda(a)$ the unitary translation operator

$$\mathcal{U}(v) \underbrace{\Psi_{s,m}(x)}_{\substack{\text{basis vector of} \\ S_3(x)}} = \Psi_{s,m}(x+v). \tag{125}$$

We have

$$\mathcal{U}^*(v)\, S_i(x)\, \mathcal{U}(v) = S_i(x+v), \qquad i = 1,..,3, \tag{126}$$

and the same for $S_\pm(x)$. Using the definition of $H(\Lambda(a))$, we see that $\mathcal{U}^*(v)\, H(\Lambda(a))\, \mathcal{U}(v) = H(\Lambda(a))$, as well as

$$\mathrm{tr}\, S_-(x)S_+(x)\, e^{-\beta H} = \tag{127}$$

$$\mathrm{tr}\, \mathcal{U}^*(-x) S_-(x)\, \mathcal{U}(-x)\, \mathcal{U}^*(-x)\, S_+(x)\, \mathcal{U}(-x)\, \mathcal{U}^*(-x)\, e^{-\beta H}\, \mathcal{U}(-x)$$

$$= \operatorname{tr} S_-(0)S_+(0) \underbrace{\mathcal{U}^*(-x)\, e^{-\beta H}\, \mathcal{U}(-x)}_{e^{-\beta H} \text{ because } \mathcal{U}^*(-x) H\, \mathcal{U}(-x) = H} = \operatorname{tr} S_-(0)S_+(0)\, e^{-\beta H}.$$

In combination with (123), the above implies

$$\left| \left\langle S_-(x')S_+(x) \right\rangle \right| \leq \left\langle S_-(0)S_+(0) \right\rangle. \tag{128}$$

We apply this reasoning to all three terms on the l.h.s. of (123) and find

$$\begin{aligned} &\left| \left\langle S_-(x')S_+(x) + S_+(x')S_-(x) + 2S_3(x')S_3(x) \right\rangle \right| \\ &\leq \left\langle S_-(0)S_+(0) + S_+(0)S_-(0) + S_3(0)^2 \right\rangle. \end{aligned} \tag{129}$$

Eq. (123) now follows from (86).

We apply (123) on the r.h.s. of (122), use $1-\cos(x) \leq \frac{x^2}{2}$ for $x \in \mathbb{R}$ and find

$$\left\langle \left[[C, H], C^* \right] \right\rangle \le \sum_{x, x' \in \Lambda(a)} k^2 (x-x')^2 J(x-x') \, 2s(s+1) \qquad (130)$$

$$+ 2h V(a) \sigma(h)$$

$$\overset{(90)}{\le} 2 V(a) \left(\ell\, s(s+1) k^2 + |h \sigma(h)| \right).$$

We are now prepared to insert our choice of A and C in (113) into the Bogoliubov inequality in (82). When we also use (115), (116) and (130), we find

$$\left\langle \hat{S}_-(-k-K) \hat{S}_+(k+K) + \hat{S}_+(k+K) \hat{S}_-(-k-K) \right\rangle \qquad (131)$$

$$\ge \frac{|V(a) \sigma(h)|^2}{\beta V(a) \left(\ell\, s(s+1) k^2 + |h \sigma(h)| \right)} .$$

We sum both sides of (131) over $k \in \underline{\Lambda}$ and use the bound

$$\sum_{k\in\Delta} \left\langle \hat{S}_-(-k-K)\hat{S}_+(k+K) + \hat{S}_+(k+K)\hat{S}_-(-k-K) \right\rangle \tag{132}$$

$$= \sum_{k\in\Delta} \sum_{x,x'\in\Lambda(a)} \Big\{ e^{i(k+K)\cdot(x-x')} \left\langle S_-(x)S_+(x')\right\rangle + e^{-i(k+K)\cdot(x-x')} \left\langle S_+(x)S_-(x')\right\rangle \Big\}$$

$$= V(a) \sum_{x\in\Lambda(a)} \left\langle S_-(x)S_+(x) + S_+(x)S_-(x) \right\rangle$$

$$\sum_{k\in\Delta} e^{ik\cdot y} = \delta_{y,0}\, V(a)$$

$$\leq V(a) \sum_{x\in\Lambda(a)} \left\langle S_-(x)S_+(x) + S_+(x)S_-(x) + 2S_3^2(x) \right\rangle$$

$$\left\langle S_3(x)^2 \right\rangle \geq 0$$

$$\overset{(86)}{\leq} V(a)^2\, s(s+1),$$

we find

$$S(s+1) \geq \frac{|\mathcal{b}(h)|^2}{\beta\, V(a)} \sum_{k\in\Delta} \frac{1}{\mu\, s(s+1) k^2 + |h\mathcal{b}(h)|} . \tag{133}$$

