# Lectures on Mathematical Methods in Statistical Mechanics

(updated 2026.05)

Si Li

**Abstract**

Lecture note for the graduate course “Mathematical Foundations of Statistical Mechanics” in Spring 2026 at Qiuzhen College, Tsinghua University.

## Contents

# 1 Contact Geometry

## Contents

Classical mechanics is rooted in symplectic geometry. In a parallel fashion, contact geometry provides the natural geometric framework for thermodynamics. For example, thermodynamic phase space is a contact manifold, and equilibrium states are Legendrian submanifolds whose generating functions give rise to thermodynamic potentials. In this section we introduce the basic language of contact geometry and explain its connection with thermodynamics.

## 1.1 Contact manifolds

Let $M$ be an orientable manifold of odd dimension $2n+1$. Let $\xi \subset TM$ be a subbundle of rank $= 2n$. At each point $x$, $\xi_x \subset T_xM$ is a hyperplane in the tangent vectors at $x$, hence can be written as:

$$\xi_x = \mathrm{Ker}(\alpha_x)$$

where $\alpha_x$ is a linear 1-form at $x$. Intrinsically, $\alpha_x$ is a generator of the real line

$$(T_xM/\xi_x)^\vee \hookrightarrow (T_xM)^\vee.$$

We will always assume $\xi$ is orientable, which is equivalent to the triviality of the line bundle $TM/\xi$. In this case we can always find a global 1-form $\alpha$ on $M$ such that $\xi = \ker \alpha$, i.e.:

$$\xi_x = \mathrm{Ker}\ \alpha_x, \qquad \forall x \in M.$$

Such $\alpha$ is defined up to a global rescaling: for any nowhere vanishing function $\lambda : M \to \mathbb{R} \setminus \{0\}$, both $\alpha$ and $\lambda\alpha$ lead to the same $\xi$.

**Definition 1.1.** The pair $(M, \xi)$ is called a **contact manifold** if its defining 1-form $\alpha$ satisfies:

$$\alpha \wedge (d\alpha)^n \neq 0 \qquad \text{everywhere on } M.$$

Such $\alpha$ is called a **contact form**.

The contact condition says $\alpha \wedge (d\alpha)^n$ defines a volume form on $M$. Note that this condition is a property of $\xi$ and does not depend on the choice of $\alpha$. In fact, for another choice $\lambda\alpha$ where $\lambda : M \to \mathbb{R} \setminus \{0\}$, we have:

$$(\lambda\alpha) \wedge (d(\lambda\alpha))^n = \lambda^{n+1}\alpha \wedge (d\alpha)^n.$$

**Example 1.2.** *On* $\mathbb{R}^{2n+1}$ *with coordinates* $\{q^i, p_i, z\}$*,*

$$\alpha = dz + \sum_{i=1}^{n} p_i dq^i$$

*defines a contact 1-form. We have*

$$d\alpha = \sum_{i=1}^{n} dp_i \wedge dq^i$$

*so* $\alpha \wedge (d\alpha)^n$ *is proportional to the standard volume form on* $\mathbb{R}^{2n+1}$*.*

**Example 1.3.** *Consider* $S^{2n+1}$ *as the unit sphere in* $\mathbb{C}^{n+1}$*,*

$$S^{2n+1} = \left\{ (z_0, z_1, \ldots, z_n) \in \mathbb{C}^{n+1} \mid \sum_{j=0}^{n} |z_j|^2 = 1 \right\}$$

*Consider the following Liouville form* $\lambda = \frac{i}{2}\sum_{j=0}^{n}(z_j d\bar{z}_j - \bar{z}_j dz_j)$ *on* $\mathbb{C}^{n+1}$*. In terms of real coordinates* $z_j = x_j + iy_j$*,*

$$\lambda = \sum_{j=0}^{n}(x_j dy_j - y_j dx_j).$$

*Then* $\alpha = \lambda|_{S^{2n+1}}$ *defines a contact 1-form. To see this, consider the Euler vector field*

$$E = \sum_{j=0}^{n} \left( x_j \frac{\partial}{\partial x_j} + y_j \frac{\partial}{\partial y_j} \right)$$

*along the radial direction normal to* $S^{2n+1}$*. Since* $\iota_E d\lambda = 2\lambda$*, we find*

$$\alpha \wedge (d\alpha)^n = \frac{1}{2n+2}\iota_E (d\lambda)^{1n+1}|_{S^{2n+1}}.$$

*Since* $(d\lambda)^{n+1}$ *is a volume form on* $\mathbb{C}^{n+1}$ *and* $E$ *is normal to* $S^{2n+1}$*, we know* $\alpha \wedge (d\alpha)^n$ *is a volume form on* $S^{2n+1}$*.*

**Theorem 1.4** (Darboux Theorem)**.** *Let* $\alpha$ *be a contact form on* $M$ *and* $p \in M$*. Then there are local coordinates* $\{x^1, \ldots, x^n, y_1, \ldots, y_n, z\}$ *near* $p$ *such that:*

$$\alpha = dz + \sum_{j=1}^{n} x_j dy_j$$

This can be proved similarly to the symplectic case via the Moser's trick. Such local coordinates will be called a Darboux coordinate of $(M, \xi)$. By Darboux Theorem, Example 1.2 can be viewed as the standard local model of contact manifolds. In Darboux coordinates,

$$\xi = \text{Span}\left\{\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial y_1} - x_1\frac{\partial}{\partial z}, \dots, \frac{\partial}{\partial y_n} - x_n\frac{\partial}{\partial z}\right\}.$$

## 1.2 Legendrian submanifolds

Let $(M, \xi)$ be a contact manifold and $\alpha$ be a contact form. At any point $x \in M$, $d\alpha|_{\xi_x}$ defines a skew-symmetric form on the tangent hyperplane $\xi_x$. By the contact condition,

$$(d\alpha)^n|_{\xi_x} \neq 0.$$

Thus $d\alpha|_{\xi_x}$ is non-degenerate on $\xi_x$. In other words, $(\xi_x, d\alpha|_{\xi_x})$ is a symplectic vector space. This symplectic structure is well-defined up to a rescaling constant. In fact, for a different contact form $\lambda\alpha$:

$$d(\lambda\alpha)|_{\xi_x} = \lambda d\alpha|_{\xi_x} = \lambda(x) d\alpha|_{\xi_x}$$

where we have used the defining relation $\alpha|_{\xi_x} = 0$.

**Definition 1.5.** A submanifold $L$ of $(M, \xi)$ is called an **isotropic submanifold** if $T_xL \subset \xi_x$ for all $x \in L$. In terms of the contact form $\alpha$, this is equivalent to:

$$\alpha|_L = 0.$$

Let $L$ be an isotropic submanifold. Then $d\alpha|_L = d(\alpha|_L) = 0$, hence $T_xL$ defines an isotropic subspace of $(\xi_x, d\alpha|_{\xi_x})$. By standard linear algebra, this implies:

$$\dim L \leq n = \frac{\dim M - 1}{2}.$$

**Definition 1.6.** An isotropic submanifold $L$ of a $(2n+1)$-dim contact manifold $(M, \xi)$ is called a **Legendrian submanifold** if $\dim L = n$ is of maximal dimension.

**Example 1.7.** *Consider the standard local model* $\mathbb{R}^{2n+1}$ *with* $\alpha = dz + \sum\limits_{i=1}^{n} x_i dy_i$. *Let* $f = f(y_i)$ *be a function of* $y_i$. *Then the equations*

$$x_i = -\frac{\partial f}{\partial y_i}, \quad z = f$$

*define a Legendrian submanifold.*

*More generally, given any partition of the set of indices* $\{1, 2, \ldots, n\}$ *into two disjoint subsets* $I$ *and* $J$*, and for a function* $\varphi = \varphi(x_I, y_J)$*, the* $(n+1)$ *equations*

$$x_j = -\frac{\partial\varphi}{\partial y_j}, \quad y_i = \frac{\partial\varphi}{\partial x_i}, \quad z = \varphi - \sum_{i\in I} x_i \frac{\partial\varphi}{\partial x_i}, \qquad i \in I, j \in J$$

*define a Legendrian submanifold* $L$*. Such* $\varphi$ *is called a* ***generating function*** *of* $L$*.*

*Conversely, let* $L \subset M$ *be a Legendrian submanifold and assume* $\{x_I, y_J\} = \{x_i, y_j\}_{i\in I, j\in J}$ *defines a coordinate on* $L$ *for some partition* $\{I, J\}$*. Thus* $L$ *is defined by relations*

$$x_j = x_j(x_I, y_J), \quad y_i = y_i(x_I, y_J), \quad z = z(x_I, y_J), \qquad i \in I, j \in J.$$

*Define the function*

$$\varphi(x_I, y_J) = z(x_I, y_J) + \sum_{i\in I} x_i y_i(x_I, y_J).$$

*By the Legendrian condition*

$$\begin{aligned} & dz + \sum_{i\in I} x_i dy_i + \sum_{j\in J} x_j dy_j|_L = 0 \\ \Longrightarrow \quad & d\varphi = \sum_{i\in I} y_i dx_i - \sum_{j\in J} x_j dy_j \text{ on } L \\ \Longrightarrow \quad & x_j = -\frac{\partial\varphi}{\partial y_j}, \quad y_i = \frac{\partial\varphi}{\partial x_i} \text{ on } L. \end{aligned}$$

*Thus* $\varphi$ *is a generating function of* $L$*.*

**Example 1.8.** *Let* $L$ *be a Legendrian Submanifold of* $(M, \xi)$*. Then locally* $L$ *is represented by a generating function in Darboux coordinates.*

### Legendre transform of generating functions

Consider the standard contact manifold $\mathbb{R}^{2n+1}$, $\alpha = dz + \sum_{i=1}^{n} x_i dy_i$. Let $L \subset \mathbb{R}^{2n+1}$ be a Legendrian submanifold. As shown in Example 1.7, $L$ can be described by a generating function. However, such generating function depends on a choice of coordinates on $L$. Legendre transform relates generating functions in different coordinates.

For example, assume both $\{y_1, \ldots, y_n\}$ and $\{x_I, y_J\}_{i\in I, j\in J}$ define coordinates on $L$. Let $f(y)$ be the generating function associated to $\{y_1, \ldots, y_n\}$ and $\tilde{f}(x_I, y_J)$ be the generating function associated to $\{x_I, y_J\}_{i\in I, j\in J}$. By Example 1.7

$$f = z|_L \text{ viewed as a function in } \{y_1, \ldots, y_n\}$$

and

$$\tilde{f} = z + \sum_{i\in I} x_i y_i|_L \text{ viewed as a function in } \{x_i, y_j\}_{i\in I, j\in J}.$$

Under the coordinate transformation

$$\{y_1, \ldots, y_n\} \to \{x_i, y_j\}_{i \in I},$$

the generating function transforms as

$$f \to \tilde{f} = f + \sum_{i \in I} x_i y_i = f - \sum_{i \in I} y_i \frac{\partial f}{\partial y_i}.$$

This is the standard form of Legendre transform. Thus the Legendre transform here describes the same Legendrian submanifold in different parametrized coordinates.

## 1.3 Thermodynamic phase space

The simplest thermodynamical phase space of a system is a contact manifold with coordinates $\{U, T, S, P, V\}$ and contact form

$$\alpha = dU - TdS + PdV.$$

Here

- $u =$ Internal Energy
- $T =$ Temperature
- $S =$ Entropy
- $P =$ Pressure
- $V =$ Volume

Equilibrium thermodynamics corresponds to a Legendrian submanifold $\mathcal{E}$ such that

$$\alpha|_{\mathcal{E}} = 0, \quad \text{i.e.} \quad dU = TdS - PdV.$$

This is the First Law of Thermodynamics. As we have learned, $\mathcal{E}$ can be described by generating functions, which are called **thermodynamical potentials**. Different parametrizations of $\mathcal{E}$ lead to different thermodynamical potentials.

1. Choose coordinates $\{S, V\}$.
   The generating function is the internal energy $U = U(S, V)$. Then the Legendrian condition $dU = TdS - PdV$ leads to

   $$T = \frac{\partial U}{\partial S}, \quad P = -\frac{\partial U}{\partial V}.$$

2. Choose Coordinates $\{T, V\}$.

   The generating function $F = F(T, V)$ is obtained by Legendre transform

   $$F = U - TS.$$

   The Legendrian condition $\alpha|_{\mathcal{E}} = 0$ becomes:

   $$dF = -SdT - PdV$$

   which leads to

   $$S = -\frac{\partial F}{\partial T}, \quad P = -\frac{\partial F}{\partial V}.$$

   The thermodynamic potential $F(T, V)$ is called the **Helmholtz free energy**.

3. Choose Coordinates $\{S, P\}$.

   The generating function $H = H(S, P)$ is obtained by Legendre transform:

   $$H = U + PV$$

   The Legendrian condition $\alpha|_{\mathcal{E}} = 0$ becomes

   $$dH = TdS + VdP$$

   which leads to

   $$T = \frac{\partial H}{\partial S}, \quad V = \frac{\partial H}{\partial P}.$$

   The thermodynamical potential $H(S, P)$ is called the **enthalpy**.

4. Choose coordinates $\{T, P\}$.

   The generating function $G = G(T, P)$ is obtained by Legendre transform

   $$G = U - TS + PV$$

   The Legendrian condition $\alpha|_{\mathcal{E}} = 0$ becomes

   $$dG = -SdT + VdP$$

   which leads to

   $$S = -\frac{\partial G}{\partial T}, \quad V = \frac{\partial G}{\partial P}.$$

   The thermodynamical potential $G(T, P)$ is called the **Gibbs free energy**.

Thus the four thermodynamical potentials $U(S, V)$, $F(T, V)$, $H(S, P)$, $G(T, P)$ and the differential relations

$$\begin{aligned} dU &= TdS - PdV \\ dF &= -SdT - PdV \\ dH &= TdS + VdP \\ dG &= -SdT + VdP \end{aligned}$$

express the same Legendrian submanifold (equilibrium) in the thermodynamic phase space via different parametrizations.

## 1.4 Thermodynamic metrics

Metric structures in thermodynamics capture various properties of physical quantities (such as fluctuations) and exhibit interesting behaviors at phase transitions. We discuss two thermodynamic metrics here.

1. **Weinhold metric** in the energy representation $U = U(S, V)$

   $$g^{(W)} = \nabla^2 U = \frac{\partial^2 U}{\partial S \partial S} dS dS + 2\frac{\partial^2 U}{\partial S \partial V} dS dV + \frac{\partial^2 U}{\partial V \partial V} dV dV$$

   or in components $g^{(W)}_{ij} = \frac{\partial^2 U}{\partial X^i \partial X^j}$, $X^i = (S, V)$.

2. **Ruppeiner metric** in the entropy representation $S = S(U, V)$

   $$g^{(R)} = -\nabla^2 S = -\left(\frac{\partial^2 S}{\partial U \partial U} dU dU + 2\frac{\partial^2 S}{\partial U \partial V} dU dV + \frac{\partial^2 S}{\partial V \partial V} dV dV\right)$$

   or in components $g^{(W)}_{ij} = -\frac{\partial^2 S}{\partial Y^i \partial Y^j}$, $Y^i = (U, V)$.

These two thermodynamic metrics are related by

$$g^{(R)} = \frac{1}{T} g^{(W)}$$

under the coordinate transformation.

Contact Riemannian geometry gives a beautiful unification of such thermodymaic metrics. We will not dive into a detailed study of Contact Riemannian geometry, but instead explain the relevant construction in this example of the thermodynmaical phase space parmatrized by $\{U, T, S, P, V\}$ with the contact 1-form $\alpha = dU - TdS + PdV$.

Consider the non-degenerate symmetric form

$$\Sigma = dTdS - dPdV + \alpha \otimes \alpha.$$

Let $L \subset M$ be a Legendre submanifold and we consider the symmetric tensor $\Sigma|_L$.

1. In the $\{S, V\}$ coordinates and $U = U(S, V)$, we have

   $$T = \frac{\partial U}{\partial S}, \quad P = \frac{\partial U}{\partial V}.$$

   Thus

   $$\Sigma|_L = d(\frac{\partial U}{\partial S}) dS + d(\frac{\partial U}{\partial V}) dV = \frac{\partial^2 U}{\partial S \partial S} dS^2 + 2\frac{\partial^2 U}{\partial S \partial V} dS dV + \frac{\partial^2 U}{\partial V \partial v} dV^2 = g^{(W)}$$

   which is the Weinhold metric.

2. In the $(U, V)$ coordinates and $S = S(U, V)$, the relation $dS = \frac{1}{T}dU + \frac{P}{T}dV$ gives
$$\frac{1}{T} = \frac{\partial S}{\partial U}, \quad \frac{P}{T} = \frac{\partial S}{\partial V}.$$
Thus
$$\begin{aligned}\Sigma|_L &= dTdS - d(\frac{P}{T}\cdot T)dV\\ &= dT(dS - \frac{P}{T}dV) - Td(\frac{P}{T})dV\\ &= -Td(\frac{1}{T})dU - Td(\frac{P}{T})dV\\ &= -Td(\frac{\partial S}{\partial U})dU - Td(\frac{\partial S}{\partial V})dV\\ &= T\left(-\frac{\partial^2 S}{\partial U\partial U}dU^2 - 2\frac{\partial^2 S}{\partial U\partial V}dUdV - \frac{\partial^2 S}{\partial V\partial V}dV^2\right)\\ &= Tg^{(R)}.\end{aligned}$$
which is Ruppeiner metric rescaled by the factor $T$.

This proves the relation $Tg^{(R)} = g^{(W)}$ and they are different parametrizations of the tensor $\Sigma|_L$.

## 1.5 Contact transformations

**Definition 1.9.** Let $(M, \xi)$ be a contact manifold with contact 1-form $\alpha$. A diffeomorphism $\varphi : M \to M$ is said to be a **contact diffeomorphism** if $\varphi$ preserves $\xi$, i.e.
$$\varphi_*\xi = \xi.$$
In terms of the contact form $\alpha$, this is equivalent to
$$\varphi^*(\alpha) = \lambda\alpha$$
where $\lambda$ is a nowhere vanishing function on $M$. If $\lambda = 1$, i.e. $\varphi^*\alpha = \alpha$, then $\varphi$ is called a **strict contact diffeomorphism**.

**Example 1.10.** *Consider* $(\mathbb{R}^3, \alpha = dz + xdy)$.

- *Translation*
$$\varphi_c : (x, y, z) \to (x, y, z + c), \quad c \in \mathbb{R}$$
*is a strict contact diffeomorphism.*
- *Rescaling*
$$\varphi_\lambda : (x, y, z) \to (x, \lambda y, \lambda z), \quad \lambda \in \mathbb{R} \setminus \{0\}$$
*is a contact diffeomorphism but not strict if* $\lambda \neq 1$.

- *Consider*

$$\varphi : (x, y, z) \to (y, -x, z + xy).$$

*We have* $\varphi^*(\alpha) = d(z+xy)+yd(-x) = \alpha$, *thus* $\varphi$ *is a strict contact diffeomorphism. Such* $\phi$ *is also called a Legendre transform.*

Infinitesimally, assume a vector field $X$ on $M$ generates a one-parameter family of contact diffeomorphisms. Then the infinitesimal form of preserving contact structure is

$$\mathcal{L}_X \alpha = \tau \alpha$$

where $\mathcal{L}_X$ is the Lie derivative and $\tau$ is a function on $M$. Equivalently,

$$\mathcal{L}_X \alpha \wedge \alpha = 0.$$

Such vector field $X$ is called a **contact vector field**.

**Example 1.11.** *Consider* $(\mathbb{R}^3, \alpha = dz + xdy)$.

- $X_1 = \frac{\partial}{\partial z}$ *is a contact vector field generating the translation in* $z$.
- $X_2 = y\frac{\partial}{\partial y} + z\frac{\partial}{\partial z}$ *is a contact vector field generating the rescaling transformation in* $(y, z)$.
- $X_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} + \frac{1}{2}(x^2 - y^2)\frac{\partial}{\partial z}$ *is a contact vector field.*

$$\mathcal{L}_{X_3}\alpha = d\iota_{X_3}\alpha + \iota_{X_3}d\alpha = d(-x^2+\frac{1}{2}x^2-\frac{1}{2}y^2)+\iota_{X_3}dx\wedge dy = d(-\frac{1}{2}x^2-\frac{1}{2}y^2)+d(\frac{1}{2}x^2+\frac{1}{2}y^2) = 0.$$

**Reeb vector field**

**Definition 1.12.** There exists a unique vector field $R_\alpha$ satisfying

1) $\iota_{R_\alpha} d\alpha = 0$
2) $\alpha(R_\alpha) = 1$

In fact, since $d\alpha|_\xi$ is symplectic, condition 1) determines $R_\alpha$ up to a constant, and condition 2) fixes the normalization constant. $R_\alpha$ is called the **Reeb vector field**.

**Example 1.13.** *Consider* $(\mathbb{R}^{2n+1}, \alpha = dz + \sum_{i=1}^n x_i dy_i)$. *The Reeb vector field is*

$$R_\alpha = \frac{\partial}{\partial z}.$$

**Example 1.14.** *Consider* $S^{2n+1} \subset \mathbb{C}^{n+1}$ *with contact form* $\alpha = \lambda|_{S^{2n+1}}$ *where* $\lambda = \sum\limits_{j=0}^{n}(x_j dy_j - y_j dx_j)$. *The Reeb vector field is*

$$R_\alpha = \sum_{j=0}^{n}(x_j\frac{\partial}{\partial y_j} - y_j\frac{\partial}{\partial x_j})\bigg|_{S^{2n+1}}.$$

**Proposition 1.15.** *Let $\lambda$ be a function on $M$. Then*

$$\mathcal{L}_{\lambda R_\alpha}\alpha = d\lambda.$$

*In particular, the Reeb vector field $R_\alpha$ is a contact vector field.*

*Proof:* $\mathcal{L}_{\lambda R_\alpha}\alpha = d\iota_{\lambda R_\alpha}\alpha + \iota_{\lambda R_\alpha}d\alpha = d\lambda.$ ☐

### Contact vector field v.s. Hamiltonian function

Let $X$ be a contact vector field on $M$. It defines a function on $M$ by

$$h_X = \iota_X\alpha.$$

The contact condition $\mathcal{L}_X\alpha = \tau\alpha$ becomes $dh_X + \iota_X d\alpha = \tau\alpha$. Applying $\iota_{R_\alpha}$ to both sides, we find $R_\alpha(h_X) = \tau$. Thus

$$\iota_X d\alpha = R_\alpha(h_X)\alpha - dh_X.$$

Now we have the combined equations

$$\begin{cases} \iota_X d\alpha = R_\alpha(h_X)\alpha - dh_X \\ \iota_X\alpha = h_X \end{cases}$$

The first equation determines $X$ modulo a vector along $R_\alpha$, and the second equation determines this ambiguity. Thus $X$ is completely determined by $h_X$.

**Definition 1.16.** Given a function $h$ on $M$, the equations

$$\begin{cases} \iota_X d\alpha = R_\alpha(h)\alpha - dh \\ \iota_X\alpha = h \end{cases}$$

determines a unique contact vector field $X$, called *the contact vector field associated to the Hamiltonian function $h$.*

*Remark* 1.17. In symplectic geometry, a Hamiltonian vector field only determines Hamiltonian functions up to a constant shift. In contact geometry, the above construction shows that there is a one-to-one correspondence between contact vector fields and Hamiltonian functions

$$X \longleftrightarrow h_X.$$

**Example 1.18.** *Consider $(\mathbb{R}^3, \alpha = dz + xdy)$. Let $h = \frac{1}{2}x^2 + \frac{1}{2}y^2$ be the Hamiltonian function. The corresponding contact vector field $X$ satisfies*

$$\begin{cases} \iota_X dx \wedge dy = -dh \\ \iota_X\alpha = h \end{cases}$$

*which is solved by $X = x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} + \frac{1}{2}(y^2 - x^2)\frac{\partial}{\partial z}$.*

## 1.6 Contact Hamiltonian equations

Let $(M, \xi)$ be a contact manifold with contact form $\alpha$ and Reeb vector field $R_\alpha$. Let $f$ be a Hamiltonian function on $M$, whose contact vector field is denoted by $X_f$. The flow equations along the vector field $X_f$ are called **contact Hamiltonian equations**.

To compare with the standard Hamiltonian equations in classical mechanics, let's consider

$$\mathbb{R}^{2n+1}, \quad \alpha = dz + \sum_{i=1}^n p_i dq^i, \quad R_\alpha = \frac{\partial}{\partial z}.$$

Let $H = H(p, q, z)$ be the Hamiltonian function. The associated contact vector field $X_H$ solves

$$\begin{cases} \iota_{X_H}(\sum\limits_{i=1}^n dp_i \wedge dq^i) = (\partial_z H)\alpha - dH \\ \iota_{X_H}\alpha = H \end{cases}$$

Let us write $X_H = A^i \frac{\partial}{\partial q^i} + B_i \frac{\partial}{\partial p_i} + C\frac{\partial}{\partial z}$. Then the above equations become

$$\begin{cases} -A^i dp_i + B_i dq^i = (\partial_z H)(dz + p_i dq^i) - dH \\ C + A^i p_i = H \end{cases}$$

which is solved by

$$A^i = \frac{\partial H}{\partial p_i}, \quad B_i = p_i \partial_z H - \frac{\partial H}{\partial q_i}, \quad C = H - p_i \frac{\partial H}{\partial p_i}.$$

Thus the contact vector field associated to $H$ is

$$X_H = \frac{\partial H}{\partial p_i}\frac{\partial}{\partial q_i} + (p_i \partial_z H - \frac{\partial H}{\partial q^i})\frac{\partial}{\partial p_i} + (H - p_i \frac{\partial H}{\partial p_i})\frac{\partial}{\partial z}.$$

The Contact Hamiltonian equations are read by

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial q_i} + p_i \frac{\partial H}{\partial z} \\ \dot{z} = H - p_i \frac{\partial H}{\partial p_i} \end{cases}$$

*Remark* 1.19. When $H = H(q, p)$ does not depend on $z$, the first two equations become

$$\begin{cases} \dot{q}_i = \frac{\partial H}{\partial p_i} \\ \dot{p}_i = -\frac{\partial H}{\partial q^i} \end{cases}$$

which is the standard Hamilton's equations. The last equation is solved by

$$-z = \int^t (p_i \dot{q}_i - H)dt$$

which is the Hamilton's principal function.

We now consider the time evolution of the Hamiltonian function $H$ itself. From the contact Hamiltonian equations

$$\begin{aligned}\frac{dH}{dt} =&\dot{q}_i\frac{\partial H}{\partial q_i}+\dot{p}_i\frac{\partial H}{\partial p_i}+\dot{z}\frac{\partial H}{\partial z}\\ =&\frac{\partial H}{\partial p_i}\frac{\partial H}{\partial q_i}+(-\frac{\partial H}{\partial q_i}+p_i\frac{\partial H}{\partial z})\frac{\partial H}{\partial p_i}+(H-p_i\frac{\partial H}{\partial p_i})\frac{\partial H}{\partial z}\\ =&H\frac{\partial H}{\partial z}=R_\alpha(H)H.\end{aligned}$$

The Hamiltonian is not conserved if $R_\alpha(H)\neq 0$, in contrast to conservative system. From

$$\begin{cases}\mathcal{L}_{X_H}\alpha=R_\alpha(H)\alpha\\ \iota_{X_H}\alpha=H\end{cases}$$

we find

$$\mathcal{L}_{X_H}(\alpha\wedge(d\alpha)^n)=(n+1)R_\alpha(H)\alpha\wedge(d\alpha)^n.$$

In particular, the contact Hamiltonian flow does not preserve the volume form $\alpha\wedge(d\alpha)^n$ (the analogue of Liouville's Theorem does not hold here). When $R_\alpha(H)<0$, this is a dissipative system.

**Example 1.20.** *Consider the contact Hamiltonian*

$$H=\frac{p^2}{2m}+V(q)-\frac{\gamma}{m}z$$

*where $V(q)$ is the mechanical potential and $\gamma$ is a constant. The contact Hamiltonian equations read*

$$\begin{aligned}\dot{q}&=\frac{p}{m}\\ \dot{p}&=-\frac{\partial V}{\partial q}-\frac{\gamma}{m}p\\ \dot{z}&=-\frac{p^2}{2m}+V(q)-\frac{\gamma}{m}z\end{aligned}$$

*The first two equations give rise to*

$$m\ddot{q}+\gamma\dot{q}+\frac{\partial V}{\partial q}=0.$$

*For $\gamma>0$, this is the damped Newtonian equation with a friction force that depends linearly on the velocity.*

## 1.7 Jacobi bracket

In symplectic geometry, the algebraic structure on functions capturing dynamics is the Poisson bracket. The analogue of Poisson bracket in contact geometry is **Jacobi bracket**.

Let $X, Y$ be two contact vector fields. It is clear that their Lie bracket $[X, Y]$ is also a contact vector field. Let $CVect(M)$ denote the space of smooth contact vector fields on $M$. Then $(CVect(M), [\cdot,\cdot])$ forms a Lie algebra. Under the isomorphism

$$\alpha\colon CVect(M) \to C^\infty(M)$$

we have induced a Lie bracket on $C^\infty(M)$ by

$$\{f, g\} = \iota_{[X_f,X_g]}.\alpha$$

This bracket $\{f, g\}$ is called the Jacobi bracket. Using Cartan's formula

$$\begin{aligned}\iota_{[X_f,X_g]}\alpha =&[\mathcal{L}_{X_f}, \iota_{X_g}]\alpha\\ =&\mathcal{L}_{X_f} g - \iota_{X_g}\mathcal{L}_{X_f}\alpha\\ =&\iota_{X_f} dg - \iota_{X_g}(R_\alpha(f)\alpha)\\ =&\iota_{X_f}(R_\alpha(g)\alpha - \iota_{X_g} d\alpha) - \iota_{X_g}(R_\alpha(f)\alpha)\\ =&-\iota_{X_f}\iota_{X_g} d\alpha + fR_\alpha(g) - gR_\alpha(f).\end{aligned}$$

Thus we find

$$\{f, g\} = -\iota_{X_f}\iota_{X_g} d\alpha + fR_\alpha(g) - gR_\alpha(f).$$

In local Darboux coordinates with $\alpha = dz + \sum_{i=1}^{n} x_i dy_i$, we have

$$\begin{cases} X_f = \frac{\partial f}{\partial x_i}\frac{\partial}{\partial y_i} + \left(x_i\frac{\partial f}{\partial z} - \frac{\partial f}{\partial y_i}\right)\frac{\partial}{\partial x_i} + \left(f - x_i\frac{\partial f}{\partial x_i}\right)\frac{\partial}{\partial z}\\ X_g = \frac{\partial g}{\partial x_i}\frac{\partial}{\partial y_i} + \left(x_i\frac{\partial g}{\partial z} - \frac{\partial g}{\partial y_i}\right)\frac{\partial}{\partial x_i} + \left(g - x_i\frac{\partial g}{\partial x_i}\right)\frac{\partial}{\partial z}\end{cases}$$

and

$$\{f, g\} = \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial y_i} - \frac{\partial f}{\partial y_i}\frac{\partial g}{\partial x_i} + \partial_z f x_i\frac{\partial g}{\partial x_i} - \partial_z g x_i\frac{\partial f}{\partial x_i} + f\partial_z g - g\partial_z f.$$

**Proposition 1.21.** *The contact Hamiltonian equations associated to $H$ is expressed in Jacobi bracket by*

$$\dot{f} = \{H, f\} + fR_\alpha H.$$

*In particular, $\dot{H} = HR_\alpha H$.*

*Proof:* The contact Hamiltonian equations read

$$\begin{aligned}\dot{f} =&X_H(f) = \iota_{X_H} df\\ =&\iota_{X_H}(R_\alpha(f)\alpha - \iota_{X_f} d\alpha)\\ =&-\iota_{X_H}\iota_{X_f} d\alpha + HR_\alpha(f)\\ =&\{H, f\} + fR_\alpha(H).\end{aligned}$$

□

*Remark* 1.22. The **Lagrange bracket** is defined by

$$(f, g) = X_f(g) = \{f, g\} + gR_\alpha(f).$$

The contact Hamiltonian flow is expressed via Lagrange bracket by $\dot{f} = (H, f)$.

**Proposition 1.23.** *The Jacobi bracket satisfies:*

1) *skew-symmetry:* $\{f, g\} = -\{g, f\}$

2) *Jacobi identity:* $\{\{f, g\}, h\} + \{\{g, h\}, f\} + \{\{h, f\}, g\} = 0$

3) $\{f, gh\} = \{f, g\}h + g\{f, h\} - gh\{f, 1\}$ *and* $\{f, 1\} = -R_\alpha(f)$.

*Thus the Jacobi bracket has all the properties of the Poisson bracket but the Leibniz rule.*

*Proof:* 1) and 2) follow from the Lie algebra structure of $(CVect(M), [\cdot, \cdot])$. We prove 3). The contact vector field of the constant function 1 is $X_1 = R_\alpha$ and

$$\{f, 1\} = -\iota_{X_f}\iota_{R_\alpha}d\alpha + fR_\alpha(1) - R_\alpha(f) = -R_\alpha(f).$$

Using $\{f, g\} = X_f(g) - gR_\alpha(f)$, we have

$$\begin{aligned}\{f, gh\} =& X_f(gh) - ghR_\alpha(f)\\ =& X_f(g)h + gX_f(h) - ghR_\alpha(f)\\ =& \{f, g\}h + g\{f, h\} + ghR_\alpha(f)\\ =& \{f, g\}h + g\{f, h\} - gh\{f, 1\}.\end{aligned}$$

□

# 2 Entropy and Partition Function

## Contents

## 2.1 Entropy in statistical thermodynamics

Entropy is one of the most important concepts in modern Science. There exist several different kinds of entropy which are closely related.

In 1865, Clausius introduced the notion of entropy in thermodynamics to make the Second Law mathematically precise. For a reversible process,

$$dS = \frac{\delta Q}{T}$$

where $T$ is the temperature and $\delta Q$ is the amount of heat transferred to the system. For systems of uniform pressure and temperature without composition change, the energy change by the transferred heat and the work done by the system

$$dU = \delta Q - PdV.$$

Thus we find the fundamental relation for thermodynamic equilibrium

$$dU = TdS - PdV.$$

For a general process (possibly irreversible): $dS \geq \frac{\delta Q}{T}$. For isolated system, the Second Law of Thermodynamics takes the form

$$\Delta S \geq 0$$

and $\Delta S = 0$ for reversible process. Thus the Clausius entropy is a macroscopic theory and measures irreversibility.

In the 1870s, Boltzmann developped the statistical interpretation of entropy. The Boltzmann entropy formula reads

$$S = k_B \log \Omega$$

where $\Omega$ = number of microscopic configurations/states compatible with macroscopic constraints. Here $k_B$ is the Boltzmann constant, $k_B = 1.380649 \times 10^{-23}$ joules/kelvin.

Consider gas in a container. The macroscopic parameters are volume V, pressure P, and temperature T. A microstate consists of a description of the positions, momentums, etc. of all particles. Boltzmann's entropy describes the system such that all microscopic configurations are equally likely, connecting the microscopic and the macroscopic views.

Boltzmann's entropy describes a single macrostate. Gibbs reformulated statistical mechanics in terms of ensembles of systems. The macroscopic state of a system is characterized by a distribution on the microstates. Assume the system has a number of states, each with a probability $p_i$ of being occupied. The Gibbs entropy formula is

$$S = -k_B \sum_i p_i \log p_i.$$

When all states have equal probability $p_i = \frac{1}{\Omega}$, Gibbs entropy formula goes back to Boltzmann entropy formula. In the continuous case, Gibbs entropy is

$$S = -k_B \int \rho(x) \log \rho(x) dx$$

where $\rho(x)$ is the probability density.

## 2.2 Shannon entropy and information

In the 1940s, Shannon established information theory and introduced his information entropy of a random variable based on Boltzmann's H-theorem. Given a probability distribution $\sum p_i = 1$ on a finite set $\mathcal{X}$, the Shannon entropy is defined by

$$H(p) = -\sum_{i \in \mathcal{X}} p_i \log p_i.$$

For a continuous distribution $p(x)$,

$$H(p) = -\int p(x) \log p(x) dx.$$

Shannon entropy matches Gibbs' formula of entropy without the Boltzmann Constant. The interpretation of Shannon entropy is that the information of an event of probability $p$ is

$$I(p) = -\log(p).$$

Then the Shannon entropy $H(p) = \sum_i p_i I(p_i)$ is the average amount of information for all possible events. As Shannon shows, such information $I(p)$ is uniquely determined (up to a choice of base for the log) if we require appropriate continuity and the following natural conditions:

- Monotone decreasing:

  $$p < q \Rightarrow I(p) > I(q), \quad \text{and} \quad I(1) = 0.$$

  This says less probable events have more information, and sure events do not contain information.

- Additivity:

  $$I(pq) = I(p) + I(q).$$

  This says for the combination of independent events, the total information is the sum of their separate informations.

**Proposition 2.1.** *Let $p, q$ be two probability distributions. Then*

$$\lambda H(p) + (1-\lambda)H(q) \leq H(\lambda p + (1-\lambda)q), \quad \textit{for} \quad 0 \leq \lambda \leq 1.$$

*This says that the Shannon entropy is a concave function of probability distributions.*

*Proof:* Since $-x\log x$ is concave function on $(0, +\infty)$, by Jensen's inequality

$$\begin{aligned} \lambda H(p) + (1-\lambda)H(q) &= \sum_i (-\lambda p_i \log p_i - (1-\lambda)q_i \log q_i) \\ &\leq \sum_i -(\lambda p_i + (1-\lambda)q_i)\log(\lambda p_i + (1-\lambda)q_i) \\ &= H(\lambda p + (1-\lambda)q). \end{aligned}$$

□

## 2.3 Relative entropy/Kullback-Leibler divergence

For two probability distributions $p$ and $q$, there is a notion of relative entropy called **Kullback-Leibler (KL) divergence**

$$D_{KL}(p||q) = \sum_{i\in\mathcal{X}} p_i \log \frac{p_i}{q_i}.$$

When $q$ is the uniform distribution ($q_i = \frac{1}{N}$ where $N$ is the cardinality of $\mathcal{X}$),

$$D_{KL}(p||q) = \log N - H(p).$$

**Proposition 2.2** (Gibbs' inequality)**.** *The KL divergence is always nonnegative*

$$D_{KL}(p||q) \geq 0.$$

*The equality holds if and only if $p = q$.*

*Proof:* Since log is a strict concave function, by Jensen's inequality

$$\sum_i p_i \log \frac{q_i}{p_i} \leq \log(\sum_i p_i \frac{q_i}{p_i}) = \log 1 = 0.$$

The equality holds when

$$\frac{q_1}{p_1} = \frac{q_2}{p_2} = \ldots = \frac{q_N}{p_N} \quad \Longrightarrow \quad p = q.$$

□

**Corollary 2.3.** *The maximum of Shannon entropy is*

$$\max_p H(p) = \log N$$

*where* $N = |\mathcal{X}|$.

*Proof:* Let $q$ be the uniform distribution on $\mathcal{X}$. Then $0 \leq D_{KL}(p||q) = \log N - H(p)$. □

Thus if we do not have further constraint, the largest Shannon entropy is related to the Boltzmann entropy. Another way to see this is to look at the optimization problem:

$$\max_p H(p) \quad \text{subject to} \quad \sum p_i = 1.$$

Introduce a Lagrangian multiplier and consider

$$H(p, \lambda) = H(p) - \lambda(\sum_i p_i - 1).$$

Then at the maximum, we have

$$\frac{\partial H(p,\lambda)}{\partial p_i} = -\log p_i - 1 - \lambda = 0 \quad \forall i$$

$$\Rightarrow \quad p_1 = p_2 = \ldots = p_N = \frac{1}{N}.$$

## 2.4 Maximum Entropy Principle

The maximum entropy principle, as stated in 1957 by Jaynes, says that we should choose the distribution that maximizes the Shannon entropy under a given set of constraints. This principle interpreted statistical mechanics as statistical inference, and links it naturally with information theory. We have seen one example above: if there is no further constraint besides $\sum p_i = 1$, then the maximal entropy principle selects the uniform distribution and the Boltzmann entropy.

Suupose we have a function $\varphi : \mathcal{X} \to \mathbb{R}$ and impose a constraint that the average

$$\langle \varphi \rangle := \sum_i p_i \varphi_i = c$$

is fixed by a constant $c$. To find the maximum Shannon entropy given such constraint, we introduce two Lagrangian multipliers

$$H(p,\lambda,\mu) = H(p) - \lambda(\sum_i p_i - 1) - \mu(\sum_i p_i\varphi_i - c)$$

and consider solving

$$\frac{\partial H(p,\lambda,\mu)}{\partial p_i} = -\log p_i - 1 - \lambda - \mu\varphi_i, \quad \forall i$$

$$\implies \quad p_i = \frac{e^{-\mu\varphi_i}}{e^{1+\lambda}}.$$

The constant $\lambda$ is determined by the normalization condition $\sum_i p_i = 1$. Thus we find

$$p_i = \frac{e^{-\mu\varphi_i}}{\sum\limits_j e^{-\mu\varphi_j}}$$

which is the Boltzmann distribution.

In thermodynamics, suppose $\varphi_i = E_i$ is the energy of the state $i$. The thermal equilibrium is achieved by maximizing the entropy subject to a constraint on its expected energy

$$\langle E\rangle = \sum_i p_i E_i = c \quad \text{fixed.}$$

In this case, the Boltzmann distribution reads

$$p_i = \frac{e^{-\beta E_i}}{\sum\limits_j e^{-\beta E_j}}.$$

The constant $\beta$ is related to the temperature by

$$\beta = \frac{1}{k_B T}.$$

Note that the maximum entropy is achieved by the relation

$$dH = \lambda d(\sum p_i - 1) + \beta d(\sum_i p_i E_i - c).$$

Inside the space of probability distributions ($\sum_i p_i = 1$ holds), this relation simplifies

$$dH = \beta d\langle E\rangle.$$

In terms of the Gibbs entropy $S = k_B H$, this gives precisely the Clausius' relation

$$dS = \frac{1}{T} d\langle E\rangle.$$

## 2.5 Partition function

Let us consider the thermal equilibrium that maximizes entropy $S$ subject to the constraint $\langle E\rangle = c$. We will unify notations and write the Boltzmann distribution as

$$p(x) = \frac{e^{-\beta E(x)}}{\int_{\mathcal{X}} e^{-\beta E(x)} dx}.$$

When $\mathcal{X}$ is discrete, we use a discrete measure. The normalization factor

$$Z(\beta) = \int_{\mathcal{X}} e^{-\beta E(x)} dx$$

is called the **partition function**.

Amazingly, the partition function knows everything about the thermal system. For example, the expected energy is

$$\langle E\rangle = \frac{\int_{\mathcal{X}} E e^{-\beta E} dx}{\int_{\mathcal{X}} e^{-\beta E} dx} = -\frac{\partial}{\partial \beta} \log Z.$$

The Gibbs entropy is

$$\begin{aligned} S =& - k_B \int_{\mathcal{X}} p(x) \log p(x) dx \\ =& - \frac{k_B}{Z} \int_{\mathcal{X}} e^{-\beta E(x)} [-\beta E(x) - \log Z] dx \\ =& k_B(\log Z + \beta\langle E\rangle) \\ =& k_B(\log Z - \beta \frac{\partial}{\partial \beta} \log Z). \end{aligned}$$

This formula resembles that for a Legendre transform.

In fact, we define the **free energy**

$$F = -\frac{1}{\beta} \log Z.$$

The above calculation can be expressed via free energy and expected energy by

$$S = -\frac{F}{T} + \frac{\langle E\rangle}{T}.$$

Now

$$F = \langle E\rangle - TS$$

represents the Legendre transform from $S$ to $T$.

## 2.6 Fisher information metric

Let $\mathcal{D}$ be the space of probability distributions on $\mathcal{X}$

$$\mathcal{P} = \{p : \mathcal{X} \to [0,1] | \sum_{i\in\mathcal{X}} p_i = 1\}.$$

We can view $\mathcal{P}$ as embedded in $S^N$ by the equation

$$\sum_{i\in\mathcal{X}}(\sqrt{p_i})^2 = 1.$$

This induces a metric on $P$ by

$$ds^2_{\mathcal{P}} = \sum_i 4(d\sqrt{p_i})^2 = \sum_i \frac{dp_i dp_i}{p_i} = \sum_i p_i d\log p_i d\log p_i.$$

Suppose we consider some statistical ensembles inside $\mathcal{P}$ parametrized by a finite number of parameters $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ and write $p(x|\theta)$ $(x \in \mathcal{X})$ for the corresponding probability density. Then $ds^2_{\mathcal{P}}$ leads to the **Fisher information metric** on the parameter space

$$g_{\alpha\beta}(\theta) = \int_{\mathcal{X}} p(x|\theta)\frac{\partial \log p(x|\theta)}{\partial\theta_\alpha}\frac{\partial \log p(x|\theta)}{\partial\theta_\beta}dx = \left\langle \frac{\partial \log p(x|\theta)}{\partial\theta_\alpha}\frac{\partial \log p(x|\theta)}{\partial\theta_\beta}\right\rangle.$$

Equivalently, using $\int p(x|\theta)dx = 1$,

$$\begin{aligned} g_{\alpha\beta} &= \int_{\mathcal{X}} \frac{\partial p(x|\theta)}{\partial\theta_\alpha}\frac{1}{p(x|\theta)}\frac{\partial p(x|\theta)}{\partial\theta_\beta}dx \\ &= \frac{\partial}{\partial\theta_\alpha}\left[\int_{\mathcal{X}}\frac{\partial p(x|\theta)}{\partial\theta_\beta}dx\right] - \int_{\mathcal{X}} p(x|\theta)\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log p(x|\theta)dx \\ &= -\int_{\mathcal{X}} p(x|\theta)\frac{\partial^2}{\partial\theta_\alpha\partial\theta_\beta}\log p(x|\theta)dx \\ &= -\left\langle\frac{\partial^2 \log p(x|\theta)}{\partial\theta_\alpha\partial\theta_\beta}\right\rangle. \end{aligned}$$

**Example 2.4.** *Consider the Boltzmann/Gibbs distribution.*

$$p(x|\beta) = \frac{1}{Z(\beta)}e^{-\beta E(x)}.$$

*We have*

$$\frac{\partial}{\partial\beta}\log p(x|\beta) = -E(x) - \frac{\partial}{\partial\beta}\log Z(\beta) = -(E(x) - \langle E\rangle).$$

*Thus the Fisher metric*

$$g_{\beta\beta} = \langle (E(x) - \langle E\rangle)^2\rangle = Var(E)$$

*gives the variance of the energy fluctuation.*

The Fisher metric is closely related to the entropy. Let $\theta_0$ be a chosen point in the parameter space and let $\theta$ be close to $\theta_0$

$$\theta = \theta_0 + \Delta\theta,$$

where $\Delta\theta$ is small. Consider the relative entropy $D_{KL}(p(\theta_0)||p(\theta))$ which achieves minimum at $\theta=\theta_0$. Its Hessian matrix at $\theta=\theta_0$ is computed by

$$\begin{aligned}
&\frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_k}D_{KL}(p(\theta_0)\|p(\theta))\bigg|_{\theta=\theta_0}\\
=&\frac{\partial}{\partial\theta_j}\frac{\partial}{\partial\theta_k}\int_{\mathcal{X}}p(\theta_0)\log\frac{p(\theta_0)}{p(\theta)}dx\bigg|_{\theta=\theta_0}\\
=&-\int_{\mathcal{X}}p(\theta_0)\ \frac{\partial^2}{\partial\theta_j\partial\theta_k}\log p(\theta)\bigg|_{\theta=\theta_0}dx\\
=&\ g_{jk}(\theta_0).
\end{aligned}$$

The positivity of $g_{jk}(\theta_0)$ is related to the positivity of Hessian matrix at minimum. Thus locally

$$D_{KL}(p(\theta_0)||p(\theta))=\frac{1}{2}\sum_{j,k}g_{jk}(\theta_0)\Delta\theta_j\Delta\theta_k+O(\Delta\theta^3).$$

**Exponential Family**

Consider the canonical ensemble arising from maximizing Shannon entropy under the constraints

$$\langle\varphi_\alpha(x)\rangle=const,\quad \alpha=1,\dots,M.$$

Here $\varphi_\alpha(x)$'s are statistical quantities whose expectations are fixed. By the same argument as before, the ensembles are given by

$$p(x|\theta)=e^{-\sum\limits_{\alpha}\theta^\alpha\varphi_\alpha(x)}/Z(\theta),$$

where

$$Z(\theta)=\int_{\mathcal{X}}e^{-\sum\limits_{\alpha}\theta^\alpha\varphi_\alpha(x)}dx$$

is the partition function. Such probability distributions are called an exponential family.

Let

$$F(\theta)=-\log Z(\theta)$$

be the free energy. Then

$$p(x|\theta)=e^{-\theta^\alpha\varphi_\alpha(x)+F(\theta)}$$

and

$$\frac{\partial\log p(x|\theta)}{\partial\theta^\alpha}=-\varphi_\alpha(x)+\frac{\partial F}{\partial\theta^\alpha}=-\varphi_\alpha(x)+\langle\varphi_\alpha\rangle.$$

Thus the Fisher metric

$$g_{\alpha\beta}(\theta)=\langle(\varphi_\alpha(x)-\langle\varphi_\alpha\rangle)(\varphi_\beta(x)-\langle\varphi_\beta\rangle)\rangle=Cov(\varphi_\alpha,\varphi_\beta)$$

represents the covariance matrix.