Let us recall the definition of Δ in (106). We would like to take the limit $a_i \to +\infty$ for $i=1\dots d$ on the r.h.s. of (133) (the l.h.s. does not depend on $a=(a_1 \dots a_d)$). To that end, we interpret the sum over $k\in\Delta$ as a Riemann sum and find

$$\lim_{\substack{a_i\to+\infty \\ \text{for } i=1\dots d}} \frac{1}{V(a)} \sum_{k\in\Delta} \frac{1}{\mu\, s(s+1) k^2 + |h\mathcal{b}(h)|} \tag{134}$$

$$\underset{\substack{\lim_{a_i\to+\infty,\ i=1\dots d} \mathcal{b}(h) = \bar{\mathcal{b}}(h)}}{=} \left(\frac{1}{2\pi}\right)^d \int_{[-\pi,\pi]^d} \frac{1}{\mu\, s(s+1) k^2 + |h\bar{\mathcal{b}}(h)|}\, dk .$$

We obtain a lower bound when we replace the set $[-\pi,\pi]^d$ on the r.h.s. by $B^d_\pi(0) = \{x\in\mathbb{R}^d \mid |x|\leq\pi\}$. Accordingly,

$$\left(\frac{1}{2\pi}\right)^d \int_{[-\pi,\pi]^d} \frac{1}{\mu s(s+1) k^2 + |h\bar{\sigma}(h)|}\, dk \tag{135}$$

$$\geq \frac{\Omega_d}{(2\pi)^d} \int_0^\pi \frac{k^{d-1}}{\mu s(s+1) k^2 + |h\bar{\sigma}(h)|}\, dk,$$

where Ω_d denotes the volume of the unit sphere in d dimensions. In combination (133), (134) and (135) imply

$$s(s+1) \geq \frac{|\bar{\sigma}(h)|^2\, \Omega_d}{\beta (2\pi)^d} \int_0^\pi \frac{k^{d-1}}{\mu s(s+1) k^2 + |h\bar{\sigma}(h)|}\, dk. \tag{136}$$

From now on we specify $d = 1, 2$. In both cases the integral on the r.h.s. of (136) can be computed explicitly and we find

$$s(s+1) \geq \frac{|\bar{\sigma}(h)|^2}{\pi\beta \left[\,|h\bar{\sigma}(h)|\, \mu s(s+1)\right]^{1/2}} \arctan\left(\left[\frac{\pi^2 \mu s(s+1)}{|h\bar{\sigma}(h)|}\right]^{1/2}\right) \tag{137}$$

for $d=1$ and

$$S(S+1) \geq \frac{|\bar{\sigma}(k)|^2}{4\pi\beta\, \mu s(s+1)} \ln\left(1+\frac{\pi^2 \mu s(s+1)}{|k\bar{\sigma}(k)|}\right) \tag{138}$$

if $d=2$. That is, we have

$$|\bar{\sigma}(k)|^3 \leq \pi^2 \mu \left[S(S+1)\right]^3 \beta^2 |k| \frac{1}{\left\{\arctan\left(\left[\frac{\pi^2 \mu s(s+1)}{|k\bar{\sigma}(k)|}\right]^{1/2}\right)\right\}^2} \tag{139}$$

for $d=1$ and

$$|\bar{\sigma}(k)|^2 \leq 4\pi\mu \left[S(S+1)\right]^2 \beta \frac{1}{\ln\left(1+\frac{\pi^2 \mu s(s+1)}{|k\bar{\sigma}(k)|}\right)} \tag{140}$$

for $d=2$.

The function $x \mapsto \arctan(x)$ is bounded for large x, nonnegative for $x>0$ and goes linearly to zero for $x \to 0$. Accordingly, the factor involving the arctan in (139) is either bounded as

$h \to 0$ or (if $h\overline{\sigma}(h) \to +\infty$) it behaves as const. $h\overline{\sigma}(h)$.
In the first case we have $\overline{\sigma}(h) \xrightarrow{h\to 0} 0$. The second case leads to an immediate contradiction because the r.h.s. of (139) would be smaller than the left hand side for h small enough. We conclude that $\overline{\sigma}(h) \xrightarrow{h \to 0} 0$ if $d=1$. A similar argument proves $\overline{\sigma}(h) \xrightarrow{h\to 0} 0$ in $d=2$. This proves Theorem 4.3.2 in the case of finite range interactions.

To prove the Theorem in the general case we approximate the function $J(x)$ by a sequence of functions $\{J_n\}_{n=1}^{\infty}$ with finite range s.t. $J_n \to J$ in $L^1(\mathbb{Z}^d)$ (This assures the convergence of our interaction in the Banach space $\mathcal{B}$ that we introduced in Section 2.2.) and s.t.

$$\sum_{x \in \mathbb{Z}^d} x^2 \, | J_n(x) - J(x) | \xrightarrow{n\to\infty} 0 . \qquad (141)$$

The fact that $g(u)$ depends continuously on the interaction in this case has been proved in Section 2, see Theorem 2.3.9. With similar techniques one can show that this also holds for $\mathcal{Z}(u)$ but we will not give that proof here. To conclude the statement of Theorem 4.3.2 in the general case it now suffices to note that (139) and (140) depend continuously on u and apart from that on no other details of J. Accordingly, we can take the limit $n \to \infty$ on both sides and the inequalities still hold in the limit. If we choose K s.t. $e^{iK \cdot x} = 1$ for all $x \in \mathbb{Z}^d$, this proves Theorem 4.3.2.