On the other hand,

$$\frac{\partial^2 \log p(x|\theta)}{\partial\theta^\alpha \partial\theta^\beta} = \frac{\partial^2 F}{\partial\theta^\alpha \partial\theta^\beta},$$

thus the Fisher metric can be also written as

$$g_{\alpha\beta} = -\frac{\partial^2 F}{\partial\theta^\alpha \partial\theta^\beta}.$$

The positivity of $g_{\alpha\beta}$ implies that $F$ is a concave function of $\theta^\alpha$.

The Shannon entropy is

$$H(\eta) = -\langle \log p(x|\theta) \rangle = \theta^\alpha \eta_\alpha - F(\theta)$$

where

$$\eta_\alpha = \langle \varphi_\alpha \rangle = \frac{\partial F}{\partial \theta^\alpha}.$$

This represents the Legendre transform of $F$. In particular, $H(\eta)$ is also a concave function. Since

$$\eta_\alpha = \frac{\partial F}{\partial \theta^\alpha},$$

the Jacobian of the coordinate transformation $\theta \to \eta$ is

$$\frac{\partial \eta}{\partial \theta} = \nabla^2 F(\theta).$$

Dually, the Jacobian of the coordinate transformation $\eta \to \theta$ is

$$\frac{\partial \theta}{\partial \eta} = \nabla^2 H(\eta)$$

It follows that $\nabla^2 F = (\nabla^2 H)^{-1}$. Hence the Fisher metric can be also expressed in the dual coordinates $\eta_\alpha$ by

$$g = -(\nabla^2 F)_{\alpha\beta} d\theta^\alpha d\theta^\beta = d\theta^\alpha d\eta_\alpha = -(\nabla^2 H)^{\alpha\beta} d\eta_\alpha d\eta_\beta.$$

## 2.7 Examples

### Classical ideal gas

We consider $N$ particles in a container of volume $V$. Let $(\vec{q}_i, \vec{p}_i)$ represent the $i$-th particle in the phase space and $H(\vec{q}, \vec{p})$ denote the classical Hamiltonian. In the canonical ensemble, the probability density on the phase space is

$$\rho(\vec{q}, \vec{p}) \propto e^{-\beta H(\vec{q}, \vec{p})} d^{3N} q d^{3N} p.$$

We consider the ideal gas ignoring all interactions. The Hamiltonian is

$$H(\vec{q}, \vec{p}) = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m}.$$

Since $d^3qd^3p$ has the unit of $h^3$, where $h$ is Planck's constant, the appropriate volume form on the phase space is

$$\frac{d^{3N}qd^{3N}p}{h^{3N}}.$$

We assume the gas consists of indistinguishable particles and this will introduce an additional $\frac{1}{N!}$ factor. Then the partition function is

$$Z(V,T) = \frac{1}{N!}\int e^{-\beta H(\vec{q},\vec{p})}\frac{d^{3N}qd^{3N}p}{h^{3N}} = \frac{1}{N!h^{3N}}\int e^{-\beta \sum\limits_{i=1}^{N}\frac{\vec{p}_i^2}{2m}}d^{3N}qd^{3N}p.$$

The integral $\int d^{3N}q = V^N$ computes the volume. On the other hand, for each $i$,

$$\int e^{-\frac{\beta}{2m}\vec{p}_i^2}d^3p_i = 4\pi\int_0^\infty e^{-\frac{\beta}{2m}p^2}p^2dp = \left(\frac{2\pi m}{\beta}\right)^{3/2}.$$

Thus the partition function becomes

$$Z(V,T) = \frac{1}{N!}\left[V\left(\frac{2\pi m}{\beta h^2}\right)^{3/2}\right]^N.$$

We obtain the Helmholtz free energy

$$F(V,T) = -\frac{1}{\beta}\log Z(V,T) = -\frac{N}{\beta}\log\left(\frac{V}{N}\left(\frac{h^2\beta}{2\pi m}\right)^{3/2}\right) + \frac{1}{\beta}\log N!\ .$$

When $N\to+\infty$, we can approximate $N!$ using Stirling's formula

$$N! \approx \left(\frac{N}{e}\right)^N\sqrt{2\pi N}$$

which leads to $\log N! \approx N\log N - N$. Thus for $N$ is sufficient large,

$$F \approx -\frac{N}{\beta}\left[\log\left(\frac{V}{N}\left(\frac{h^2\beta}{2\pi m}\right)^{3/2}\right)-1\right] = -Nk_BT\left[\log\left(\frac{V}{N}\left(\frac{h^2}{2\pi mk_BT}\right)^{3/2}\right)-1\right].$$

Using the fundamental relation

$$dF = -SdT - PdV$$

we can derive

$$P = -\frac{\partial F}{\partial V} = \frac{Nk_BT}{V}$$

which gives the classical result for ideal gas

$$PV = Nk_BT.$$

The entropy is given by

$$S = -\frac{\partial F}{\partial T} = Nk_B\left[\log\left(\frac{V}{N}\left(\frac{2\pi mk_BT}{h^2}\right)^{3/2}\right)+\frac{5}{2}\right]$$

and the (expected) energy is

$$U = -\frac{\partial}{\partial\beta}\log Z = F + TS = \frac{3}{2}Nk_BT.$$

Let us compute the corresponding Ruppeneir metric (see Section 1.4). The entropy $S = S(U, V)$ is viewed as a function of $U, V$. We have

$$S = Nk_B\left[\log V + \frac{3}{2}\log U\right] + \text{const}$$

$$\Longrightarrow \qquad g_{UU} = -\partial_U^2 S = \frac{3}{2}\frac{Nk_B}{U^2}, \qquad g_{VV} = -\partial_V^2 S = \frac{Nk_B}{V^2}, \qquad g_{UV} = 0.$$

Thus the Ruppeneir metric is

$$ds^2 = \frac{3}{2}\frac{Nk_B}{U^2}dU^2 + \frac{Nk_B}{V^2}dV^2$$

and its curvature $= 0$.

### Van der Waals gas

The ideal gas model ignores the size of each particle and interactions between them. We can improve the result by approximating the above effects. First, we assume each particle has a finite minimum volume $b$. Then the effective volume of the system is $V - Nb$. Second, we assume each particle experiences a net attractive force to all other particles. Based upon a mean field approximation, the energy associated to each particle can be linearly approximated by $\mathcal{E} = -a\frac{N}{V}$. The partition function is improved to

$$Z = \frac{1}{N!}\left[\frac{V - Nb}{\lambda^3}\right]^N e^{-\beta N\mathcal{E}}$$

where $\lambda = \sqrt{\frac{h^2\beta}{2\pi m}} = \sqrt{\frac{h^2}{2\pi m k_B T}}$ is also called the thermal de Broglie wave length. The Helmholtz free energy becomes

$$F = -\frac{1}{\beta}\log Z \overset{\text{Stirling}}{\approx} Nk_BT\left[\log\left(\frac{N\lambda^3}{V - bN}\right) - 1\right] - a\frac{N^2}{V}.$$

The pressure is now improved to be

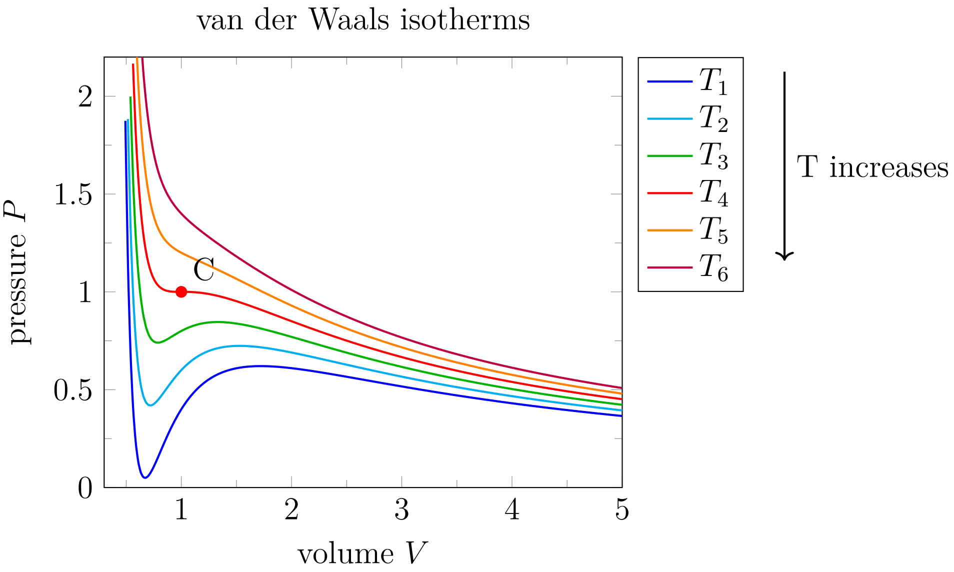
$$P = -\frac{\partial F}{\partial V} = \frac{Nk_BT}{V - bN} - a\frac{N^2}{V^2}$$

i.e.

$$\left(P + a\frac{N^2}{V^2}\right)(V - bN) = Nk_BT.$$

This is the Van der Waals equation of state.

We can plot the isotherms for van der Waals gas in the P-V diagram.



There are parts where the stability condition $\frac{\partial P}{\partial V} < 0$ is violated. This indicates phase transition:

$$\text{the locus}: \quad \frac{\partial P}{\partial V} = 0 \quad \text{(called the spinodal curve)}$$

is the boundary separating different phases. The critical point appears at

$$\frac{\partial P}{\partial V} = 0, \quad \frac{\partial^2 P}{\partial V^2} = 0.$$

From the Van der Waals equation,

$$\begin{aligned}
\frac{\partial P}{\partial V} &= -\frac{Nk_B T}{(V-bN)^2} + 2a\frac{N^2}{V^3} \\
\frac{\partial^2 P}{\partial V^2} &= \frac{2Nk_B T}{(V-bN)^3} - 6a\frac{N^2}{V^4}
\end{aligned}$$

Solving $\frac{\partial P}{\partial V} = \frac{\partial^2 P}{\partial V^2} = 0$, we find the critical point at

$$V_c = 3bN, \quad T_c = \frac{8a}{27bk_B}, \quad P_c = \frac{a}{27b^2}.$$

Let's compute the Ruppeneir metric. We have the entropy

$$S = Nk_B\left[\log\frac{V-bN}{N\lambda^3} + \frac{5}{2}\right] \qquad \text{where} \quad \lambda = \sqrt{\frac{h^2}{2\pi m k_B T}}$$

and the energy

$$U = \frac{3}{2}Nk_B T - a\frac{N^2}{V}$$

Thus we can solve the relation $S = S(U, V)$

$$S = Nk_B\left[\log(V-bN) + \frac{3}{2}\log\left(U + a\frac{N^2}{V}\right)\right] + \text{const}$$

from which we can read off the metric

$$
\begin{aligned}
g^{(R)}_{UU} &= -\frac{\partial^2 S}{\partial U^2} = \frac{3Nk_B}{2(U+aN^2/V)^2}\\
g^{(R)}_{UV} &= -\frac{\partial^2 S}{\partial U \partial V} = -\frac{3Nk_B}{2(U+aN^2/V)^2}\frac{aN^2}{V^2}\\
g^{(R)}_{VV} &= -\frac{\partial^2 S}{\partial V^2} = \frac{Nk_B}{(V-bN)^2} - \frac{3Nk_B}{2}\left[\frac{2aN^2}{V^3(U+aN^2/V)} - \frac{a^2N^4}{V^4(U+aN^2/V)^2}\right]
\end{aligned}
$$

Let us compute the determinant

$$
\begin{aligned}
\det(g^{(R)}) &= g^{(R)}_{uu} g^{(R)}_{vv} - (g^{(R)}_{uv})^2\\
&= \frac{3(Nk_B)^2}{2(U+aN^2/V)^2}\left[\frac{1}{(V-Nb)^2} - \frac{3aN^2}{(U+aN^2/V)V^3}\right]\\
&= \frac{2}{3T^2}\left[\frac{1}{(V-Nb)^2} - \frac{2aN}{k_BTV^3}\right]
\end{aligned}
$$

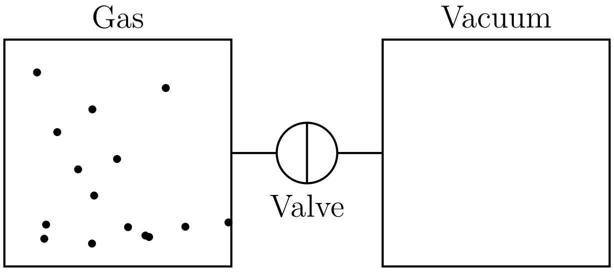
We find

$$
\det g^{(R)} = 0 \iff \left.\frac{\partial P}{\partial V}\right|_T = 0
$$

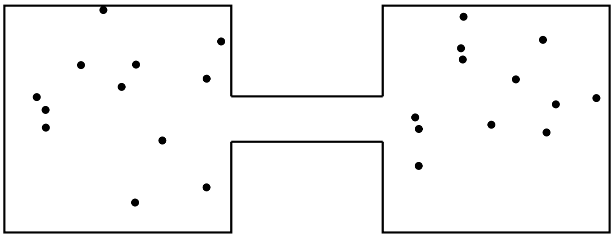
i.e. at the spinodal curve. The curvature becomes $\infty$.

**Joule Effect**

Consider the free expansion of gas



Initial state (valve closed)



Final state (valve open, gas expanded)

from a smaller volume $V_1$ to a larger volume $V_2$. Since the gas expands freely, there is no work done $\Delta W = 0$. For insulated chambers, there is no heat transfer $\Delta Q = 0$. Thus

$$
\Delta U = 0
$$

during the process.

For ideal gas, $\Delta U = 0$ implies that the initial and final equilibrium state has the same temperature. This is not quite true since we expect the free expansion will cool the gas. Let us instead consider Van der Waals gas. $\Delta U = 0$ implies that the only effect during free expansion is a transfer of energy between the potential energy and kinetic energy of the particles. On the other hand, the entropy will increase. For Van der Waals gas, the entropy is

$$S = Nk_B \left[\log\left(\frac{V - bN}{N\lambda^3}\right) + \frac{5}{2}\right]$$

and the energy is

$$U = F + TS = \frac{3}{2}Nk_BT - a\frac{N^2}{V}.$$

Thus the initial temperature $T_1$ and final temperature $T_2$ are related by

$$\frac{3}{2}Nk_BT_1 - a\frac{N^2}{V_1} = \frac{3}{2}Nk_BT_2 - a\frac{N^2}{V_2}$$
$$\implies \quad T_2 - T_1 = \frac{2aN}{3k_B}\left[\frac{1}{V_2} - \frac{1}{V_1}\right].$$

In particular, expansion in volume will cool the gas as expected. This formula gives the right qualitative physics of the Joule effect, but only moderate quantitative accuracy – good at low–moderate densities, poor near criticality or at high pressures.

### Classical solid

Solids can be modelled by an ordered array of atoms on a lattice site. As a first approximation, we can introduce a quadratic potential to trap the atom such that each atom is described by a 3d harmonic oscillator with Hamiltonian

$$H(\vec{p}, \vec{q}) = \frac{\vec{p}^2}{2m} + \frac{m\omega^2}{2}\vec{q}^2.$$

The probability density in the canonical ensemble is

$$\rho(\vec{q}, \vec{p}) \propto e^{-\beta H(\vec{q},\vec{p})}d^{3N}qd^{3N}p.$$

Since these $N$ atoms sit at different places and should be viewed as distinguishable objects. Thus the partition function is

$$Z = \int e^{-\beta\sum_{i=1}^{N}\left(\frac{\vec{p}_i^2}{2m} + \frac{m\omega^2}{2}\vec{q}_i^2\right)}\frac{d^{3N}qd^{3N}p}{h^{3N}} = (Z_1)^N$$

where $Z_1$ is the partition function for a single atom

$$\begin{aligned} Z_1 &= \int e^{-\beta\left(\frac{\vec{p}^2}{2m} + \frac{m\omega^2}{2}\vec{q}^2\right)}\frac{d^3qd^3p}{h^3} \\ &= \frac{1}{h^3}\left[\left(\frac{2\pi m}{\beta}\right)^{3/2}\left(\frac{2\pi}{m\omega^2\beta}\right)^{3/2}\right] = \left(\frac{2\pi}{\beta\omega h}\right)^3. \end{aligned}$$

Therefore

$$Z = \left(\frac{2\pi}{\beta\omega h}\right)^{3N}.$$

The Helmholtz free energy is

$$F = -\frac{1}{\beta}\log Z = -\frac{3N}{\beta}\log\frac{2\pi}{\beta\omega h} = \frac{3N}{\beta}\log\frac{\beta\omega h}{2\pi}.$$

The energy is

$$U = -\frac{\partial}{\partial\beta}\log Z = \frac{3N}{\beta} = 3Nk_BT.$$

The entropy is

$$S = \frac{1}{T}(U-F) = 3Nk_B\left(\log\frac{2\pi k_BT}{\omega h}+1\right).$$

The quantity

$$C_V = \left.\frac{\partial U}{\partial T}\right|_V = 3Nk_B$$

is called the heat capacity. This formula of $C_V$ gives the correct high temperature result but not the correct low temperature result. At low temperature, we need quantum mechanical correction.

**Quantum ideal gas**

We consider ideal gas of $N$ quantum particles. Assume we have quantum states of energies $\epsilon_0, \epsilon_1, \ldots$, for each particle. Then the partition function can be written as

$$Z_N = \sum_{\substack{n_j\geq 0\\ \sum\limits_{j=0}^{\infty} n_j=N}} C(N;n_0,n_1,\ldots)e^{-\beta\left(\sum\limits_{j=0}^{\infty} n_j\epsilon_j\right)}$$

Here the summation means there are $n_j$ particles in the energy state $\epsilon_j$. The combinatorial coefficient $C(N;n_0,n_1,\ldots)$ is a statistical factor. For distinguishable particles,

$$C_{dis}(N;n_0,n_1,\ldots,n_\infty) = \frac{N!}{\prod\limits_{j=0}^{\infty} n_j!}.$$

There are three typical statistics for indistinguishable particles that play important roles in statistical mechanics.

1. Maxwell-Boltzmann

$$C_{MB}(N;n_0,n_1,\ldots) = \frac{C_{dis}(N;n_0,n_1,\ldots)}{N!} = \frac{1}{\prod\limits_{j=0}^{\infty} n_j!}$$

$n_j = 0, 1, 2, \ldots$

2. Bose-Einstein
$$C_{BE}(N; n_0, n_1, \dots) = 1$$
$n_j = 0, 1, 2, \dots$

3. Fermi-Dirac
$$C_{FD}(N; n_0, n_1, \dots) = 1$$
$n_j = 0, 1$

In Bose-Einstein and Fermi-Dirac, each different configuration of occupation numbers $\{n_j\}$ corresponds to one possible state. For Bose-Einstein, $n_j$ can be arbitrary nonnegative integer; for Fermi-Dirac, $n_j = 0$ or 1 according to Pauli's principle. At high temperature, we expect all the occupation numbers to be $n_j \ll N$ (so approximately $n_j = 0$ or 1), and the above three statistics become equivalent at high temperature. At low temperature, they have completely different behavior. We can also consider the grand partition function

$$\mathcal{Z}(V, T, \mu) = \sum_{N=0}^{\infty} Z_N e^{\beta \mu N} = \sum_{n_j \geq 0} C\left(\sum n_j; n_0, n_1, \dots\right) e^{-\beta \sum\limits_{j=0}^{\infty} n_j(\epsilon_j - \mu)}$$

Here $\mu$ is called the chemical potential.

For Maxwell-Boltzmann gas, the grand partition is

$$\mathcal{Z}_{MB}(V, T, \mu) = \prod_{j=0}^{\infty} \exp(e^{-\beta(\epsilon_j - \mu)})$$

and the grand free energy is

$$F_{MB} = -\frac{1}{\beta} \log \mathcal{Z} = -\frac{1}{\beta} \sum_{j=0}^{\infty} e^{-\beta(\epsilon_j - \mu)}.$$

For Bose-Einstein gas, the grand partition is

$$\mathcal{Z}_{BE}(V, T, \mu) = \prod_{j=0}^{\infty} \left( \sum_{n_j=0}^{\infty} e^{-\beta n_j(\epsilon_j - \mu)} \right) = \prod_{j=0}^{\infty} \frac{1}{1 - e^{-\beta(\epsilon_j - \mu)}}$$

and the grand free energy is

$$F_{MB} = -\frac{1}{\beta} \log \mathcal{Z} = \frac{1}{\beta} \sum_{j=0}^{\infty} \log(1 - e^{-\beta(\epsilon_j - \mu)}).$$

For Fermi-Dirac gas, the grand partition is

$$\mathcal{Z}_{FD}(V, T, \mu) = \prod_{j=0}^{\infty} \left( \sum_{n_j=0,1} e^{-\beta n_j(\epsilon_j - \mu)} \right) = \prod_{j=0}^{\infty} (1 + e^{-\beta(\epsilon_j - \mu)})$$

and the grand free energy is

$$F_{FD} = -\frac{1}{\beta} \log \mathcal{Z} = -\frac{1}{\beta} \sum_{j=0}^{\infty} \log(1 + e^{-\beta(\epsilon_j - \mu)}).$$

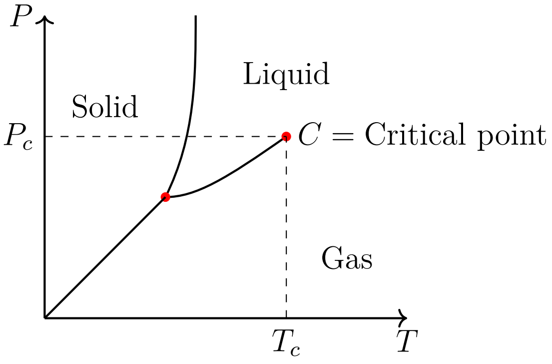
# 3 The Thermodynamic Limit

## Contents

## 3.1 Phase transitions

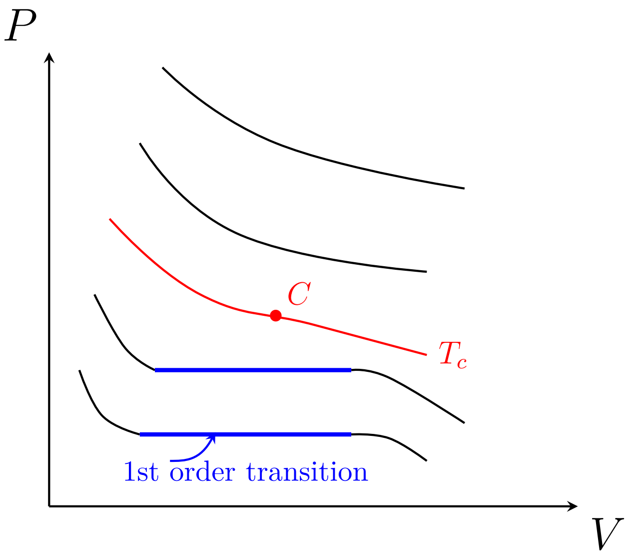
Phase transitions describe situations in which the physics properties of a macroscopic system undergo a qualitative change as external parameters—such as temperature, pressure, or magnetic field—are varied. Such behavior is among the most striking manifestations of collective behavior in *many-body systems*.

One such example is a simple fluid whose phase diagram is shown below. As the temperature and pressure are varied, water can exist as a solid, a liquid or a gas.



Such three stable phases are separated by phase boundaries when two phases coexist. When crossing the coexistence curve by changing the pressure or temperature, there is a jump in the density and a latent heat of the state. This is a typical ”first-order phase transition”.

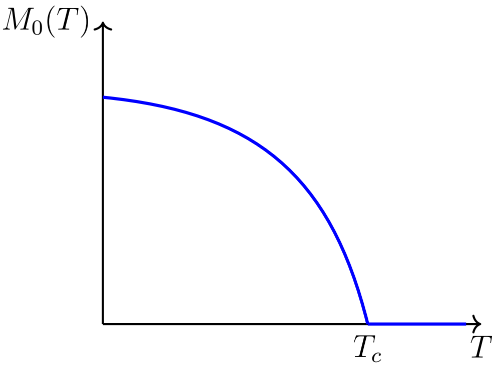
As the temperature increases, the curve of liquid-gas coexistence will end at the critical point $C$ where the distinction between liquid and gas, such as the density, becomes zero. At a temperature above $T_c$, one can pass from the liquid state to the gas state without any discontinuity. The critical point is a typical ”second-order phase transition”. The corresponding isotherms in the $P-V$ diagram is shown below.



Another typical example of phase transition is the ferromagnetic system. A ferromagnetic material (such as iron) consists of microscopic magnetic moments (spins) that tend to align with each other due to interactions. It can be described by a residual magnetization that remains when an external magnetizing field $H$ is reduced to zero.

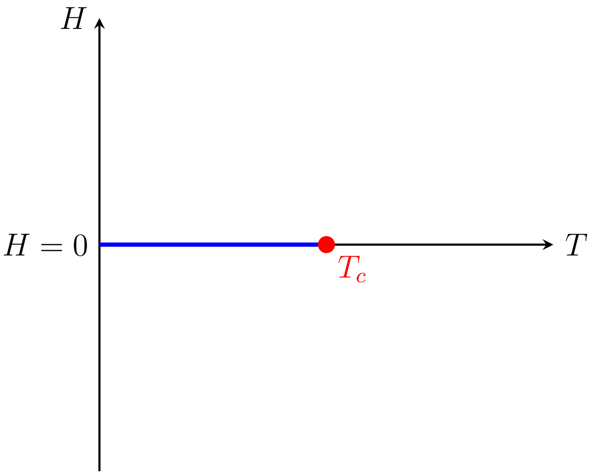
Formally, let $M(T, H)$ denote the equilibrium magnetization in the external magnetic field $H$ at temperature $T$. Then the spontaneous magnetization can be described by

$$M_0(T) = \lim_{H \to 0^+} M(T, H)$$

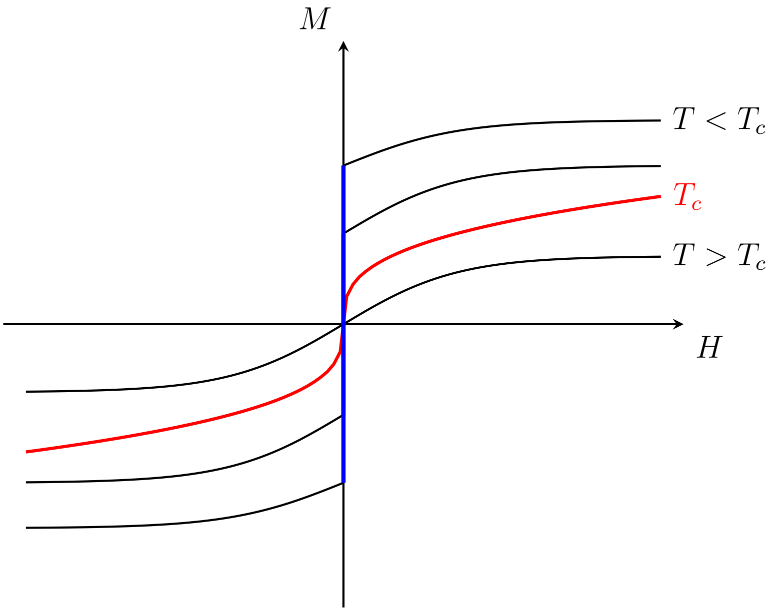


At low temperature, the spontaneous magnetization $M_0(T) \neq 0$ which characterizes a ferromagnetic system. As the temperature is increased, $M_0(T)$ decreases and finally vanishes at a critical temperature $T_c$ known as the Curie temperature. When $T > T_c$, thermal fluctuation dominates and $M_0(T) = 0$.

Note that if we change the sign of the field $H$, the sign of $M_0(T)$ changes. Thus the phase diagram for a ferromagnetic can be shown as below.



The boundary curve for phase transition looks very like the coexistence curve of liquid-gas. If we plot the isotherms in the $M - H$ plane, this picture is again similar to the liquid-gas. This is not a coincidence as we shall see.



As we have discussed, various macroscopic physical quantities can be expressed by the partition function

$$Z = \sum e^{-\beta E}$$

and its derivatives. Usually $E$ will be an analytic function of parameters in the model. Thus $e^{-\beta E}$ will be analytic.

Q: How can derivatives of $Z$ exhibit discontinuity as in the phase transition phenomena?

If $Z$ is a finite sum, this is impossible since a finite sum of analytic functions will still be an analytic function. However, we are dealing with a huge number of particles like $10^{23}$ many. We can treat the partition function as a suitable limit and $\Sigma$ becomes an infinite sum. Mathematically, for an infinite sum,

$$\frac{d}{dx}\sum \neq \sum \frac{d}{dx}$$

derivative may not commute with the sum. Thus it is possible for $Z$ or its derivatives to display discontinuity in parameters in the limit behavior.

What kind of limit shall we take? The physical considerations justify the so-called **thermodynamic limit**. Consider a sequence of finite systems indexed by a region $\Lambda \subset \mathbb{R}^n$. Let $\mathcal{N} = N(\Lambda)$ denote the number of particles inside $\Lambda$. Then the thermodynamic limit is defined formally as:

- $V \to \infty$ (volume goes to infinity)
- $N \to \infty$
- the density $\frac{N(\Lambda)}{|\Lambda|} = \rho$ is fixed.

The free energy per particle is defined by

$$f(\beta, V, N) = \frac{F}{N} = -\frac{1}{N\beta}\log Z$$

It will be also convenient to consider

$$g(\beta, V, N) = \frac{F}{V} = -\frac{1}{V\beta}\log Z = \rho f$$

which can be viewed as the free energy density. Before we analyze the existence problem of thermodynamic limit of $f$ (or $g$) and its properties, let us look at some simple examples.

**Example 3.1** (classical ideal gas)**.** *The partition function is*

$$Z = \frac{1}{N!}\left[\frac{V}{\lambda^3}\right]^N = \frac{N^N}{N!}\left[\frac{V}{N\lambda^3}\right]^N$$

*where* $\lambda = \sqrt{\frac{h^2\beta}{2\pi m}}$. *Thus*

$$\frac{1}{N}\log Z = \log\left(\frac{V}{N\lambda^3}\right) + \frac{1}{N}\log\frac{N^N}{N!}.$$

*In the thermodynamic limit, using Stirling's formula*

$$\lim_{V,N\to\infty} f = \frac{1}{\beta}(\log(\rho\lambda^3) - 1)$$
$$\lim_{V,N\to\infty} g = \frac{\rho}{\beta}(\log(\rho\lambda^3) - 1)$$

**Example 3.2** (Van der Waals gas). *The partition function is*

$$Z = \frac{1}{N!}\left[\frac{V-Nb}{\lambda^3}\right]^N e^{-\beta N\xi} \qquad (\xi = -a\frac{N}{V})$$

*and*

$$\frac{1}{N}\log Z = \log\frac{V-Nb}{N\lambda^3} + a\beta\frac{N}{V} + \frac{1}{N}\log\frac{N^N}{N!}.$$

*In the thermodynamic limit*

$$\lim_{V,N\to\infty} f = \frac{1}{\beta}\left(\log\frac{\lambda^3}{\rho^{-1}-b} - 1\right) - a\rho$$

**Example 3.3.** *We consider a simple mathematical example to illustrate how discontinuity can appear in the thermodynamic limit. Consider the following partition function*

$$Z_N(y) = \frac{(1+y)^N(1-y^N)}{1-y}$$

*where $y$ is the parameter of the model. It is clear that $Z_N(y)$ is analytic in $y$. Consider*

$$g(y) = \lim_{N\to\infty}\frac{1}{N}\log Z_N(y) = \begin{cases}\log(1+y) & \textit{if } y<1\\ \log(1+y)+\log y & \textit{if } y>1\end{cases}$$

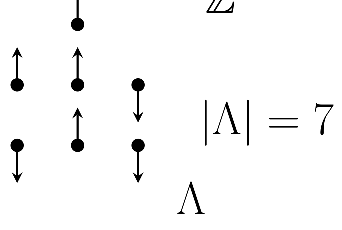
*$g(y)$ is continuous but*

$$y\frac{\partial}{\partial y}g(y) = \begin{cases}\frac{y}{1+y} & \textit{if } y<1\\ \frac{y}{1+y}+1 & \textit{if } y>1\end{cases}$$

*displays a jump at $y=1$ which can be viewed as a first-order phase transition. As we will see, this is a special case of Lee-Yang theory.*

## 3.2 Thermodynamic limit of lattice spin systems

Consider the d-dim lattice $\mathbb{Z}^d$. Let $S$ be a finite set (e.g. $S = \{\pm 1\}$) representing the configuration at each site of the lattice. For each finite region $\Lambda \subset \mathbb{Z}^d$, we denote $\Omega_\Lambda = S^\Lambda$ for all configurations on $\Lambda$. We use $|\Lambda|$ for the number of sites inside $\Lambda$.



For any $\Lambda$, we associate a Hamiltonian function

$$H_\Lambda : \Omega_\Lambda \to \mathbb{R}.$$

We assume $H_\Lambda$ has the form

$$H_\Lambda(\sigma) = \sum_{X \subset \Lambda} \Phi(X, \sigma_X)$$

where $\Phi$ represents the interaction. For example, if $\Phi(X, \cdot)$ is only non-zero for $|X| = 2$, then this represents a pair-wise interaction. The partition function on $\Lambda$ is defined by

$$Z_\Lambda = \sum_{\sigma \in \Omega_\Lambda} e^{-\beta H_\Lambda(\sigma)}$$

and we denote the following free energy density

$$g_\Lambda(\beta) = \frac{1}{|\Lambda|} \log Z_\Lambda.$$

We make two assumptions on the interactions:

(A) **Translation invariance**:

$$\Phi(X + y, \sigma^{(y)}) = \Phi(X, \sigma), \quad \forall y \in \mathbb{Z}^d.$$

Here $\sigma^{(y)}(x + y) = \sigma(x),\ \forall x \in X$.

(B) **Finite range**: $\exists R > 0$ such that

$$\Phi(X) = 0 \quad \text{if} \quad \operatorname{diam}(X) > R.$$

Here $\operatorname{diam}(X)$ is the diameter of $X$.

This finite range condition can be relaxed by decay properties of $\Phi(X)$ when $\operatorname{diam}(X) \to \infty$. We will focus on finite range situations to simplify the discussion and illustrate the main idea.

**Definition 3.4.** We write $\Lambda_n \nearrow \mathbb{Z}^d$ for an increasing sequence $\Lambda_1 \subset \Lambda_2 \subset \cdots \subset \Lambda_n \subset \dots$ of finite regions such that $\bigcup_{n=1}^{\infty} \Lambda_n = \mathbb{Z}^d$. A sequence $\Lambda_n \nearrow \mathbb{Z}^d$ is called a **Van Hove Sequence** if

$$\frac{|\partial \Lambda_n|}{|\Lambda_n|} \to 0$$

where $\partial \Lambda_n$ is the boundary of $\Lambda_n$.

**Example 3.5.** *The sequence of standard cubes*

$$Q_n = [-n, n]^d \cap \mathbb{Z}^d$$

*is a Van Hove Sequence.*

**Theorem 3.6.** *Assume $\Phi$ is translation invariant and finite range. Let $\Lambda_n$ be a Van Hove Sequence. Then $P(\beta) = \lim\limits_{n\to\infty} \frac{1}{|\Lambda_n|} \log Z_{\Lambda_n}$ exists and is independent of the shape of $\Lambda_n$.*

*Proof:* Since $\Phi$ is translation invariant and finite range:

$$C_0 := \sup_{x,\sigma_x} |\Phi(x, \sigma_x)| < \infty$$

$$N_R := \#\{X \ni 0 \text{ such that } \Phi(X) \neq 0\} < \infty$$

## Step 1: Stability and Boundedness

For any finite region $\Lambda$,

$$|H_\Lambda(\sigma)| = \left| \sum_{X \subset \Lambda} \Phi(X, \sigma_X) \right| \leq \sum_{y \in \Lambda} \sum_{X \ni y} |\Phi(X, \sigma_X)| \leq C_0 N_R |\Lambda|.$$

Thus

$$Z_\Lambda = \sum_{\sigma \in \Omega_\Lambda} e^{-\beta H_\Lambda(\sigma)} \leq \sum_{\sigma \in \Omega_\Lambda} e^{\beta C_0 N_R |\Lambda|} = |S|^{|\Lambda|} e^{\beta C_0 N_R |\Lambda|}.$$

Similarly,

$$Z_\Lambda \geq |S|^{|\Lambda|} e^{-\beta C_0 N_R |\Lambda|}.$$

It follows that $P_\Lambda$ is uniformly bounded:

$$-\beta C_0 N_R + \log |S| \leq P_\Lambda \leq \beta c_0 N_R + \log |S|$$

$$\Rightarrow \quad |P_\Lambda| \leq C_1$$

for some constant $C_1$ independent of $\Lambda$.

## Step 2: Almost additivity

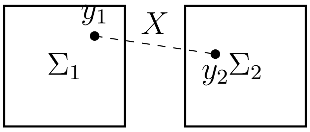
Let $\Sigma_1, \Sigma_2 \subset \mathbb{Z}^d$ be disjoint and $\Sigma = \Sigma_1 \cup \Sigma_2$. We can decompose the Hamiltonian as

$$H_\Sigma(\sigma) = H_{\Sigma_1}(\sigma) + H_{\Sigma_2}(\sigma) + H_{int}(\sigma)$$

where

$$H_{int}(\sigma) = \sum_{\substack{X \cap \Sigma_1 \neq \emptyset \\ X \cap \Sigma_2 \neq \emptyset \\ X \subset \Sigma}} \Phi(X, \sigma_X)$$

represents terms that intersect both $\Sigma_1$ and $\Sigma_2$.



Here for nonzero $\Phi(X)$ in $H_{int}$, $\exists y_1 \in X \cap \Sigma_1, y_2 \in X \cap \Sigma_2$ such that $\mathrm{dist}(y_1, y_2) \leq R$.

Define

$$\partial_a \Sigma = \{x \in \Sigma \mid \mathrm{dist}(x, \Sigma) \leq a\}.$$

Then $y_2 \in \partial_R \Sigma_1$ and $y_1 \in \partial_R \Sigma_2$. We have the estimate

$$|H_{int}(\sigma)| \leq \sum_{y \in \partial_R \Sigma_1} \sum_{X \ni y} |\Phi(X, \sigma_X)| \leq |\partial_R \Sigma_1| N_R C_0 \leq C(R,d) |\partial \Sigma_1| N_R C_0$$

where $C(R,d)$ is a constant depending on $R$ and $d$. Thus there exists a constant $C_2$ independent of $\Sigma$ such that

$$\beta |H_{int}| \leq C_2 |\partial \Sigma_1|.$$

so

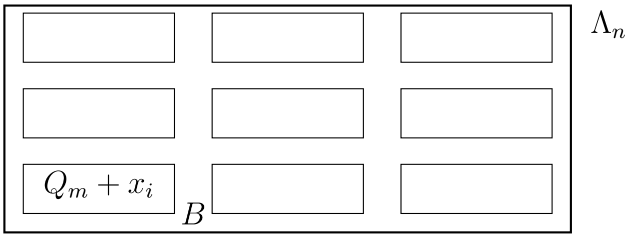
$$e^{-C_2|\partial\Sigma_1|} Z_{\Sigma_1} Z_{\Sigma_2} \leq Z_\Sigma \leq e^{C_2|\partial\Sigma_1|} Z_{\Sigma_1} Z_{\Sigma_2}.$$

We derive the following almost additivity property:

$$|\log Z_\Sigma - \log Z_{\Sigma_1} - \log Z_{\Sigma_2}| \leq C_2 |\partial \Sigma_1|.$$

## Step 3: Tiling Estimate

Let $Q_m = [-m, m]^d \cap \mathbb{Z}^d$ be the Van Hove sequence of standard cubes. We fix $m$ and consider a tiling of $\Lambda_n$ by $Q_m$ for $n$ large enough.



We can find disjoint translates $Q_m + x_1, Q_m + x_2, \ldots, Q_m + x_k \subset \Lambda_n$ and a remainder set $B \subset \Lambda_n$ such that

$$\Lambda_n = \bigcup_{i=1}^k (Q_m + x_i) \cup B$$

$$|B| \leq C(m) |\partial \Lambda_n|$$

where $C(m) \sim \mathrm{diam}(Q_m)$. By the almost additivity and translation invariance:

$$\begin{aligned}\log Z_{\Lambda_n} &= \sum_{i=1}^k \log Z_{Q_m + x_i} + \log Z_B + \mathcal{E}_{n,m} \\ &= k \log Z_{Q_m} + \log Z_B + \mathcal{E}_{n,m}\end{aligned}$$

where the error term

$$|\mathcal{E}_{n,m}| \leq C_2 k |\partial Q_m| + C(m)|\partial \Lambda_n| \leq C_2 |\Lambda_n| \frac{|\partial Q_m|}{|Q_m|} + C(m)|\partial \Lambda_n|.$$

Divide $\log Z_{\Lambda_n}$ by $|\Lambda_n|$, we find

$$P_{\Lambda_n} = \frac{k|Q_m|}{|\Lambda_n|} P_{Q_m} + \frac{|B|}{|\Lambda_n|} P_B + \frac{\mathcal{E}_{n,m}}{|\Lambda_n|} = P_{Q_m} + \frac{|B|}{|\Lambda_n|}(P_B - P_{Q_m}) + \frac{\mathcal{E}_{n,m}}{|\Lambda_n|}.$$

Take $\limsup\limits_n$ and use the boundedness of $P$ and the Van Hove property:

$$\limsup_n P_{\Lambda_n} \leq P_{Q_m} + C_2 \frac{|\partial Q_m|}{|Q_m|}$$

which holds for any fixed $m$. Now we take $\liminf\limits_m$ and use the Van Hove property of $Q_m$

$$\Rightarrow \quad \limsup_n P_{\Lambda_n} \leq \liminf_m P_{Q_m}.$$

Similarly, if we take $\liminf\limits_n$ first

$$\liminf_n P_{\Lambda_n} \geq P_{Q_m} - C_2 \frac{|\partial Q_m|}{|Q_m|}$$

and then take $\limsup\limits_m$, we get

$$\liminf_n P_{\Lambda_n} \geq \limsup_m P_{Q_m}.$$

It follows that $\lim\limits_n P_{\Lambda_n}, \lim\limits_m P_{Q_m}$ exist and

$$\lim_n P_{\Lambda_n} = \lim_m P_{Q_m}.$$

The equality $\lim_n P_{\Lambda_n} = \lim_m P_{Q_m}$ implies that the thermodynamic limit is independent of the shape of the chosen Van Hove sequence. ☐

*Remark* 3.7. The finite range condition can be relaxed to the absolutely summable condition:

$$\sum_{X \ni 0} \frac{1}{|X|} \sup_{\sigma_X} |\Phi(X, \sigma_X)| < \infty$$

The proof is similar. See [39].

## 3.3 Thermodynamic limit of continuous particle systems

We consider a classical continuous system of $N$ particles in a finite region $\Omega \subset \mathbb{R}^d$ with Hamiltonian

$$\mathcal{H}_N = \sum_{i=1}^N \frac{\vec{p}_i^{\,2}}{2m} + \mathcal{U}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) \qquad \vec{x}_i \in \mathbb{R}^d.$$

The partition function reads

$$\begin{aligned} Z_N(\beta,\Omega) &= \frac{1}{N!}\int_{\Omega^N\times\mathbb{R}^{dN}} e^{-\beta\mathcal{H}_N}\frac{d\vec{x}d\vec{p}}{h^{dN}} \\ &= \frac{1}{N!}\left[\int_{\mathbb{R}^d} e^{-\beta\frac{\vec{p}^2}{2m}}\frac{d\vec{p}}{h^d}\right]^N \int_{\Omega^N} d\vec{x}_1\dots d\vec{x}_N e^{-\beta\mathcal{U}(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_N)} \\ &= \frac{1}{\lambda^{dN}}Y_N(\beta,\Omega) \end{aligned}$$

where $\lambda = \left(\frac{\beta h^2}{2m}\right)^{\frac{1}{2}}$ and

$$Y_N = \frac{1}{N!}\int_{\Omega^N} d\vec{x}_1\dots d\vec{x}_N e^{-\beta\mathcal{U}(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_N)}.$$

Let $|\Omega|$ denote the volume of $\Omega$ and

$$\rho = \frac{N}{|\Omega|}$$

be the particle density to be fixed in the thermodynamic limit. Let

$$g(\beta,\rho,\Omega) = \frac{1}{|\Omega|}\log Z_N = \frac{1}{|\Omega|}\log Y_N - \rho d\log\lambda$$

and we are interested in the existence of its thermodynamic limit

$$g(\beta,\rho) \overset{?}{:=} \lim_{\substack{|\Omega|\to\infty \\ N=\rho|\Omega|}} \frac{1}{|\Omega|}\log Z_N(\beta,\Omega).$$

We follow the strategy as discussed in Fisher [13] and Ruelle [39].

**Definition 3.8.** We say the potential $\mathcal{U}$ is stable if there exists $B>0$ such that

$$\mathcal{U}(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_N) \geq -BN \qquad \forall N.$$

We will give a class of stable pair interaction in Example 3.15 below.

**Proposition 3.9.** *Assume* $\mathcal{U}$ *is stable. Then*

$$g(\beta,\rho,\Omega) \leq \rho(1+\beta B - \log(\rho\lambda^d))$$

*is bounded from above.*

*Proof:* For stable $\mathcal{U}$,

$$Y_N \leq \frac{|\Omega|^N}{N!}e^{\beta BN} \qquad \forall N.$$

Using $\log N! \geq N\log N - N$, we find

$$\begin{aligned} g(\beta,\rho,\Omega) &\leq \frac{1}{|\Omega|}\left[\beta BN + N\log|\Omega| - N\log N + N\right] - \rho d\log\lambda \\ &= \rho\left[1+\beta B - \log(\rho\lambda^d)\right]. \end{aligned}$$

$\square$

**Definition 3.10.** Let $\Phi_{N_1,N_2}$ denote the following mutual potential

$$\begin{aligned}&\Phi_{N_1,N_2}(\vec{x}'_1,\dots,\vec{x}'_{N_1};\vec{x}''_1,\dots,\vec{x}''_{N_2})\\ =&\mathcal{U}_{N_1+N_2}(\vec{x}'_1,\dots,\vec{x}'_{N_1},\vec{x}''_1,\dots,\vec{x}''_{N_2})-\mathcal{U}_{N_1}(\vec{x}'_1,\dots,\vec{x}'_{N_1})-\mathcal{U}_{N_2}(\vec{x}''_1,\dots,\vec{x}''_{N_2}).\end{aligned}$$

We say $\mathcal{U}$ is <u>tempered</u> if $\exists \varepsilon>0, R_0>0, A\geq 0$ such that

$$\Phi_{N_1,N_2}(\vec{x}'_1,\dots,\vec{x}'_{N_1};\vec{x}''_1,\dots,\vec{x}''_{N_2})\leq N_1N_2Ar^{-d-\varepsilon}$$

whenever $|\vec{x}'_i-\vec{x}''_j|\geq r\geq R_0$ for all $i=1,\dots,N_1$, $j=1,\dots,N_2$.