$\boxed{\mathbb{Z}}$

Remarks: When we chose K s.t. $e^{iK \cdot x} \in \{-1, 1\}$ in such a way that $e^{iK \cdot x} e^{iK \cdot y} = -1$ for next neighboring lattice sites x and y, we can also exclude anti-ferromagnetic order in the system. More precisely, if we define two sublattices by $\mathcal{L}_\pm = \{ x \in \mathbb{Z}^d \mid e^{iK\cdot x} = \pm 1 \}$ we can exclude that there is an order where spins on $\mathcal{L}_+$ tend to point in the positive e_3-direction while spins on $\mathcal{L}_-$ tend to point in the negative e_3-direction (or vice versa). The relevant order parameter has been defined in (111) and the symmetry breaking term in (112).

Exercise: Check that the bound in (136) does not allow us to conclude that $Z_0(h) \xrightarrow{h\to 0} 0$ if $d \geq 3$.

4.4. Three additional statements without proof

At the end of the lecture I would like to mention three additional statements that I could not prove in the given time, but which are good to know. The first concerns the absence of a first order phase transition in the model that we studied in Section 4.2 if the temperature is chosen large enough.

Theorem 4.4.1.: Let the potential Φ satisfy $\Phi(X)=0$ if $N(X)=1$ or $N(X)>k$ for some $k\in\mathbb{N}$ and let $\|\Phi\|<+\infty$.

We define

$$p_\Phi(z,\beta) = \beta^{-1} \lim_{\alpha\to\infty} \frac{1}{N(\Lambda_\alpha)} \ln \sum_{X\subset\Lambda_\alpha} z^{N(X)} \exp\left[-\beta U(X)\right]. \quad (142)$$

Then, if β is sufficiently small, $p_\Phi(z,\beta)$ is a real analytic function of $z > 0$.

Remark: • This proves the absence of a first order phase transition for sufficiently large temperatures.

• It can be shown that in this case $p_\Phi(z,\beta)$ is also real analytic in β.

The second statement concerns the existence of a first order phase transition in the same lattice gas (or spin) model for sufficiently low temperatures but with fewer assumptions on the interaction.

Theorem 4.4.2: We use the same notation and assumptions as in Theorem 4.2.1 but replace (4.1) by $\Phi^2 \leq 0$ and $\Phi^2(\xi_1) < 0$, $\Phi^2(\xi_2) < 0$ for some linearly independent vectors ξ_1, ξ_2. Then a first order phase transition occurs at $z=1$ for sufficiently small β^{-1}.

The proof of Theorem 4.4.2 uses correlation inequalities called Griffiths inequalities. These inequalities show that if we consider a ferromagnetic interaction between the spins, that is, an interaction that favors spin alignement, then the spins are positively correlated. The precise statement is captured in the following theorem.

Theorem 4.4.3.: Consider a finite set $\Lambda = \{x_1 \dots x_n\}$ and let σ_x take the values ± 1 for each $x \in \Lambda$. For each $X \subseteq \Lambda$, let

$$\sigma^X = \prod_{x \in X} \sigma_x \tag{143}$$

and let $J_X \geq 0$ or $J_X = +\infty$, $J_\emptyset = 0$. We define

$$\mathcal{U}'(\sigma_{x_1}, \dots, \sigma_{x_n}) = -\sum_{X \subseteq \Lambda} J_X \sigma^X, \tag{144}$$

$$Z = \sum_{\sigma \in \{-1,1\}^n} \exp\left[-\mathcal{U}'(\sigma_{x_1}, \dots, \sigma_{x_n})\right], \tag{145}$$

$$\langle \sigma^X \rangle = \frac{\sum_{\sigma \in \{-1,1\}^n} \sigma^X \exp\left[-\mathcal{U}'(\sigma_{x_1}, \dots, \sigma_{x_n})\right]}{Z}. \tag{146}$$

With this notation we have

$$\frac{\partial \langle \sigma^X \rangle}{\partial J_Y} = \langle \sigma^X \sigma^Y \rangle - \langle \sigma^X \rangle \langle \sigma^Y \rangle \tag{147}$$

for all $X, Y \subseteq \Lambda$ and the following inequalities hold

$$\langle \sigma^X \rangle \geq 0, \tag{148}$$

$$\langle \sigma^X \sigma^Y \rangle - \langle \sigma^X \rangle \langle \sigma^Y \rangle \geq 0. \tag{149}$$

This ends our discussion of classical spin systems as well as our lecture on mathematical statistical mechanics.