**Example 3.11.** *Assume the potential consists of pair interaction*

$$\mathcal{U}(\vec{x}_1,\dots,\vec{x}_N)=\sum_{1\leq i<j\leq N}\phi(\vec{x}_j-\vec{x}_i).$$

*Then the mutual potential is*

$$\Phi_{N_1,N_2}(\vec{x}'_1,\dots,\vec{x}'_{N_1};\vec{x}''_1,\dots,\vec{x}''_{N_2})=\sum_{\substack{1\leq i\leq N_1\\1\leq j\leq N_2}}\phi(\vec{x}''_j-\vec{x}'_i).$$

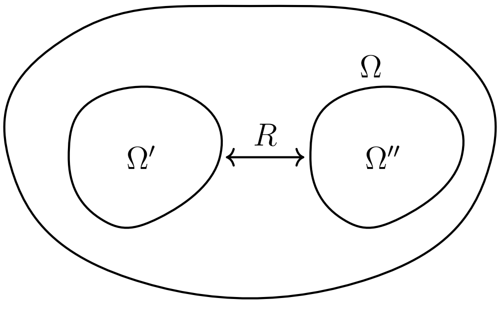
*If the pair interaction $\phi$ has the property*

$$\phi(\vec{x})\leq\frac{A}{|\vec{x}|^{d+\varepsilon}}\qquad \textit{for } |\vec{x}|\geq R_0,$$

*then $\mathcal{U}$ is tempered. In such case the tail*

$$\int_{|\vec{x}|\geq r}\phi(\vec{x})d\vec{x}\leq A\int_{|\vec{x}|\geq r}\frac{d\vec{x}}{|\vec{x}|^{d+\varepsilon}}\xrightarrow{r\to+\infty}0$$

*which says the positive part of the interaction energy at large distance is negligible.*



**Proposition 3.12.** *Assume $\mathcal{U}$ is tempered. Let $\Omega',\Omega''\subset\Omega$ be two disjoint subdomains of $\Omega$ separated by a distance $R\geq R_0$. Then*

$$Z_{N'+N''}(\Omega)\geq Z_{N'}(\Omega')Z_{N''}(\Omega'')e^{-N'N''\beta A/R^{d+\varepsilon}}.$$

*Proof:* We can split the integral in $Z_N(\Omega)$

$$\begin{aligned}
Z_N(\Omega) &= \frac{\lambda^{-dN}}{N!}\int_{\Omega^N} d\vec{x}_1\dots d\vec{x}_N e^{-\beta\mathfrak{U}(\vec{x}_1,\dots,\vec{x}_N)}\\
&\geq \frac{\lambda^{-dN}}{N!}\sum_{N'+N''=N}\frac{N!}{N'!N''!}\int_{(\Omega')^{N'}} d\vec{x}'_1\dots d\vec{x}'_{N'}\int_{(\Omega'')^{N''}} d\vec{x}''_1\dots d\vec{x}''_{N''}\\
&\qquad e^{-\beta\mathfrak{U}(\vec{x}'_1,\dots,\vec{x}'_{N'})}e^{-\beta\mathfrak{U}(\vec{x}''_1,\dots,\vec{x}''_{N''})}e^{-\beta\Phi_{N'N''}(\vec{x}'_1,\dots,\vec{x}'_{N'};\vec{x}''_1,\dots,\vec{x}''_{N''})}\\
&\geq Z_{N'}(\Omega')Z_{N''}(\Omega'')e^{-\beta N'N''A/R^{d+\varepsilon}}.
\end{aligned}$$

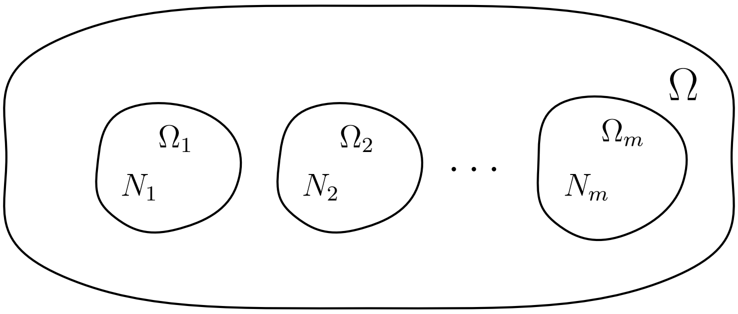
□

We will assume $\mathfrak{U}$ is stable and tempered. Proposition 3.12 implies

$$g(\beta,\rho,\Omega)\geq\frac{|\Omega'|}{|\Omega|}g(\beta,\rho',\Omega')+\frac{|\Omega''|}{|\Omega|}g(\beta,\rho'',\Omega'')-\frac{N'N''}{|\Omega|}\beta A/R^{d+\varepsilon}$$

where

$$\rho=\frac{N}{|\Omega|},\quad \rho'=\frac{N'}{|\Omega'|},\quad \rho''=\frac{N''}{|\Omega''|}.$$



In general, suppose we have disjoint subdomains

$$\Omega_1\cup\Omega_2\cup\dots\cup\Omega_m\subset\Omega$$

which are separated from each other by a distance $R\geq R_0$. Assume $\Omega_i$ contains $N_i$ particles and $\rho_i=N_i/|\Omega_i|$. Let

$$\rho=\sum_{i=1}^m w_i\rho_i \qquad w_i=\frac{|\Omega_i|}{|\Omega|}.$$

By successively applying Proposition 3.12,

$$\begin{aligned}
g(\beta,\rho,\Omega) &\geq \sum_{i=1}^m w_i g(\beta,\rho_i,\Omega_i)-\frac{\beta A}{|\Omega|R^{d+\varepsilon}}\sum_{i=1}^m N_i(N_{i+1}+\dots+N_m)\\
&\geq \sum_{i=1}^m w_i g(\beta,\rho_i,\Omega_i)-\frac{\beta AN^2}{|\Omega|R^{d+\varepsilon}}
\end{aligned}$$

we obtain

$$g(\beta,\rho,\Omega)\geq\sum_{i=1}^m w_i g(\beta,\rho_i,\Omega_i)-\beta A\rho^2\frac{|\Omega|}{R^{d+\varepsilon}}. \qquad (\dagger)$$

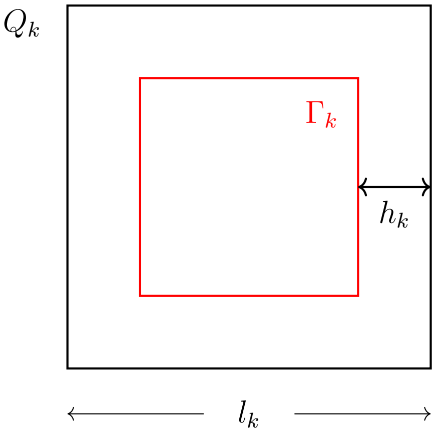
### Thermodynamic limit for standard cubes

Let $Q_k$ be a cube of length

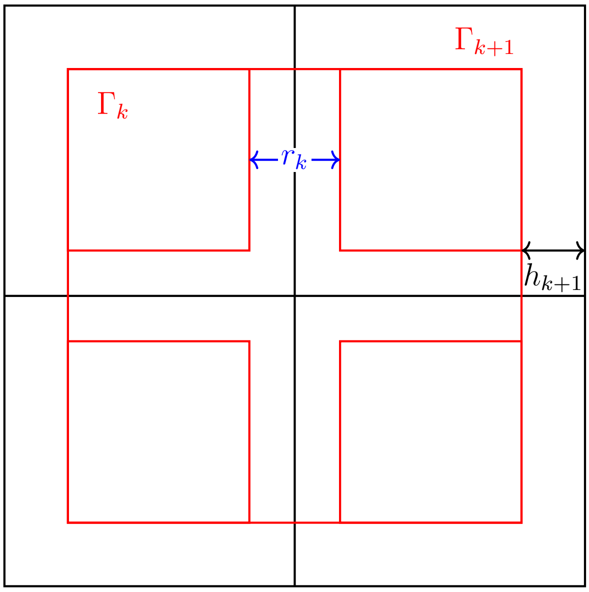
$$l_k = 2^k l_0.$$

Let $\Gamma_k$ be a cube of length $l_k - 2h_k$ where

$$h_k = (2\theta)^k h_0, \qquad \frac{1}{2} < \theta < 1.$$



Now $Q_{k+1}$ contains $2^d$ many $Q_k$'s and we place $\Gamma_k$ inside $\Gamma_{k+1}$ as pictured



Such $\Gamma_k$'s are separated by a distance

$$r_k = (l_{k+1} - 2h_{k+1}) - 2(l_k - 2h_k) = 2^{k+2}\theta^k(1-\theta)h_0.$$

We choose $h_0$ large enough such that $r_k \geq R_0$ for all $k \geq 0$.

We place a particle density $\rho'$ inside $2^{d-1}$ $\Gamma_k$'s and $\rho''$ inside $2^{d-1}$ $\Gamma_k$'s inside $\Gamma_{k+1}$. Applying the above inequality (†), we obtain

$$g(\beta, \rho, \Gamma_{k+1}) \geq \frac{1}{2} g(\beta, \rho', \Gamma_k) + \frac{1}{2} g(\beta, \rho'', \Gamma_k) - \beta A \rho^2 \xi_{k+1}.$$

Here

$$\rho = \frac{1}{2}\rho' + \frac{1}{2}\rho''$$

and

$$\xi_{k+1} = \frac{|\Gamma_{k+1}|}{r_k^{d+\varepsilon}} = \frac{\left(2^{k+1}(l_0 - \theta^k h_0)\right)^d}{\left(2^{k+2}\theta^k(1-\theta)h_0\right)^{d+\varepsilon}} \le \left(\frac{1}{2^\varepsilon \theta^{d+\varepsilon}}\right)^k \frac{(2l_0)^d}{(4(1-\theta)h_0)^{d+\varepsilon}}.$$

We will choose $\frac{1}{2} < \theta < 1$ such that $2^\varepsilon\theta^{d+\varepsilon} = (2\theta)^\varepsilon\theta^d > 1$. This implies

$$\sum_{k=1}^{\infty} \xi_k < +\infty.$$

To prove the existence of thermodynamic limit, we set $\rho' = \rho'' = \rho$ to get

$$g(\beta, \rho, \Gamma_{k+1}) \ge g(\beta, \rho, \Gamma_k) - \beta A \rho^2 \xi_{k+1}.$$

This implies that the sequence

$$y_k = g(\beta, \rho, \Gamma_k) + \beta A \rho^2 \sum_{i=1}^{k} \xi_i$$

is monotonic increasing. By Proposition 3.9, $y_k$ is bounded from above, so the limit

$$\lim_{k\to\infty} y_k \quad \text{exists.}$$

Hence the thermodynamic limit

$$g(\beta, \rho) := \lim_{k\to\infty} g(\beta, \rho, \Gamma_k) = \lim_{k\to\infty} y_k - \beta A \rho^2 \sum_{i=1}^{k} \xi_i$$

exists. Moreover, by taking the limit $k \to \infty$ of the relation

$$g(\beta, \rho, \Gamma_{k+1}) \ge \frac{1}{2} g(\beta, \rho', \Gamma_k) + \frac{1}{2} g(\beta, \rho'', \Gamma_k) - \beta A \rho^2 \xi_{k+1}$$

we obtain

$$g(\beta, \rho) \ge \frac{1}{2} g(\beta, \rho') + \frac{1}{2} g(\beta, \rho'')$$

which says that $g(\beta, \rho)$ is concave in $\rho$.

To show the existence of thermodynamic limit of general Van Hove sequences and independence of the shape of the domains, we can perform a similar tiling argument as in the lattice case. We refer to Fisher [13] and Ruelle [39] for details.

**Proposition 3.13.** *The thermodynamic limit $g(\beta, \rho)$ has the following properties:*

*(1) $g$ is a continuous concave function in $\rho$ in a defining domain $0 \le \rho < \rho_{max}$ ($\rho_{max}$ could be $+\infty$). In particular, the left and right derivatives*

$$\partial_\rho^- g := \lim_{\delta\to 0^-} \frac{g(\rho+\delta) - g(\rho)}{\delta}, \quad \partial_\rho^+ g := \lim_{\delta\to 0^+} \frac{g(\rho+\delta) - g(\rho)}{\delta}$$

*are finite and exist everywhere, and $\partial_\rho^- g \ge \partial_\rho^+ g$. The equality holds almost everywhere.*

*(2) The free energy per particle*

$$f(\beta,\rho) = -\lim_N \frac{1}{N\beta}\log Z_N = -\frac{1}{\beta\rho}g(\beta,\rho)$$

*is a non-decreasing function in $\rho$.*

*Proof:* (1) $g$ is concave, hence continuous in the interior of its defining domain. We compute its limit at $\rho = 0$. Since

$$y_k = g(\beta,\rho,\Gamma_k) + \beta A\rho^2 \sum_{i=1}^{k} \xi_i$$

is non-decreasing,

$$g(\beta,\rho) \geq g(\beta,\rho,\Gamma_k) + \beta A\rho^2 \sum_{i=k+1}^{\infty} \xi_i.$$

This implies

$$\liminf_{\rho\to 0} g(\beta,\rho) \geq 0.$$

On the other hand, the upper bound from stability

$$g(\beta,\rho) \leq \rho(1 + \beta B - \log(\rho\lambda^d))$$

implies

$$\limsup_{\rho\to 0} g(\beta,\rho) \leq 0.$$

It follows that

$$\lim_{\rho\to 0} g(\beta,\rho) = 0.$$

(2) By concavity, for any $0 \leq \lambda \leq 1$

$$g(\beta,\lambda\rho) \geq (1-\lambda)g(\beta,0) + \lambda g(\beta,\rho) = \lambda g(\beta,\rho)$$

$$\implies \quad f(\beta,\lambda\rho) \leq f(\beta,\rho)$$

□

*Remark* 3.14. Let $v = \frac{1}{\rho}$ denote the specific volume. Then the pressure

$$P = -\frac{\partial f}{\partial v} = \rho^2 \frac{\partial f}{\partial \rho}$$

exists almost everywhere and $\geq 0$. In terms of $g$

$$P = -\frac{\rho^2}{\beta}\frac{\partial}{\partial\rho}\left(\frac{1}{\rho}g\right) = \frac{1}{\beta}\left(g - \rho\frac{\partial g}{\partial\rho}\right)$$

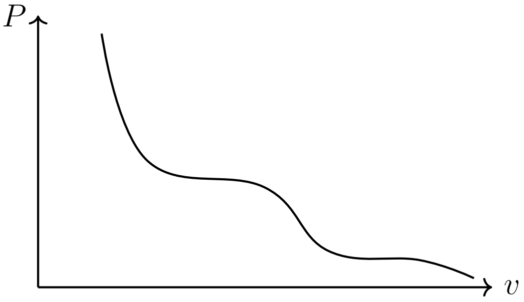
thus

$$\frac{\partial P}{\partial\rho} = -\frac{\rho}{\beta}\frac{\partial^2 g}{\partial\rho^2} \geq 0 \qquad \text{by concavity}$$

i.e.

$$\frac{\partial P}{\partial v} \leq 0.$$

So the pressure is non-increasing in $v$. The isotherm in the $P - V$ diagram looks like



**Example 3.15.** *A potential function* $\phi(\vec{x})$ *satisfying*

$$\phi(\vec{x}) = +\infty \qquad \textit{if} \quad |\vec{x}| \leq r_0$$

*is called a* <u>*hard core*</u> *potential. This condition says that particles are not allowed to approach each other within a distance of* $r_0$*. We further assume*

$$\phi(\vec{x}) \geq -f(|\vec{x}|) \qquad \textit{for } |\vec{x}| \geq r_0$$

*where*

- $f(r) \geq 0$ *and decreasing*
- $f(r)$ *satisfies*
$$\int_{r_0}^{+\infty} f(r) r^{d-1} dr < +\infty.$$

*We show that a pair hard core interaction satisfying the above condition is stable.*

In fact,

$$\mathcal{U}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) = \sum_{i<j} \phi(\vec{x}_j - \vec{x}_i) = \frac{1}{2} \sum_{i=1}^{N} \left( \sum_{j \neq i} \phi(\vec{x}_j - \vec{x}_i) \right).$$

We can assume $|\vec{x}_j - \vec{x}_i| \geq r_0$ for all $i \neq j$. Then

$$\mathcal{U}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) \geq -\frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i} f(|\vec{x}_j - \vec{x}_i|).$$

Thus it suffices to prove

$$\sup_{|\vec{x}_j - \vec{x}_i| \geq r_0} \sum_{j \neq i} f(|\vec{x}_j - \vec{x}_i|) \leq 2B.$$

Let us fix a particle $x_i = 0$ and consider

$$S = \sum_{j \neq i} f(|\vec{x}_j - \vec{x}_i|).$$

Let $\delta > 0$ and we consider the shell

$$A_k = \{\vec{x} \in \mathbb{R}^d \mid k\delta \leq |\vec{x}| < (k+1)\delta\}.$$

Because of hard core, each particle $\vec{x}_j$ occupies a volume $v_0 \sim r_0^d$. Thus the number of $\vec{x}_j$'s inside the shell $A_k$ is bounded by

$$\#\{\vec{x}_j \in A_k\} \leq \frac{\mathrm{Vol}(A_k)}{v_0} \leq C_1 \frac{((k+1)^d - k^d)\delta^d}{r_0^d} \leq C_2 k^{d-1}\delta^d$$

for some constant $C_2$ depending on $d$ and $r_0$. By the decreasing property of $f(r)$

$$\sum_{\vec{x}_j \in A_k} f(|\vec{x}_j - \vec{x}_i|) \leq f(k\delta)\#\{\vec{x}_j \in A_k\} \leq C_2 k^{d-1}\delta^d f(k\delta).$$

Summing up all shells $A_k$,

$$S \leq C_2 \sum_{k \geq \frac{r_0}{\delta}} f(k\delta)k^{d-1}\delta^d \leq C \int_{r_0}^{+\infty} f(r) r^{d-1} dr$$

as required for stability.

Now we impose a third condition

$$\phi(\vec{x}) \leq \frac{A}{|\vec{x}|^{d+\varepsilon}} \qquad \text{for } |\vec{x}| \geq R_0$$

for some $R_0 > r_0$, $\varepsilon > 0$, $A \geq 0$. By Example 3.11, this condition implies that the potential

$$\mathfrak{U}(\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_N) = \sum_{i<j} \phi(\vec{x}_j - \vec{x}_i)$$

is tempered. Thus, for such pair potential $\phi$, the thermodynamic limit $g(\beta, \rho)$ exists.

Note that the hard core condition implies that the defining domain of $g(\beta, \rho)$ is

$$0 \leq \rho < \rho_{\max}$$

for some $\rho_{\max} < +\infty$ in this case.

# 4 The Ising Model

## Contents

The Ising model is one of the simplest interacting systems in equilibrium statistical mechanics. Let $\Lambda \subset \mathbb{Z}^d$ be a finite region. For each site $i \in \Lambda$, one places a spin

$$\sigma_i \in \{+1, -1\}.$$

The Hamiltonian is

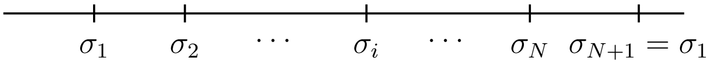
$$H_\Lambda(\sigma) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_{i \in \Lambda} \sigma_i$$

where $\langle i, j \rangle$ denote nearest-neighbor pairs in $\Lambda$. $J$ is the nearest-neighbor coupling. The interaction is ferromagnetic for $J > 0$ and antiferromagnetic for $J < 0$. The constant $h$ represents external magnetic field.

This model was proposed by Wilhelm Lenz in 1920 and studied by his student Ernst Ising. Ising proved that there is no phase transition in 1D at $T > 0$ and incorrectly conjectured that this holds in higher dimensions. In 1944, Lars Onsager solved the 2D Ising model (zero field) exactly which demonstrated the existence of a phase transition at critical temperature $T_c$. Nowadays the 2D square-lattice Ising model is one of the simplest statistical model to illustrate a phase transition and qualitative results for real physical systems.

## 4.1 *The one-dimensional Ising model*

We start with solving the 1D Ising model. Consider a ring of $N$ spins with periodic boundary condition $\sigma_{N+1} \equiv \sigma_1$ (the difference between boundary conditions here becomes negligible in the thermodynamic limit).



The Hamiltonian for this periodic spin chain is

$$H_N(\sigma) = -J\sum_{i=1}^N \sigma_i\sigma_{i+1} - h\sum_{i=1}^N \sigma_i$$

and the corresponding partition function reads

$$Z_N(\beta, h) = \sum_{\sigma_1,\dots,\sigma_N=\pm 1} \exp\left(\beta J\sum_{i=1}^N \sigma_i\sigma_{i+1} + \beta h\sum_{i=1}^N \sigma_i\right).$$

## 4.2 Transfer matrix

This partition function can be computed via the method of *transfer matrix*. Let us first rewrite

$$\sum_{i=1}^N \sigma_i = \frac{1}{2}\sum_{i=1}^N (\sigma_i + \sigma_{i+1})$$

so

$$Z_N(\beta, h) = \sum_{\sigma_1,\dots,\sigma_N} \prod_{i=1}^N \exp\left(\beta J\sigma_i\sigma_{i+1} + \frac{\beta h}{2}(\sigma_i + \sigma_{i+1})\right).$$

Define the transfer matrix $T$ indexed by $\sigma, \sigma' \in \{\pm 1\}$ by

$$T_{\sigma,\sigma'} = \exp\left(\beta J\sigma\sigma' + \frac{\beta}{2}h(\sigma + \sigma')\right).$$

In the ordered basis $\{+1, -1\}$,

$$T = \begin{pmatrix} e^{\beta J+\beta h} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J-\beta h} \end{pmatrix}.$$

The partition function can be written as a trace

$$Z_N = \mathrm{Tr}\; T^N.$$

The eigenvalues of the $2 \times 2$ matrix $T$ are

$$\lambda_\pm = e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}.$$

Therefore

$$Z_N = \lambda_+^N + \lambda_-^N$$

and

$$F_N = -\frac{1}{\beta}\log Z_N = -\frac{1}{\beta}\log\left(\lambda_+^N + \lambda_-^N\right).$$

The free energy per site is

$$f_N(\beta, h) = \frac{1}{N}F_N = -\frac{1}{\beta N}\log(\lambda_+^N + \lambda_-^N).$$

In the thermodynamic limit

$$f(\beta, h) = \lim_{N\to\infty} f_N(\beta, h) = -\frac{1}{\beta}\log \lambda_+.$$

Since $\lambda_+$ is analytic for every finite $\beta$ and real $h$, the free energy $f(\beta, h)$ is analytic. This implies that there is no phase transition at $T > 0$ for 1D Ising model.

The internal energy per site is

$$u_N = -\frac{1}{N}\frac{\partial}{\partial\beta}\log Z_N$$

and the specific heat is

$$c_N = \frac{\partial u_N}{\partial T} = k_B \frac{\beta^2}{N}\frac{\partial^2}{\partial\beta^2}\log Z_N.$$

In the thermodynamic limit

$$u = -\frac{\partial}{\partial\beta}\log\lambda_+, \qquad c = k_B\beta^2\frac{\partial^2}{\partial\beta^2}\log\lambda_+.$$

At $h = 0$, we have $\lambda_+ = 2\cosh\beta J$ and

$$\begin{aligned} u(\beta, 0) &= -J\tanh\beta J \\ c(\beta, 0) &= k_B(\beta J)^2 \mathrm{sech}^2\beta J \end{aligned}$$

Both are smooth in $\beta$ and there is no singularity.

## 4.3 Magnetization

The total magnetization of a spin configuration is

$$M(\sigma) = \sum_{i=1}^N \sigma_i$$

and its expectation value can be computed by

$$\langle M\rangle = \sum_{\sigma_1,\dots,\sigma_N} M(\sigma)\frac{e^{-\beta H_N(\sigma)}}{Z_N} = \frac{1}{\beta}\frac{\partial}{\partial h}\log Z_N = -\frac{\partial}{\partial h}F_N.$$

The magnetization density is

$$m_N(\beta, h) = \frac{1}{N}\langle M\rangle = -\frac{\partial}{\partial h}f_N = \langle\sigma_i\rangle, \qquad \text{for any } i.$$

In the thermodynamic limit

$$\begin{aligned} m(\beta, h) &= -\frac{\partial}{\partial h}f(\beta, h) = \frac{1}{\beta}\frac{\partial}{\partial h}\log\lambda_+ \\ &= \frac{1}{\beta}\frac{\partial}{\partial h}\log\left[e^{\beta J}\cosh(\beta h) + \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}}\right] \\ &= \frac{1}{\beta}\frac{\partial}{\partial h}\log\left[\cosh(\beta h) + \sqrt{\sinh^2(\beta h) + e^{-4\beta J}}\right] \\ &= \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}}. \end{aligned}$$

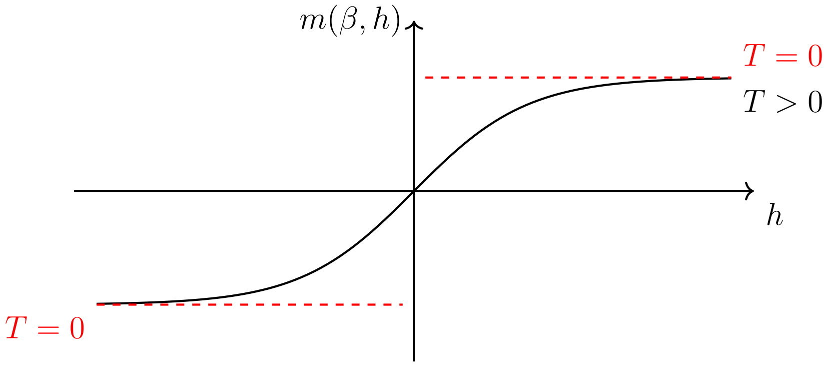
We analyze the spontaneous magnetization which is about the limit $\lim_{h\to 0^\pm}$. We assume $J > 0$ (ferromagnetism).

- At finite $N$, the Hamiltonian with zero field $H_N = -J\sum_{i=1}^{N}\sigma_i\sigma_{i+1}$ is invariant under the global spin-flip symmetry $\sigma_i \to -\sigma_i$. Thus

$$\lim_{h\to 0^\pm} m_N(\beta,h) = m_N(\beta,0) = 0.$$

- In the thermodynamic limit

$$\lim_{h\to 0^\pm} m(\beta,h) = \begin{cases} 0 & \text{if } \beta < \infty \\ \pm 1 & \text{if } \beta = \infty \text{ or } T = 0 \end{cases}$$



So at zero temperature $T = 0$, the system will exhibit spontaneous magnetization. This is expected since at $T = 0$, we have two ground states (in the case $J > 0$)

$$\uparrow\uparrow\uparrow\dots\uparrow \qquad \text{and} \qquad \downarrow\downarrow\dots\downarrow$$

If we turn on external field $h$ and gradually send it back to zero, the system will stay in the ground state that aligns with the direction of $h$. In this case,

$$\begin{array}{ccc} \lim_{h\to 0^\pm}\lim_{N\to\infty} m_N(+\infty,h) & \neq & \lim_{N\to\infty}\lim_{h\to 0^\pm} m_N(+\infty,h) \\ \| & & \| \\ \pm 1 & & 0 \end{array}$$

## 4.4 Correlation functions

The magnetization

$$m_N(\beta,h) = \langle\sigma_i\rangle$$

is about the 1-point correlation function. In general, we can consider the n-point function

$$\langle\sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_n}\rangle \qquad i_1 < i_2 < \cdots < i_n.$$

**2-point fucntion**

The most important is the two-point function

$$\langle \sigma_i \sigma_j \rangle$$

and its connected version

$$\langle \sigma_i \sigma_j \rangle_c := \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle.$$

- If $\langle \sigma_i \sigma_j \rangle_c = 0$, the fluctuations at different sites are uncorrelated.
- If $\langle \sigma_i \sigma_j \rangle_c$ decays slowly with $|i-j|$, the system has long-range order.
- If $\langle \sigma_i \sigma_j \rangle_c$ decays exponentially, correlations are short-ranged.

In the following, we show that the connected correlations in 1D Ising model decays exponentially at any finite temperature $T > 0$. By the periodic boundary condition and the translate invariance of the Hamiltonian function

$$\langle \sigma_i \sigma_j \rangle = \langle \sigma_0 \sigma_r \rangle \qquad \text{where } r = |j-i|.$$

By definition

$$\langle \sigma_0 \sigma_r \rangle = \frac{1}{Z_N} \sum_{\sigma_i} \sigma_0 \sigma_r e^{\beta J \sum\limits_{i=1}^{N} \sigma_i \sigma_{i+1} + \frac{\beta}{2} h \sum\limits_{i=1}^{N} (\sigma_i + \sigma_{i+1})} = \frac{1}{Z_N} \sum_{\sigma_i} \sigma_0 \sigma_r \prod_{i=1}^{N} T_{\sigma_i, \sigma_{i+1}}.$$

Introduce the $2 \times 2$ matrix

$$S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Then the above two-point function can be written as

$$\langle \sigma_0 \sigma_r \rangle = \frac{\mathrm{Tr}(S T^r S T^{N-r})}{\mathrm{Tr}\ T^N}.$$

Let $|\pm\rangle$ be the two normalized eigenvectors of $T$

$$T|\pm\rangle = \lambda_\pm |\pm\rangle$$

so we have the spectral decomposition

$$T = \lambda_+ |+\rangle\langle+| + \lambda_- |-\rangle\langle-|.$$

Let us compute $\langle \sigma_0 \sigma_r \rangle$ in the thermodynamic limit $N \to +\infty$. In $N \to +\infty$, $T^{N-r}$ and $T^N$ are dominated by

$$T^{N-r} \sim \lambda_+^{N-r} |+\rangle\langle+|, \quad T^N \sim \lambda_+^N |+\rangle\langle+|.$$

Thus

$$\begin{aligned}\lim_{N\to+\infty}\langle\sigma_0\sigma_r\rangle &= \lim_{N\to+\infty}\frac{\lambda_+^{N-r}\langle+|ST^rS|+\rangle}{\lambda_+^N\langle+|+\rangle} = \frac{1}{\lambda_+^r}\langle+|ST^rS|+\rangle\\ &= \frac{1}{\lambda_+^r}\left(\langle+|ST^r|+\rangle\langle+|S|+\rangle + \langle+|ST^r|-\rangle\langle-|S|+\rangle\right)\\ &= \langle+|S|+\rangle^2 + \left(\frac{\lambda_-}{\lambda_+}\right)^r\langle+|S|-\rangle\langle-|S|+\rangle\\ &= \langle+|S|+\rangle^2 + \left(\frac{\lambda_-}{\lambda_+}\right)^r|\langle+|S|-\rangle|^2.\end{aligned}$$

Note that the one-point function is

$$m = \lim_{N\to\infty}\langle\sigma_0\rangle = \lim_{N\to\infty}\frac{\mathrm{Tr}(ST^N)}{\mathrm{Tr}\ T^N} = \langle+|S|+\rangle$$

so

$$\lim_{N\to\infty}\langle\sigma_0\sigma_r\rangle = m^2 + \left(\frac{\lambda_-}{\lambda_+}\right)^r|\langle+|S|-\rangle|^2, \quad \lim_{N\to\infty}\langle\sigma_0\sigma_r\rangle_c = \left(\frac{\lambda_-}{\lambda_+}\right)^r|\langle+|S|-\rangle|^2.$$

Using $S^2 = I$, we have

$$1 = \langle+|+\rangle = \langle+|S^2|+\rangle = \langle+|S|+\rangle\langle+|S|+\rangle + \langle+|S|-\rangle\langle-|S|+\rangle = m^2 + |\langle+|S|-\rangle|^2$$

$$\implies \quad |\langle+|S|-\rangle|^2 = 1 - m^2.$$

Thus we find

$$\lim_{N\to\infty}\langle\sigma_0\sigma_r\rangle_c = (1-m^2)\left(\frac{\lambda_-}{\lambda_+}\right)^r$$

which decays exponentially as $r \to +\infty$. At zero field $h = 0$ and $T > 0$, we have

$$\lim_{N\to\infty}\langle\sigma_0\sigma_r\rangle_c\Big|_{h=0} = (\tanh\beta J)^r.$$

The correlation length $\xi$ is defined by the decay property

$$\langle\sigma_0\sigma_r\rangle \sim e^{-r/\xi}.$$

The above result shows

$$\xi = \frac{1}{\log\lambda_+/\lambda_-}.$$

At zero field $h = 0$,

$$\xi = -\frac{1}{\log\tanh\beta J} = \frac{1}{\log\coth\beta J}$$

which is finite for any $T > 0$, and diverges as $T \to 0$.

**n-point function**

Now we move on to compute n-point function

$$\langle \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_n}\rangle \qquad i_1 < i_2 < \dots < i_n.$$

By the same argument as before via the transfer matrix

$$\langle \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_n}\rangle = \frac{\mathrm{Tr}(T^{i_1}ST^{i_2-i_1}S\dots ST^{N-i_n})}{\mathrm{Tr}\ T^N}.$$

For simplicity, let us consider the case at zero field $h=0$. In this case

$$T = \begin{pmatrix} e^{\beta J} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta J} \end{pmatrix}$$

and the normalized eigenvectors are

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

with eigenvalues

$$\lambda_+ = 2\cosh(\beta J), \quad \lambda_- = 2\sinh(\beta J).$$

The $S$-operator swaps two eigenvectors

$$S|+\rangle = |-\rangle, \quad S|-\rangle = |+\rangle.$$

It follows that the $n$-point function is zero unless $n=2m$ even. This can be also seen by the reflection symmetry $\sigma_i \to -\sigma_i$ at zero field $h=0$.

Assume $n=2m$ is even. In the thermodynamic limit, we have

$$\begin{aligned}
\lim_{N\to\infty}\langle \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_{2m}}\rangle &= \langle +|T^{i_1}ST^{i_2-i_1}S\dots T^{-i_{2m}}|+\rangle \\
&= \lambda_+^{i_1}\lambda_-^{i_2-i_1}\lambda_+^{i_3-i_2}\dots\lambda_-^{i_{2m}-i_{2m-1}}\lambda_+^{-i_{2m}} \\
&= \left(\frac{\lambda_-}{\lambda_+}\right)^{i_2-i_1}\left(\frac{\lambda_-}{\lambda_+}\right)^{i_4-i_3}\dots\left(\frac{\lambda_-}{\lambda_+}\right)^{i_{2m}-i_{2m-1}} \\
&= \prod_{j=1}^{m}\left(\frac{\lambda_-}{\lambda_+}\right)^{i_{2j}-i_{2j-1}}
\end{aligned}$$

Note that the $2m$-point function factors

$$\lim_{N\to\infty}\langle \sigma_{i_1}\sigma_{i_2}\dots\sigma_{i_{2m}}\rangle = \lim_{N\to\infty}\langle \sigma_{i_1}\sigma_{i_2}\rangle\langle \sigma_{i_3}\sigma_{i_4}\rangle\dots\langle \sigma_{i_{2m-1}}\sigma_{i_{2m}}\rangle$$

which does not obey the Wick rule on summing over all pairings. For example, we have

$$\lim_{N\to\infty}\langle \sigma_i\sigma_j\sigma_k\sigma_l\rangle = \lim_{N\to\infty}\langle \sigma_i\sigma_j\rangle\langle \sigma_k\sigma_l\rangle, \qquad i<j<k<l.$$

This formula says that in 1D Ising model the information carried by a many-spin product is transmitted along the chain in a completely ordered way. This reflects the fact that the 1D Ising model has only nearest-neighbor coupling with only one spatial direction and no loops.

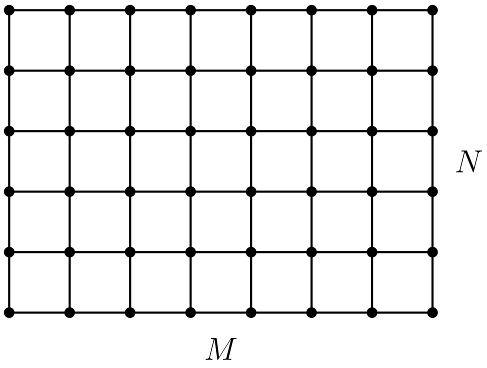
## 4.5 *The two-dimensional Ising model*

We move on to discuss 2D Ising model on the square-lattice. We focus on the zero-field case ($h = 0$), which is classically exactly solvable.

Let $\Lambda_{M,N} \subset \mathbb{Z}^2$ be an $M \times N$ rectangular region. We put periodic boundary conditions in both directions. A spin configuration on $\Lambda_{M,N}$ is described by

$$\{\sigma_{m,n}\} \qquad 1 \leq m \leq M, \quad 1 \leq n \leq N$$

and we will identify $\sigma_{M+1,n} \equiv \sigma_{1,n}$ and $\sigma_{m,N+1} \equiv \sigma_{m,1}$ by the periodic boundary condition. This can be also viewed as Ising model on the torus.



We consider the anisotropic nearest-neighbor Hamiltonian

$$H(\sigma) = -J_h \sum_{m=1}^{M} \sum_{n=1}^{N} \sigma_{m,n}\sigma_{m,n+1} - J_v \sum_{m=1}^{M} \sum_{n=1}^{N} \sigma_{m,n}\sigma_{m+1,n}$$

where $J_h$ represents the horizontal coupling and $J_v$ represents the vertical coupling. The standard isotropic model is the special case $J_h = J_v = J$. The partition function is

$$Z_{M,N}(K_h, K_v) = \sum_{\sigma} \exp\left( K_h \sum_{m,n} \sigma_{m,n}\sigma_{m,n+1} + K_v \sum_{m,n} \sigma_{m,n}\sigma_{m+1,n} \right)$$

where we write

$$K_h = \beta J_h, \quad K_v = \beta J_v.$$

The thermodynamic free energy density is

$$f(K_h, K_v) = -\frac{1}{\beta} \lim_{M,N\to\infty} \frac{1}{MN} \log Z_{M,N}(K_h, K_v).$$

Remarkably, Onsager [36] found the following exact formula for this free energy!

$$-\beta f = \log 2 + \frac{1}{2\pi^2} \int_0^{\pi} \int_0^{\pi} \log\Big(\cosh 2K_h \cosh 2K_v - \sinh 2K_h \cos\theta_1 - \sinh 2K_v \cos\theta_2\Big) d\theta_1 d\theta_2.$$

## 4.6 Kramers-Wannier duality

Before deriving Onsager's solution, we discuss a duality relation (Kramers-Wannier duality) for square-lattice Ising model that will directly locate the critical temperature.

### High temperature expansion

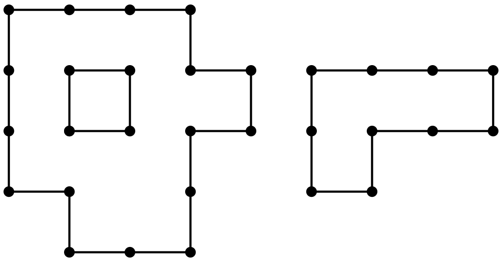
For two spin variables $\sigma, \sigma' \in \{\pm 1\}$, we have

$$e^{K\sigma\sigma'} = \cosh K + \sigma\sigma' \sinh K = \cosh K(1 + \sigma\sigma' \tanh K).$$

Therefore

$$Z_{M,N} = (\cosh K_h \cosh K_v)^{MN} \sum_{\sigma} \prod_{m,n} (1 + \sigma_{m,n}\sigma_{m,n+1} \tanh K_h) \prod_{m,n} (1 + \sigma_{m,n}\sigma_{m+1,n} \tanh K_v).$$

When the products are expanded, a monomial in $\sigma$'s survives the spin sum if and only if each spin appears with even degree. Such a spin configuration forms a closed bond configuration $\Gamma$.



$$\Gamma: \quad |\Gamma_h| = 16, \quad |\Gamma_v| = 14.$$

Thus the partition function can be written as

$$Z_{M,N} = (2 \cosh K_h \cosh K_v)^{MN} \sum_{\Gamma} (\tanh K_h)^{|\Gamma_h|} (\tanh K_v)^{|\Gamma_v|}$$

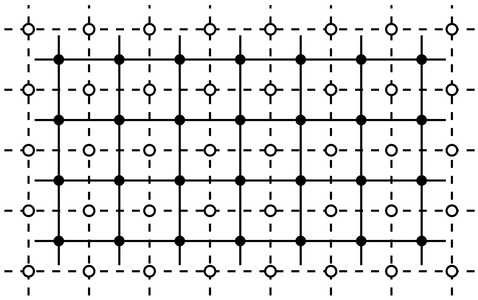
where the summation is over all closed bond configurations and

$$|\Gamma_h| = \#\{\text{horizontal bonds}\}, \quad |\Gamma_v| = \#\{\text{vertical bonds}\}.$$

Note that

$$\tanh K_h = \tanh \beta J_h, \quad \tanh K_v = \tanh \beta J_v$$

which are small when $\beta \to 0$ (i.e. $T \to +\infty$). Thus this formula can be viewed as the high temperature expansion.



black lattice $\overset{\text{dual}}{\longleftrightarrow}$ white lattice

**Low temperature expansion**

The partition function $Z_{M,N}$ is formulated in a rectangular region $\Lambda_{M,N}$ in the lattice $\mathbb{Z}^2$. We can consider the dual lattice by placing one dual vertex at the center of each face in $\mathbb{Z}^2$, and drawing a dual edge crossing each primal edge in $\mathbb{Z}^2$.

At low temperature, $K_h$ and $K_v$ are large, and the dominant configurations are the two ordered ground states: all spin $+1$ and all spin $-1$.

$$\uparrow\uparrow\uparrow\dots\uparrow \qquad \text{and} \qquad \downarrow\downarrow\dots\downarrow$$

The corresponding contribution to the partition function is

$$e^{(K_h+K_v)MN}.$$

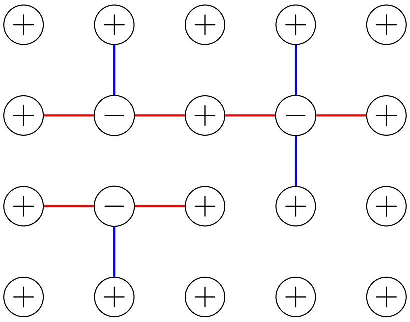
For each spin configuration $\sigma$ on $\Lambda_{M,N}$, let

$$D(\sigma)=\{\text{nearest-neighbor } (x,y)\in\Lambda_{M,N} \mid \sigma_x\neq\sigma_y\}$$

which will be called the set of disagreeing bonds. Let

$$D(\sigma)_h\subset D(\sigma), \quad D(\sigma)_v\subset D(\sigma)$$

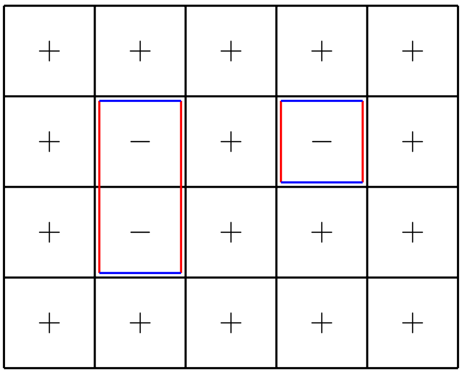
denote the subset of horizontal and vertical disagreeing bonds respectively.



$D(\sigma)_h=$ red $\qquad D(\sigma)_v=$ blue

We can draw $D(\sigma)$ in the dual lattice. Then $D(\sigma)$ becomes a closed bond configuration $\Gamma^*(\sigma)$ in the dual lattice such that

$$D(\sigma)_h \Leftrightarrow \Gamma^*(\sigma)_v, \qquad D(\sigma)_v \Leftrightarrow \Gamma^*(\sigma)_h.$$



$D(\sigma)$ in dual lattice

Comparing with the ground state, each horizontal disagreeing bond will increase the Hamiltonian by $2J_h$ and each vertical disagreeing bond will increase the Hamiltonian by $2J_v$. Thus a spin configuration $\sigma$ contributes to the partition function by

$$e^{(K_h+K_v)MN}e^{-2K_h|D(\sigma)_h|-2K_v|D(\sigma)_v|}.$$

The summation over all spin configuration $\sigma$ becomes the summation over all closed bond configurations $\Gamma^*(\sigma)$ in the dual lattice. Thus the partition function can be rewritten dually as

$$Z_{M,N} = e^{(K_h+K_v)MN}\sum_{\Gamma^*} e^{-2K_v|\Gamma^*_h|}e^{-2K_h|\Gamma^*_v|}.$$

Note that when $T$ is small, $K_h, K_v$ are large and so $e^{-2K_v}$ and $e^{-2K_h}$ are small. Thus this formula can be viewed as the low temperature expansion.

For convenience, let us denote

$$Z^{\mathrm{red}}_{M,N}(K_h, K_v) := \sum_{\Gamma}(\tanh K_h)^{|\Gamma_h|}(\tanh K_v)^{|\Gamma_v|}$$

and

$$\psi(K_h, K_v) = -\frac{1}{\beta}\lim_{M,N\to\infty}\frac{1}{MN}\log Z^{\mathrm{red}}_{M,N}(K_h, K_v).$$

We have

$$Z_{M,N} = (2\cosh K_h \cosh K_v)^{MN} Z^{\mathrm{red}}_{M,N}$$

and

$$-\beta f(K_h, K_v) = \log(2\cosh K_h \cosh K_v) - \beta\psi(K_h, K_v).$$

By the dual description in the low temperature expansion

$$-\beta f(K_h, K_v) = K_h + K_v - \beta\psi(K_v^*, K_h^*)$$

where the dual coupling $K_h^*, K_v^*$ are defined by

$$\tanh K_h^* = e^{-2K_h}, \quad \tanh K_v^* = e^{-2K_v}$$

which is equivalent to the symmetric form

$$\sinh(2K_h^*)\sinh(2K_h) = 1, \quad \sinh(2K_v^*)\sinh(2K_v) = 1.$$

This is the Kramers-Wannier duality relation. Under this duality, we have

$$\log(2\cosh K_h \cosh K_v) - \beta\psi(K_h, K_v) = K_h + K_v - \beta\psi(K_v^*, K_h^*)$$

The duality map sends

- high temperature $K \ll 1$ to low temperature $K^* \gg 1$
- low temperature $K \gg 1$ to high temperature $K^* \ll 1$.

If the model has a unique phase transition (we will see this by Onsager's formula), then the critical point must be fixed by this duality

$$(K_h, K_v) = (K_v^*, K_h^*).$$

Thus at the critical temperature $K_h, K_v$ will satisfy the relation

$$\sinh(2K_h)\sinh(2K_v) = 1.$$

In the isotropic case $J_h = J_v = J$, the critical temperature $T_c$ satisfies

$$\sinh\left(\frac{2J}{k_B T_c}\right) = 1$$

from which we find $T_c$ for the phase transition

$$T_c = \frac{2J}{k_B \log(1+\sqrt{2})}.$$

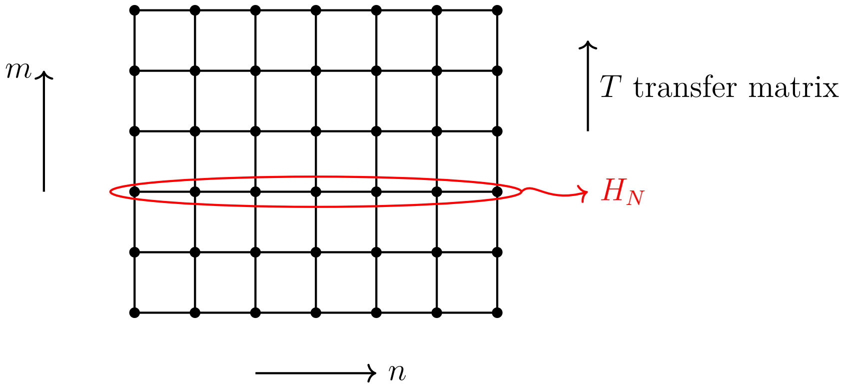
## 4.7 Onsager's exact free energy

### Transfer matrix

The free energy per site $f(K_h, K_v)$ can be computed exactly via the method of transfer matrix. The idea is to group spin configurations of each row by the Hilbert space

$$H_N = (\mathbb{C}^2)^{\otimes N}$$

and turn the partition as a trace of operators on $H_N$ in the vertical direction.



Let $\mu$ denote a row spin configuration

$$\mu = (\sigma_1, \sigma_2, \cdots, \sigma_N) \qquad \sigma_i \in \{\pm 1\}.$$

Such $\mu$ gives an orthonormal basis $\{|\mu\rangle\}$ of $H_N$. Consider the operator

$$T : H_N \to H_N$$

whose matrix entry is given by

$$\langle \mu | T | \mu' \rangle = \exp\left( K_h \sum_{j=1}^{N} \sigma_j \sigma_{j+1} + K_v \sum_{j=1}^{N} \sigma_j \sigma'_j \right).$$

Then the partition function can be written as

$$Z_{M,N} = \sum_{\sigma} \exp\left( K_h \sum_{m,n} \sigma_{m,n}\sigma_{m,n+1} + K_v \sum_{m,n} \sigma_{m,n}\sigma_{m+1,n} \right) = \mathrm{Tr}(T^M).$$

The problem is to find the largest eigenvalue of $T$ as $M, N \to \infty$.

We can further decompose $T$ as

$$T = T_h T_v$$

where $T_h$ is diagonal

$$\langle \mu | T_h | \mu' \rangle = \delta_{\mu,\mu'} \exp\left( K_h \sum_{j=1}^{N} \sigma_j \sigma_{j+1} \right)$$

and $T_v$ has the matrix entry

$$\langle\mu|T_v|\mu'\rangle = \exp\left(K_v\sum_{j=1}^N \sigma_j\sigma'_j\right).$$

Thus

$$\langle\mu|T_hT_v|\mu'\rangle = \langle\mu|T_h|\mu\rangle\langle\mu|T_v|\mu'\rangle = \langle\mu|T|\mu'\rangle$$

as expected.

### Spinor formalism

We use Kaufman's spinor formalism [23] to compute eigenvalues of the transfer matrix. Introduce the Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

We extend them to linear operators

$$\sigma_x^{(j)}, \sigma_y^{(j)}, \sigma_z^{(j)} : H_N \to H_N$$

which act on the $j$-th component $\mathbb{C}^2$ of $H_N = (\mathbb{C}^2)^{\otimes N}$ by $\sigma_x, \sigma_y, \sigma_z$ and on the other component as identity. We can express $T_h$ and $T_v$ in terms of the Pauli matrices. For $T_h$, we simply have

$$T_h = \exp\left(K_h\sum_{j=1}^N \sigma_z^{(j)}\sigma_z^{(j+1)}\right).$$

For $T_v$, its matrix entry factorizes

$$\langle\mu|T_v|\mu'\rangle = \prod_{j=1}^N e^{K_v\sigma_j\sigma'_j}$$

So

$$T_v = A_v^{\otimes N}$$

where $A_v$ is the $2\times 2$ matrix

$$A_v = \begin{pmatrix} e^{K_v} & e^{-K_v} \\ e^{-K_v} & e^{K_v} \end{pmatrix}.$$

Since $\det A_v = 2\sinh(2K_v)$, we have

$$\frac{1}{\sqrt{2\sinh(2K_v)}}A_v = \exp(K_v^*\sigma_x) = \begin{pmatrix} \cosh K_v^* & \sinh K_v^* \\ \sinh K_v^* & \cosh K_v^* \end{pmatrix}$$

for some $K_v^*$. Computing the ratio of the first row, we find

$$\tanh K_v^* = e^{-2K_v}.$$

So $K_v^*$ is precisely the dual coupling of $K_v$. Thus

$$T_v = (2\sinh(2K_v))^{N/2} \exp\left(K_v^* \sum_{j=1}^N \sigma_x^{(j)}\right)$$

and the full transfer matrix $T$ becomes

$$T = (2\sinh(2K_v))^{N/2} \exp\left(K_h \sum_{j=1}^N \sigma_z^{(j)}\sigma_z^{(j+1)}\right) \exp\left(K_v^* \sum_{j=1}^N \sigma_x^{(j)}\right).$$

Define $2N$ operators on $H_N$ (Jordan–Wigner transformation)

$$\Gamma_{2j-1} = \left(\prod_{i<j} \sigma_x^{(i)}\right)\sigma_z^{(j)} \quad , \quad \Gamma_{2j} = \left(\prod_{i<j} \sigma_x^{(i)}\right)\sigma_y^{(j)}.$$

They satisfy the Clifford relations

$$\Gamma_a\Gamma_b + \Gamma_b\Gamma_a = 2\delta_{ab}.$$

Observe

$$\sigma_z^{(j)}\sigma_z^{(j+1)} = i\Gamma_{2j}\Gamma_{2j+1}, \quad \sigma_x^{(j)} = i\Gamma_{2j-1}\Gamma_{2j}.$$

Thus the transfer matrix $T$ can be written as (we identify $\Gamma_{2N+1} = \Gamma_1$)

$$T = (2\sinh(2K_v))^{N/2} \exp\left(iK_h \sum_{j=1}^N \Gamma_{2j}\Gamma_{2j+1}\right) \exp\left(iK_v^* \sum_{j=1}^N \Gamma_{2j-1}\Gamma_{2j}\right).$$

The upshot is that $T$ is expressed as exp of quadratic expressions in $\Gamma$'s, i.e., elements of $\mathrm{Spin}(2N)$. By the Baker-Campbell-Hausdorff formula, we can write

$$\exp\left(iK_h \sum_{j=1}^N \Gamma_{2j}\Gamma_{2j+1}\right) \exp\left(iK_v^* \sum_{j=1}^N \Gamma_{2j-1}\Gamma_{2j}\right) = \exp\left(\frac{1}{4}\sum_{a,b=1}^{2N} A_{ab}\Gamma_a\Gamma_b\right)$$

for some skew-symmetric matrix $A$. To compute $A$, let us denote

$$\Phi_h := \exp\left(iK_h \sum_{j=1}^N \Gamma_{2j}\Gamma_{2j+1}\right), \quad \Phi_v := \exp\left(iK_v^* \sum_{j=1}^N \Gamma_{2j-1}\Gamma_{2j}\right).$$

and

$$\Phi = \Phi_h \cdot \Phi_v.$$

**Lemma 4.1.** *Let* $C = (C_{ab})$ *be a skew-symmetric* $2N \times 2N$ *matrix. Let* $\mathcal{U} = \exp\left(\frac{1}{4}\sum\limits_{a,b=1}^{2N} C_{ab}\Gamma_a\Gamma_b\right)$ *and* $X = e^{-C} = (X_{ab})$*. Then*

$$\mathcal{U}\Gamma_a\mathcal{U}^{-1} = \sum_b X_{ab}\Gamma_b.$$

*Proof:* This is essentially the map $\mathrm{Spin}(2N) \to SO(2N)$. Let $Y = \frac{1}{4}\sum_{a,b} C_{ab}\Gamma_a\Gamma_b$ so $\mathcal{U} = e^Y$. By the Clifford relations and the matrix identity

$$[A_1A_2, A_3] = A_1\{A_2, A_3\} - \{A_1, A_3\}A_2, \quad \text{where } \{A_1, A_2\} = A_1A_2 + A_2A_1,$$

we find

$$[Y, \Gamma_a] = -\sum_b C_{ab}\Gamma_b.$$

It follows that

$$\mathcal{U}\Gamma_a\mathcal{U}^{-1} = e^Y\Gamma_a e^{-Y} = e^{\mathrm{ad}_Y}\Gamma_a = \sum_b X_{ab}\Gamma_b.$$

□

Our goal is to compute eigenvalues of $T$ or

$$\Phi = \exp\left(\frac{1}{4}\sum_{a,b} A_{ab}\Gamma_a\Gamma_b\right)$$

on the Hilbert space $H_N$, which is the standard spin representation.

**Proposition 4.2.** *Let* $\pm\lambda_1, \pm\lambda_2, \cdots, \pm\lambda_N$ *be eigenvalues of* $A$*. Then eigenvalues of* $\Phi$ *are*

$$\exp\left(\pm\frac{1}{2}\lambda_1 \pm \frac{1}{2}\lambda_2 \pm \cdots \pm \frac{1}{2}\lambda_N\right).$$

*Proof:* Since the matrix $A$ is skew symmetric, we can find $O \in O(2N)$ such that

$$A = O^{-1}\begin{pmatrix} 0 & \omega_1 & & & & & \\ -\omega_1 & 0 & & & & & \\ & & 0 & \omega_2 & & & \\ & & -\omega_2 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \omega_N \\ & & & & & -\omega_N & 0 \end{pmatrix} O$$

Eigenvalues of $A$ are $\pm i\omega_1, \pm i\omega_2, \ldots, \pm i\omega_N$.

Let

$$\tilde{\Gamma}_a = \sum_b O_{ab}\Gamma_b.$$

Since $O \in O(2N)$, $\{\tilde{\Gamma}_a\}$ still satisfy the Clifford relations

$$\tilde{\Gamma}_a\tilde{\Gamma}_b + \tilde{\Gamma}_b\tilde{\Gamma}_a = 2\delta_{ab}$$

and $\Phi$ can be rewritten as

$$\Phi = \exp\left(\frac{1}{2}\sum_{j=1}^{N}\omega_j\tilde{\Gamma}_{2j-1}\tilde{\Gamma}_{2j}\right).$$

Define the fermionic ladder operators

$$c_j^+ = \frac{1}{2}(\tilde{\Gamma}_{2j-1} + i\tilde{\Gamma}_{2j}), \quad c_j = \frac{1}{2}(\tilde{\Gamma}_{2j-1} - i\tilde{\Gamma}_{2j})$$

which satisfy the relations

$$(c_j^+)^2 = c_j^2 = 0, \qquad c_j^+c_j + c_jc_j^+ = 1.$$

Let $|0\rangle \in H_N$ denote the vacuum state satisfying

$$c_j|0\rangle = 0 \qquad \forall j = 1, \cdots, N.$$

Then the full Hilbert $H_N$ has a basis

$$|\alpha_1\alpha_2\ldots\alpha_N\rangle := (c_1^+)^{\alpha_1}(c_2^+)^{\alpha_2}\ldots(c_N^+)^{\alpha_N}|0\rangle, \qquad \alpha_1, \alpha_2, \cdots, \alpha_N \in \{0, 1\}.$$

Since

$$\frac{1}{2}\tilde{\Gamma}_{2j-1}\tilde{\Gamma}_{2j} = i\left(c_j^+c_j - \frac{1}{2}\right) \implies \Phi = \exp\left(i\sum_{j=1}^{N}\omega_j\left(c_j^+c_j - \frac{1}{2}\right)\right).$$

Note that

$$c_j^+c_j - \frac{1}{2}: \quad \begin{aligned} |0\rangle &\longmapsto -\frac{1}{2}|0\rangle \\ c_j^+|0\rangle &\longmapsto \frac{1}{2}c_j^+|0\rangle \end{aligned}$$

It follows that $|\alpha_1\alpha_2\ldots\alpha_N\rangle$ are all eigenvectors of $\Phi$ with eigenvalues

$$\prod_{j=1}^{N} e^{i\omega_j\left(\alpha_j - \frac{1}{2}\right)}.$$

The proposition follows. □

By Lemma 4.1 and Proposition 4.2, we need to compute eigenvalues of the adjoint action

$$\mathrm{Ad}_\Phi : V \longrightarrow V$$

where

$$V = \mathrm{Span}\{\Gamma_1, \dots, \Gamma_{2N}\} \simeq \mathbb{C}^{2N} \quad \text{and} \quad \mathrm{Ad}_\Phi(\Gamma_a) = \Phi \Gamma_a \Phi^{-1}.$$

Since $\Phi = \Phi_h \cdot \Phi_v$, we can decompose

$$\mathrm{Ad}_\Phi = \mathrm{Ad}_{\Phi_h} \circ \mathrm{Ad}_{\Phi_v}.$$

Recall

$$\Phi_h := \exp\left(iK_h \sum_{j=1}^N \Gamma_{2j}\Gamma_{2j+1}\right), \quad \Phi_v := \exp\left(iK_v^* \sum_{j=1}^N \Gamma_{2j-1}\Gamma_{2j}\right).$$

$\mathrm{Ad}_{\Phi_h}$ preserves $\mathrm{Span}\{\Gamma_{2j}, \Gamma_{2j+1}\}$ and $\mathrm{Ad}_{\Phi_v}$ preserves $\mathrm{Span}\{\Gamma_{2j-1}, \Gamma_{2j}\}$. By Lemma 4.1,

$$\mathrm{Ad}_{\Phi_h}\begin{pmatrix}\Gamma_{2j}\\ \Gamma_{2j+1}\end{pmatrix} = \begin{pmatrix}\cosh 2K_h & -i\sinh 2K_h\\ i\sinh 2K_h & \cosh 2K_h\end{pmatrix}\begin{pmatrix}\Gamma_{2j}\\ \Gamma_{2j+1}\end{pmatrix}$$

$$\mathrm{Ad}_{\Phi_v}\begin{pmatrix}\Gamma_{2j-1}\\ \Gamma_{2j}\end{pmatrix} = \begin{pmatrix}\cosh 2K_v^* & -i\sinh 2K_v^*\\ i\sinh 2K_v^* & \cosh 2K_v^*\end{pmatrix}\begin{pmatrix}\Gamma_{2j-1}\\ \Gamma_{2j}\end{pmatrix}$$

To diagonalize $\mathrm{Ad}_\Phi$, consider the discrete Fourier transform and define vectors $\alpha_q, \beta_q$ by

$$\Gamma_{2j-1} = \frac{1}{\sqrt{N}}\sum_{q=1}^N e^{\frac{2\pi i}{N}qj}\alpha_q, \quad \Gamma_{2j} = \frac{1}{\sqrt{N}}\sum_{q=1}^N e^{\frac{2\pi i}{N}qj}\beta_q, \qquad q = 1, 2, \cdots, N.$$

On Fourier modes, we have

$$\mathrm{Ad}_{\Phi_v}\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix} = \begin{pmatrix}\cosh 2K_v^* & -i\sinh 2K_v^*\\ i\sinh 2K_v^* & \cosh 2K_v^*\end{pmatrix}\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix}$$

$$\mathrm{Ad}_{\Phi_h}\begin{pmatrix}\beta_q\\ e^{\frac{2\pi i q}{N}}\alpha_q\end{pmatrix} = \begin{pmatrix}\cosh 2K_h & -i\sinh 2K_h\\ i\sinh 2K_h & \cosh 2K_h\end{pmatrix}\begin{pmatrix}\beta_q\\ e^{\frac{2\pi i q}{N}}\alpha_q\end{pmatrix}$$

i.e.

$$\mathrm{Ad}_{\Phi_h}\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix} = \begin{pmatrix}\cosh 2K_h & ie^{-2\pi iq/N}\sinh 2K_h\\ -ie^{2\pi iq/N}\sinh 2K_h & \cosh 2K_h\end{pmatrix}\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix}$$

Therefore $\mathrm{Ad}_\Phi$ preserves $\mathrm{Span}\{\alpha_q, \beta_q\}$ and

$$\mathrm{Ad}_\Phi\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix} = \underbrace{\begin{pmatrix}\cosh 2K_v^* & -i\sinh 2K_v^*\\ i\sinh 2K_v^* & \cosh 2K_v^*\end{pmatrix}\begin{pmatrix}\cosh 2K_h & ie^{-2\pi iq/N}\sinh 2K_h\\ -ie^{2\pi iq/N}\sinh 2K_h & \cosh 2K_h\end{pmatrix}}_{=:B_q}\begin{pmatrix}\alpha_q\\ \beta_q\end{pmatrix}.$$

Let $e^{\pm\gamma(q)}$ denote eigenvalues of $B_q$. Then

$$e^{\gamma(q)} + e^{-\gamma(q)} = \text{Tr } B_q$$

which implies

$$\cosh\gamma(q) = \cosh 2K_h \cosh 2K_v^* - \sinh 2K_h \sinh 2K_v^* \cos\frac{2\pi q}{N}.$$

The set $\{\pm\gamma(q)\}_{q=1}^N$ give all eigenvalues of $\mathrm{Ad}_\Phi$.

**Computation of the free energy**

The partition function is

$$Z_{M,N} = \mathrm{Tr}(T^M) = \text{Tr }\left((2\sinh(2K_v))^{MN/2}\Phi^M\right).$$

The largest eigenvalue of $\Phi$ is

$$\lambda_{\max} = \prod_{q=1}^N e^{\frac{1}{2}\gamma(q)}.$$

Thus in the thermodynamic limit

$$-\beta f = \lim_{M,N\to\infty}\frac{1}{MN}\log Z_{M,N} = \frac{1}{2}\log(2\sinh 2K_v) + \lim_{N\to\infty}\frac{1}{N}\log\lambda_{\max}.$$

The second term becomes an integral

$$\lim_{N\to\infty}\frac{1}{N}\sum_{q=1}^N\frac{1}{2}\gamma(q) = \frac{1}{4\pi}\int_0^{2\pi}\gamma(\theta)d\theta = \frac{1}{2\pi}\int_0^{\pi}\gamma(\theta)d\theta$$

where

$$\cosh\gamma(\theta) = \cosh 2K_h\cosh 2K_v^* - \sinh 2K_h \sinh 2K_v^*\cos\theta.$$

Thus we obtain the exact formula

$$-\beta f = \frac{1}{2}\log(2\sinh 2K_v) + \frac{1}{2\pi}\int_0^\pi \cosh^{-1}\left(\cosh 2K_h\cosh 2K_v^* - \sinh 2K_h\sinh 2K_v^*\cos\theta\right)d\theta.$$

We can rewrite this formula via the original coupling $K_v$ in a symmetric form. The dual $K_v^*$ is given by

$$\sinh(2K_v^*) = \frac{1}{\sinh(2K_v)}$$

and we have

$$\cosh(2K_v^*) = \frac{\cosh(2K_v)}{\sinh(2K_v)}.$$

To simplify notations, let us denote

$$C_h = \cosh 2K_h, \quad S_h = \sinh 2K_h$$

$$C_v = \cosh 2K_v, \quad S_v = \sinh 2K_v$$

Then

$$\cosh\gamma(\theta) = \frac{C_h C_v - S_h \cos\theta}{S_v}.$$

For $x > 1$, we have the integral formula

$$\cosh^{-1}(x) = \frac{1}{\pi}\int_0^\pi \log(2(x - \cos\tilde{\theta}))d\tilde{\theta}$$

Therefore

$$\gamma(\theta) = \frac{1}{\pi}\int_0^\pi \log\left(2\left(\frac{C_h C_v - S_h\cos\theta}{S_v} - \cos\tilde{\theta}\right)\right)d\tilde{\theta}$$

and

$$\begin{aligned}-\beta f &= \frac{1}{2}\log(2S_v) + \frac{1}{2\pi^2}\int_0^\pi\int_0^\pi \log\left[2\left(\frac{C_h C_v - S_h\cos\theta}{S_v} - \cos\tilde{\theta}\right)\right]d\theta d\tilde{\theta}\\ &= \log 2 + \frac{1}{2\pi^2}\int_0^\pi\int_0^\pi \log(C_h C_v - S_h\cos\theta - S_v\cos\tilde{\theta})d\theta d\tilde{\theta}\end{aligned}$$

which is now symmetric in $K_h \leftrightarrow K_v$. Thus we have arrived at Onsager's exact solution of free energy for 2D Ising model at zero field

$$-\beta f = \log 2 + \frac{1}{2\pi^2}\int_0^\pi\int_0^\pi \log\Big(\cosh 2K_h\cosh 2K_v - \sinh 2K_h\cos\theta_1 - \sinh 2K_v\cos\theta_2\Big)d\theta_1 d\theta_2.$$

A phase transition in the thermodynamic limit shows up as a non-analyticity of the free energy or its derivatives. The singularity occurs when the minimum of the log argument vanishes. The minimum occurs at $\theta_1 = \theta_2 = 0$ and note that

$$(\cosh 2K_h\cosh 2K_v)^2 - (\sinh 2K_h + \sinh 2K_v)^2 = (\sinh 2K_h\sinh 2K_v - 1)^2 \geq 0.$$

Thus the non-analyticity shows up when

$$\sinh 2K_h\sinh 2K_v = 1.$$

This is precisely the fixed point of Kramers-Wannier duality.

### Internal energy and specific heat

Let us compute the internal energy at zero field. We consider the isotropic case

$$K_h = K_v = K = \beta J.$$

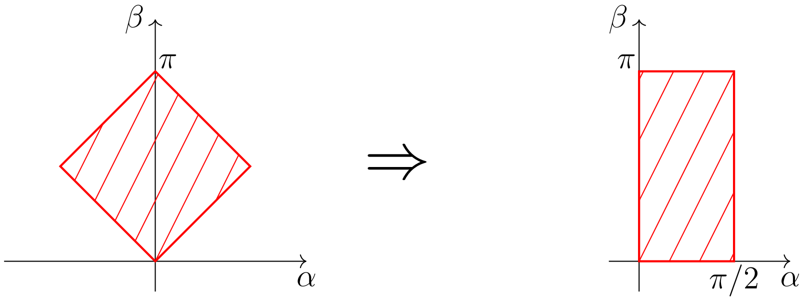
The free energy becomes

$$\begin{aligned}-\beta f &= \log 2 + \frac{1}{2\pi^2}\int_0^\pi\int_0^\pi \log\left(\cosh^2 2K - \sinh 2K(\cos\theta_1 + \cos\theta_2)\right)d\theta_1 d\theta_2\\ &= \log(2\cosh 2K) + \frac{1}{2\pi^2}\int_0^\pi\int_0^\pi \log\left[1 - \frac{\tau}{2}(\cos\theta_1 + \cos\theta_2)\right]d\theta_1 d\theta_2\end{aligned}$$

where

$$\tau = \frac{2\sinh 2K}{\cosh^2 2K}.$$

We can further simplify by a change of variable $\alpha = \frac{1}{2}(\theta_1 - \theta_2)$, $\beta = \frac{1}{2}(\theta_1 + \theta_2)$



$$-\beta f = \log(2\cosh 2K) + \frac{1}{\pi^2}\int_0^{\pi/2} d\alpha \int_0^{\pi} d\beta \ \log[1-\tau\cos\alpha\cos\beta].$$

We integrate $\beta$ using

$$\int_0^{\pi} \log(x - y\cos\beta) d\beta = \pi \log\left[\frac{x+\sqrt{x^2-y^2}}{2}\right]$$

thus

$$\begin{aligned}
-\beta f &= \log(2\cosh 2K) + \frac{1}{\pi}\int_0^{\pi/2} \log\left[\frac{1}{2}\left(1+\sqrt{1-\tau^2\cos^2\alpha}\right)\right] d\alpha \\
&= \log(2\cosh 2K) + \frac{1}{\pi}\int_0^{\pi/2} \log\left[\frac{1}{2}\left(1+\sqrt{1-\tau^2\sin^2\theta}\right)\right] d\theta.
\end{aligned}$$

The internal energy is

$$\begin{aligned}
u &= \frac{\partial}{\partial\beta}(\beta f) = J\frac{\partial}{\partial K}(\beta f) \\
&= -2J\tanh 2K - \frac{J}{\pi}\int_0^{\pi/2} \frac{\tau'(K)}{\tau}\left[1 - \frac{1}{\sqrt{1-\tau^2\sin^2\theta}}\right] d\theta \\
&= -J\coth 2K\left[1 + \frac{2}{\pi}(2\tanh^2 2K - 1)\int_0^{\pi/2}\frac{1}{\sqrt{1-\tau^2\sin^2\theta}} d\theta\right].
\end{aligned}$$

The critical temperature is at

$$\sinh^2 2K = 1 \iff \tau = 1.$$

Near the critical temperature $T_c$, the elliptic integral

$$\int_0^{\pi/2}\frac{1}{\sqrt{1-\tau^2\sin^2\theta}} d\theta \sim \frac{1}{2}\log\frac{16}{1-\tau^2} \quad \text{as } \tau \to 1^-$$

and

$$\begin{aligned}
1-\tau^2 &\propto (T-T_c)^2 \\
2\tanh^2 2K - 1 &\propto (T-T_c).
\end{aligned}$$

Thus $U$ is continuous at $T = T_c$. However, the analyticity breaks down at $T = T_c$. In particular, the specific heat per spin

$$c(T) = \frac{\partial u}{\partial T}$$

will have a logarithmic divergence

$$c(T) \propto \log |T - T_c|$$

near the critical temperature $T_c$. Thus for 2D Ising model at zero field near $T_c$, the first derivative of the free energy (internal energy) is continuous while the second derivative (specific heat) is singular. This illustrates a continuous (second-order) phase transition.

## 4.8 Spontaneous magnetization

We study the spontaneous magnetization of 2D Ising model on the square lattice, which is defined by

$$m(\beta) = \lim_{h \to 0^+} \lim_{M,N \to \infty} \frac{1}{MN} \sum_{m,n} \langle \sigma_{m,n} \rangle = \lim_{h \to 0^+} \lim_{M,N \to \infty} \frac{1}{\beta} \frac{1}{MN} \frac{\partial}{\partial h} \log Z_{M,N}.$$

By translation symmetry, we can also write $m(\beta)$ as a single spin expectation

$$m(\beta) = \lim_{h \to 0^+} \lim_{M,N \to \infty} \langle \sigma_{0,0} \rangle.$$

Note that the order of limit is important. Otherwise

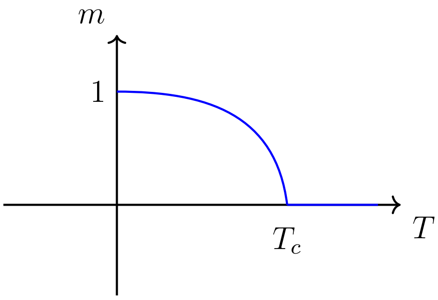
$$\lim_{h \to 0^+} \langle \sigma_{0,0} \rangle = 0 \qquad \text{at finite } M, N$$

by the $\mathbb{Z}_2$-symmetry $\sigma_{m,n} \to -\sigma_{m,n}$ at zero field $h = 0$.

In a conference at Cornell in 1948 and another conference in Florence in 1949, Onsager announced the following astonishing result on the simple exact formula

$$m = \begin{cases} \left[1 - \frac{1}{\sinh^2(2K_h)\sinh^2(2K_v)}\right]^{1/8} & T < T_c \\ 0 & T \geq T_c \end{cases}$$

where $T_c$ is the critical temperature.



Onsager never published a proof of this result. In 1952, Yang [46] published a full derivation of this result based on different method. We present a calculation of this famous formula following the simplification by Montroll-Potts-Ward [34] in terms of Szegő's limit formula for Toeplitz determinants.

What has been computed in Onsager and Yang is essentially

$$m^2 = \lim_{|r|\to\infty} \langle \sigma_0 \sigma_r \rangle$$

for the two-point function at large distance. Strictly speaking, this definition is not exactly the one defined above. These two definitions are expected to be equal, which was proved rigorously a few years later. See [4, 25].

### Transfer matrix and Pfaffian

Recall the partition function of 2D Ising model on the square lattice with periodic boundary condition can be formulated as

$$Z_{M,N} = \mathrm{Tr}(T^M)$$

where $T$ is the transfer matrix on the Hilbert space $H_N = (\mathbb{C}^2)^{\otimes N}$. We consider two spins in the same row, separated by a distance $r$, whose correlation is

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = \frac{1}{Z_{M,N}} \sum_{\sigma} \sigma_{1,1}\sigma_{1,r+1} \exp\left( K_h \sum_{m,n} \sigma_{m,n}\sigma_{m,n+1} + K_v \sum_{m,n} \sigma_{m,n}\sigma_{m+1,n} \right).$$

As we have see, the transfer matrix can be written via Clifford algebra

$$T = (2\sinh(2K_v))^{N/2}\Phi$$

where $\Phi$ has the form

$$\Phi = \Phi_h \Phi_v = \exp\left( iK_h \sum_{j=1}^{N} \Gamma_{2j}\Gamma_{2j+1} \right) \exp\left( iK_v^* \sum_{j=1}^{N} \Gamma_{2j-1}\Gamma_{2j} \right).$$

Thus the correlation can be written as

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = \frac{\mathrm{Tr}\ \sigma_z^{(1)}\sigma_z^{(r+1)}\Phi^M}{\mathrm{Tr}\ \Phi^M}.$$

It turns out to be convenient to work with

$$\tilde{\Phi} = \Phi_h^{\frac{1}{2}}\Phi_v\Phi_h^{\frac{1}{2}} = \Phi_h^{-\frac{1}{2}}\Phi\Phi_h^{\frac{1}{2}}$$

which is real and symmetric. Since

$$\sigma_z^{(j)}\Phi_h = \Phi_h\sigma_z^{(j)}$$

commutes, the two-spin correlation is the same as

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = \frac{\mathrm{Tr}\ \sigma_z^{(1)}\sigma_z^{(r+1)}\tilde{\Phi}^M}{\mathrm{Tr}\ \tilde{\Phi}^M}.$$

Since $\sigma_{m,n}^2 = 1$, we can write $\sigma_{1,1}\sigma_{1,r+1}$ as a spin chain $\sigma_{1,1}\sigma_{1,2}\sigma_{1,2}\sigma_{1,3}\cdots\sigma_{1,r}\sigma_{1,r+1}$. Using the relation $\sigma_z^{(j)}\sigma_z^{(j+1)} = i\Gamma_{2j}\Gamma_{2j+1}$, we arrive at

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = i^r\frac{\mathrm{Tr}\left(\Gamma_2\Gamma_3\Gamma_4\Gamma_5\ldots\Gamma_{2r}\Gamma_{2r+1}\tilde{\Phi}^M\right)}{\mathrm{Tr}\ \tilde{\Phi}^M}.$$

**Proposition 4.3** (Wick's Theorem for fermions)**.** *Let* $Q = \frac{1}{4}\sum_{a,b=1}^{2N} Q_{ab}\Gamma_a\Gamma_b$ *where* $Q_{ab} = -Q_{ba}$*. Let* $\eta_1,\ldots,\eta_{2N}$ *be Grassmann variables* ($\eta_a\eta_b = -\eta_b\eta_a$)*. Then*

$$\frac{Tr(e^{\sum_a \eta_a\Gamma_a}e^Q)}{Tr\ e^Q} = \exp\left(\frac{1}{2}\sum_{a,b} C_{ab}\eta_a\eta_b\right)$$

*where* $C_{ab}$ *is the two-point function*

$$C_{ab} = \begin{cases} Tr(\Gamma_a\Gamma_b e^Q)/\,Tr\ e^Q & a \neq b \\ 0 & a = b \end{cases}$$

*Proof:* We can assume

$$Q = \begin{pmatrix} 0 & \omega_1 & & & & & \\ -\omega_1 & 0 & & & & & \\ & & 0 & \omega_2 & & & \\ & & -\omega_2 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \omega_N \\ & & & & & -\omega_N & 0 \end{pmatrix}$$

Otherwise we can perform a linear orthonormal change of basis. This reduces the proof to the case $N = 1$ which is easily proved. □

**Corollary 4.4.** *Assume the same set-up in Proposition 4.3. Then for any* $k_1 < k_2 < \cdots < k_{2m}$*, we have*

$$\frac{Tr\left(\Gamma_{k_1}\Gamma_{k_2}\ldots\Gamma_{k_{2m}}e^Q\right)}{Tr\ e^Q} = Pf(A)$$

*where* $Pf(A)$ *is the Pfaffian of the skew-symmetric* $2m\times 2m$ *matrix*

$$A_{ij} = \frac{Tr\left(\Gamma_{k_i}\Gamma_{k_j}e^Q\right)}{Tr\ e^Q}, \quad i \neq j.$$

Let us denote

$$A_j = \Gamma_{2j-1}, \quad B_j = \Gamma_{2j}, \quad j = 1, \dots, N.$$

We will always use periodic index so $A_{N+1} \equiv A_1$ and $B_{N+1} \equiv B_1$. Then

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = i^r \frac{\left(\mathrm{Tr}\ B_1A_2B_2A_3\dots B_rA_{r+1}\Phi^M\right)}{\mathrm{Tr}\ \Phi^M}.$$

By Corollary 4.4, this correlation is determined by

$$\langle A_iA_j\rangle, \quad \langle B_iB_j\rangle, \quad \langle B_iA_j\rangle.$$

We can use symmetries of $\tilde{\Phi}$ to simplify these two point correlations.

(1) Translation symmetry: $\Gamma_a \to \Gamma_{a+2}$. This shifts

$$A_i \mapsto A_{i+1}, \quad B_i \mapsto B_{i+1}.$$

Thus $\langle A_iA_j\rangle$, $\langle B_iB_j\rangle$, $\langle B_iA_j\rangle$ only depend on their relative position $j-i$.

(2) Transpose symmetry: $\tilde{\Phi}^t = \tilde{\Phi}$. By the construction of $\Gamma_a$, we have

$$A_i^t = A_i, \quad B_i^t = -B_i.$$

By translation symmetry, transpose symmetry and Clifford relation,

$$\langle A_iA_j\rangle = \langle A_j^tA_i^t\rangle = \langle A_jA_i\rangle = 2\delta_{ij} - \langle A_iA_j\rangle$$

$$\Longrightarrow \quad \langle A_iA_j\rangle = \delta_{ij}.$$

Similarly,

$$\langle B_iB_j\rangle = \delta_{ij}.$$

Let us denote the $r \times r$ matrix

$$G_{kj} = i\langle B_kA_{j+1}\rangle \quad 1 \le k, j \le r.$$

By Corollary 4.4, $\langle \sigma_{1,1}\sigma_{1,r+1}\rangle$ is computed by the Pfaffian of

$$\begin{pmatrix} 0 & G \\ -G & 0 \end{pmatrix}.$$

Precisely, we find

$$\langle \sigma_{1,1}\sigma_{1,r+1}\rangle = (-1)^{r(r-1)/2} \det G.$$

Note that by translation symmetry

$$G_{kj} = i\langle B_0A_{k-j+1}\rangle$$

which is a Toeplitz matrix. As we will see, this Toeplitz property allows us to compute the large distance limit.

### Dominant eigenvector

Let $|\Omega\rangle$ be the eigenvector of $\tilde{\Phi}$ with the largest eigenvalue. We are interested in the thermodynamic limit $M \to \infty$

$$\lim_{M\to\infty} \langle B_0 A_n\rangle = \lim_{M\to\infty} \frac{\text{Tr } B_0 A_n \tilde{\Phi}^M}{\text{Tr } \tilde{\Phi}^M} = \langle \Omega | B_0 A_n | \Omega\rangle.$$

We will compute this thermodynamic correlation in the next.

Note that $A_j$ and $B_j$ are Hermitian matrices. Let us decompose into Fourier modes

$$A_j = \frac{1}{\sqrt{N}} \sum_\theta e^{i\theta j}(c_\theta + c^\dagger_{-\theta}), \quad B_j = \frac{i}{\sqrt{N}} \sum_\theta e^{i\theta j}(c_\theta - c^\dagger_{-\theta})$$

where the sum is over all $\theta$ such that $e^{iN\theta} = 1$. The Fourier modes $c_\theta, c^+_\theta$ satisfy the anti-commutator relations ($\{U, V\} := UV + VU$)

$$\{c_\theta, c^\dagger_\theta\} = 1 \qquad c_\theta^2 = (c^\dagger_\theta)^2 = 0$$

so they are $N$-copies of fermionic ladder operators. In terms of Fourier modes,

$$\Phi_h = \exp\left(iK_h \sum_j B_j A_{j+1}\right) = \exp\left(K_h \sum_\theta e^{i\theta}(c^\dagger_\theta - c_{-\theta})(c_\theta + c^\dagger_{-\theta})\right)$$
$$\Phi_v = \exp\left(iK_v^* \sum_j A_j B_j\right) = \exp\left(K_v^* \sum_\theta (c_\theta + c^\dagger_{-\theta})(c^\dagger_\theta - c_{-\theta})\right).$$

It is now straight-forward to compute the adjoint action

$$\begin{aligned}\text{Ad}_{\Phi_h}\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix} &= \exp\left[2K_h \begin{pmatrix} -\cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta\end{pmatrix}\right]\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix} \\ &= \begin{pmatrix} \cosh 2K_h - \sinh 2K_h \cos\theta & -i\sinh 2K_h \sin\theta \\ i\sinh 2K_h \sin\theta & \cosh 2K_h + \sinh 2K_h\cos\theta\end{pmatrix}\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix} \\ \text{Ad}_{\Phi_v}\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix} &= \begin{pmatrix} e^{2K_v^*} & \\ & e^{-2K_v^*}\end{pmatrix}\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix}\end{aligned}$$

Thus

$$\text{Ad}_{\tilde{\Phi}}\begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix} = M_\theta \begin{pmatrix} c_\theta \\ c^\dagger_{-\theta}\end{pmatrix}$$

where $M_\theta$ is the $2\times 2$ matrix

$$\begin{aligned}M_\theta =& \exp\left[K_h\begin{pmatrix} -\cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta\end{pmatrix}\right]\exp\left[\begin{pmatrix} 2K_v^* & \\ & -2K_v^*\end{pmatrix}\right]\exp\left[K_h\begin{pmatrix} -\cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta\end{pmatrix}\right] \\ =& \left(\cosh K_h I_2 + \sinh K_h \begin{pmatrix} -\cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta\end{pmatrix}\right)\left(\cosh 2K_v^* I_2 + \sinh 2K_v^* \begin{pmatrix} 1 & 0 \\ 0 & -1\end{pmatrix}\right) \\ & \left(\cosh K_h I_2 + \sinh K_h \begin{pmatrix} -\cos\theta & -i\sin\theta \\ i\sin\theta & \cos\theta\end{pmatrix}\right).\end{aligned}$$

Note that the matrix $M_\theta$ is of the form

$$M_\theta = \frac{1}{2}\mathrm{Tr}\ M_\theta + \begin{pmatrix} a & ib \\ -ib & -a \end{pmatrix} = \cosh\gamma(\theta) + \begin{pmatrix} a & ib \\ -ib & -a \end{pmatrix}, \quad a, b \in \mathbb{R}$$

where $e^{\pm\gamma(\theta)}$ $(\gamma(\theta) > 0)$ are eigenvalues of $M_\theta$. Explicit computation gives

$$\begin{aligned} -a &= \cosh 2K_v^* \sinh 2K_h \cos\theta - \sinh 2K_v^* \left(\cosh^2 K_h + \sinh^2 K_h \cos 2\theta\right) \\ -b &= \cosh 2K_v^* \sinh 2K_h \sin\theta - \sinh 2K_v^* \sinh^2 K_h \sin 2\theta \end{aligned}$$

and

$$\cosh\gamma(\theta) = \cosh 2K_h \cosh 2K_v^* - \sinh 2K_h \sinh 2K_v^* \cos\theta.$$

Let

$$\begin{pmatrix} u_\pm(\theta) & v_\pm(\theta) \end{pmatrix} M_\theta = e^{\pm\gamma(\theta)} \begin{pmatrix} u_\pm(\theta) & v_\pm(\theta) \end{pmatrix}$$

denote the left eigenvectors of $M_\theta$. Since $M_\theta$ is Hermitian, we can normalize

$$|u_+|^2 + |v_+|^2 = |u_-|^2 + |v_-|^2 = 1, \qquad \bar{u}_+ u_- + \bar{v}_+ v_- = 0.$$

Then

$$\eta_\pm(\theta) = u_\pm(\theta) c_\theta + v_\pm(\theta) c^+_{-\theta}.$$

are eigenvectors under the adjoint action

$$\tilde{\Phi}\eta_\pm\tilde{\Phi}^{-1} = e^{\pm\gamma(\theta)}\eta_\pm.$$

Since $\tilde{\Phi}$ is Hermitian,

$$\implies \quad \tilde{\Phi}^{-1}\eta_\pm^\dagger\tilde{\Phi} = e^{\pm\gamma(\theta)}\eta_\pm^\dagger$$

i.e. we can choose $\eta_\pm(\theta)$ such that

$$\eta_\pm^\dagger(\theta) = \eta_\mp(-\theta), \quad \text{i.e.} \quad \begin{aligned} u_\mp(-\theta) &= \overline{v_\pm(\theta)} \\ v_\mp(-\theta) &= \overline{u_\pm(\theta)} \end{aligned}.$$

By the normalization condition, we have the anti-commutator relation

$$\{\eta_+(\theta), \eta_+^\dagger(\theta)\} = 1$$

and these eigenvectors of $\tilde{\Phi}$ form new ladder operators. Since $\eta_+(\theta)$'s have eigenvalues $> 1$, the eigenstate $|\Omega\rangle$ of the largest eigenvalue of $\tilde{\Phi}$ is characterized by

$$\eta_+(\theta)|\Omega\rangle = 0 \qquad \forall\theta.$$

**Computation of $i\langle\Omega|B_0A_n|\Omega\rangle$**

In Fourier modes, we have

$$i\langle\Omega|B_0A_n|\Omega\rangle = \frac{1}{N}\sum_\theta e^{-in\theta}\langle(c^\dagger_{-\theta}-c_\theta)(c_{-\theta}+c^\dagger_\theta)\rangle.$$

Let

$$g(\theta) = \langle\Omega|(c^+_{-\theta}-c_\theta)(c_{-\theta}+c^+_\theta)|\Omega\rangle$$

so

$$i\langle\Omega|B_0A_n|\Omega\rangle = \frac{1}{N}\sum_\theta e^{-in\theta}g(\theta)$$

In the limit $N\to\infty$, this will become an integral

$$\frac{1}{2\pi}\int_0^{2\pi} e^{-in\theta}g(\theta)d\theta.$$

The projection of $c_{-\theta}+c^+_\theta$ to $\eta_-(-\theta)$ is

$$\left(\overline{u_-(-\theta)}+\overline{v_-(-\theta)}\right)\eta_-(-\theta) = (u_+(\theta)+v_+(\theta))\eta^+_+(\theta)$$

and the projection of $c^+_{-\theta}-c_\theta$ to $\eta_+(\theta)$ is

$$\left(-\overline{u_+(\theta)}+\overline{v_+(\theta)}\right)\eta_+(\theta).$$

Thus

$$g(\theta) = (u_+(\theta)+v_+(\theta))\left(-\overline{u_+(\theta)}+\overline{v_+(\theta)}\right).$$

Now the matrix $M_\theta$ has the form

$$M_\theta = \frac{1}{2}\mathrm{Tr}\ M_\theta I_2 + \begin{pmatrix} a & ib \\ -ib & -a\end{pmatrix}.$$

The eigenvector equation becomes

$$\begin{pmatrix} u_+ & v_+\end{pmatrix}\begin{pmatrix} a & ib \\ -ib & -a\end{pmatrix} = \begin{pmatrix} u_+ & v_+\end{pmatrix}\sqrt{a^2+b^2}$$

from which we find

$$\frac{v_+}{u_+} = i\frac{\sqrt{a^2+b^2}-a}{b} = i\xi.$$

By the normalization condition $|u_+|^2+|v_+|^2=1$, we find

$$|u_+|^2 = \frac{1}{1+\xi^2} \qquad |v_+|^2 = \frac{\xi^2}{1+\xi^2}$$
$$\overline{u_+}v_+ = \frac{i\xi}{1+\xi^2} \qquad u_+\overline{v_+} = \frac{-i\xi}{1+\xi^2}$$

Thus

$$\begin{aligned}
g(\theta) &= -\frac{1}{1+\xi^2} + \frac{\xi^2}{1+\xi^2} - 2i\frac{\xi}{1+\xi^2} \\
&= \frac{\xi^2 - 1 - 2i\xi}{1+\xi^2} = \frac{(\xi - i)^2}{(\xi+i)(\xi-i)} \\
&= \frac{\xi - i}{\xi + i} = \frac{|z| + z}{|z| + \overline{z}} \qquad\qquad \text{where } z = -a - ib \\
&= \frac{z}{|z|} = \left(\frac{z}{\overline{z}}\right)^{\frac{1}{2}}
\end{aligned}$$

which is a phase factor determined by $M_\theta$. Now

$$\begin{aligned}
z &= \cosh 2K_v^* \sinh 2K_h e^{i\theta} - \sinh 2K_v^* \cosh^2 K_h - \sinh 2K_v^* \sinh^2 K_h e^{2i\theta} \\
&= -\sinh 2K_v^* \sinh^2 K_h (e^{i\theta} - r_+)(e^{i\theta} - r_-)
\end{aligned}$$

where

$$r_\pm = \frac{(\tanh K_v^*)^{\mp 1}}{\tanh K_h} = e^{\pm 2K_v} \coth K_h.$$

Introduce

$$\alpha_1 = e^{-2K_v} \tanh K_h = r_+^{-1}, \qquad \alpha_2 = e^{-2K_v} \coth K_h = r_-.$$

Then

$$z \propto (1 - \alpha_1 e^{i\theta})(e^{i\theta} - \alpha_2)$$

and

$$g(\theta) = \left(\frac{(1 - \alpha_1 e^{i\theta})(e^{i\theta} - \alpha_2)}{(1 - \alpha_1 e^{-i\theta})(e^{-i\theta} - \alpha_2)}\right)^{\frac{1}{2}} = e^{i\theta} \left(\frac{(1 - \alpha_1 e^{i\theta})(1 - \alpha_2 e^{-i\theta})}{(1 - \alpha_1 e^{-i\theta})(1 - \alpha_2 e^{i\theta})}\right)^{\frac{1}{2}}.$$

Thus we find

$$\lim_{M,N\to\infty} G_{kj} = \lim_{M,N\to\infty} i\langle B_0 A_{k-j+1}\rangle = \frac{1}{2\pi}\int_0^{2\pi} e^{-i\theta(k-j)} \varphi(e^{i\theta}) d\theta$$

where

$$\varphi(e^{i\theta}) = \left[\frac{(1 - \alpha_1 e^{i\theta})(1 - \alpha_2 e^{-i\theta})}{(1 - \alpha_1 e^{-i\theta})(1 - \alpha_2 e^{i\theta})}\right]^{\frac{1}{2}}.$$

**Toeplitz matrix and Szegö limit theorem**

We consider $T < T_c$, In this case

$$0 < \alpha_1 < 1, \qquad 0 < \alpha_2 < 1.$$

We are interested in computing $\det(G)$ for the $r \times r$ matrix $G$ in the large distance limit $r \to \infty$. Since $G$ is a Toeplitz, such limit can be computed by Szegö's Theorem. Let

$$\psi(z) = \log \varphi(z)$$

and

$$c_n = \frac{1}{2\pi}\int_0^{2\pi} \psi(e^{i\theta})e^{-in\theta}d\theta$$

be the Fourier modes of $\psi$. Then Szegö's limit theorem says

$$\det G \sim \exp\left(rc_0 + \sum_{n=1}^{\infty} nc_n c_{-n}\right) \qquad \text{as } r \to \infty.$$

In our case, we have

$$\psi(z) = \frac{1}{2}\left[\log(1-\alpha_1 z) + \log(1-\alpha_2 z^{-1}) - \log(1-\alpha_1 z^{-1}) - \log(1-\alpha_2 z)\right]$$

We expand on $|z| = 1$ and since $|\alpha_1| < 1, |\alpha_2| < 1$

$$\psi(z) = \sum_{n=1}^{\infty} \frac{\alpha_2^n - \alpha_1^n}{2n} z^n + \sum_{n=1}^{\infty} \frac{\alpha_1^n - \alpha_2^n}{2n} z^{-n}.$$

The Fourier coefficients are read by

$$c_0 = 0, \qquad c_n = \frac{\alpha_2^n - \alpha_1^n}{2n}, \qquad c_{-n} = \frac{\alpha_1^n - \alpha_2^n}{2n}, \qquad n > 0.$$

By Szegö's limit theorem,

$$\lim_{r\to\infty} \det G = \exp\left(\sum_{n=1}^{\infty} nc_n c_{-n}\right).$$

Now

$$\begin{aligned}\sum_{n=1}^{\infty} nc_n c_{-n} &= -\sum_{n=1}^{\infty} \frac{(\alpha_1^n - \alpha_2^n)^2}{4n} = -\sum_{n=1}^{\infty} \frac{\alpha_1^{2n} - 2\alpha_1^n\alpha_2^n + \alpha_2^{2n}}{4n} \\ &= \frac{1}{4}\left[\log(1-\alpha_1^2) - 2\log(1-\alpha_1\alpha_2) + \log(1-\alpha_2^2)\right] \\ &= \frac{1}{4}\log\frac{(1-\alpha_1^2)(1-\alpha_2^2)}{(1-\alpha_1\alpha_2)^2}.\end{aligned}$$

We find the spontaneous magnetization

$$m = \left(\lim_{r\to\infty} \det G\right)^{\frac{1}{2}} = \left[\frac{(1-\alpha_1^2)(1-\alpha_2^2)}{(1-\alpha_1\alpha_2)^2}\right]^{\frac{1}{8}}.$$

We can further simplify this expression. Recall

$$\begin{aligned}\alpha_1 &= e^{-2K_v}\tanh K_h \\ \alpha_2 &= e^{-2K_v}\coth K_h\end{aligned}$$

so

$$\begin{aligned}\frac{(1-\alpha_1^2)(1-\alpha_2^2)}{(1-\alpha_1\alpha_2)^2} &= \frac{(1-e^{-4K_v}\tanh^2 K_h)(1-e^{-4K_v}\coth^2 K_h)}{(1-e^{-4K_v})^2}\\ &= \frac{1-e^{-4K_v}(\tanh^2 K_h+\coth^2 K_h)+e^{-8K_v}}{(1-e^{-4K_v})^2}\\ &= 1-\frac{(\tanh^2 K_h+\coth^2 K_h-2)}{(e^{2K_v}-e^{-2K_v})^2}\\ &= 1-\frac{1}{\sinh^2(2K_h)\sinh^2(2K_v)}\end{aligned}$$

Thus for $T<T_c$, we obtain the famous formula

$$m=\left[1-\frac{1}{\sinh^2(2K_h)\sinh^2(2K_v)}\right]^{\frac{1}{8}}.$$

Near $T_c$, define

$$k=\frac{1}{\sinh 2K_h \sinh 2K_v}.$$

As $T\nearrow T_c$, $k\nearrow 1$, so

$$m\sim(1-k^2)^{\frac{1}{8}}\sim(T_c-T)^{\frac{1}{8}}$$

which shows the critical exponent is $\frac{1}{8}$.

## 4.9 Peierls estimate

We discuss the famous Peierls argument [37] on the existence of a phase transition at low temperature for the ferromagnetic Ising model in dimension $d\geq 2$.

### Boundary conditions

Consider the nearest-neighbor ferromagnetic Ising model on a finite region $\Lambda\subset\mathbb{Z}^d$. A spin configuration in $\Lambda$ is an assignment

$$\sigma=(\sigma_x)_{x\in\Lambda}\ ,\quad \sigma_x\in\{\pm1\}.$$

We impose a boundary condition outside $\Lambda$ by fixing a choice

$$\eta=(\eta_y)_{y\notin\Lambda}\ ,\quad \eta_y\in\{\pm1\}.$$

We define the Hamiltonian on $\Lambda$ under a fixed boundary condition $\eta$ outside $\Lambda$ by

$$H_\Lambda(\sigma|\eta)=-J\sum_{\substack{\langle x,y\rangle\\ x,y\in\Lambda}}\sigma_x\sigma_y-J\sum_{\substack{\langle x,y\rangle\\ x\in\Lambda,y\notin\Lambda}}\sigma_x\eta_y$$

where the sums run over nearest-neighbor bonds and we assume $J > 0$ for the ferromagnetic case. The partition function is given by

$$Z_\Lambda^\eta = \sum_\sigma e^{-\beta H_\Lambda(\sigma|\eta)}$$

and the associated Gibbs measure is

$$\mu_\Lambda^\eta(\sigma) = \frac{1}{Z_\Lambda^\eta} e^{-\beta H_\Lambda(\sigma|\eta)}.$$

We will mostly consider the plus boundary condition

$$\eta_y = +1 \qquad \text{for all } y \notin \Lambda$$

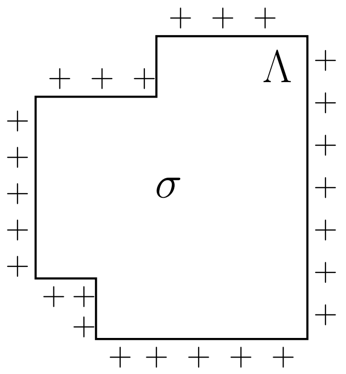
where the Hamiltonian, partition function and Gibbs measure will be denoted by

$$H_\Lambda(\sigma|+), \quad Z_\Lambda^+, \quad \mu_\Lambda^+(\sigma).$$

By spin-flip symmetry we also consider the minus boundary condition

$$\eta_y = -1 \qquad \text{for all } y \notin \Lambda$$

and $H_\Lambda(\sigma|-)$, $Z_\Lambda^-$, $\mu_\Lambda^-(\sigma)$ respectively.



Plus boundary condition

We are interested in the thermodynamic limit

$$\mu^+ = \lim_{\Lambda \nearrow \mathbb{Z}^d} \mu_\Lambda^+, \qquad \mu^- = \lim_{\Lambda \nearrow \mathbb{Z}^d} \mu_\Lambda^-.$$

Our goal is to show that

$$\mu^+ \neq \mu^- \qquad \text{at low temperature.}$$

This illustrates a phase transition.

Peierls' idea is to show that for sufficiently large $\beta$ (low temperature)

$$\mu^+(\sigma_0 = -1) < \frac{1}{2}\ , \quad \mu^-(\sigma_0 = -1) > \frac{1}{2}$$

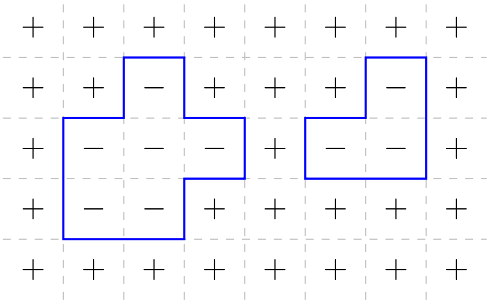
where 0 is the origin. Hence $\mu^+ \neq \mu^-$.

**Peierls estimate in $d = 2$**

We begin with $d = 2$ where the geometry is most transparent. We consider the dual lattice $(\mathbb{Z}^2)^*$ of $\mathbb{Z}^2$ which can be identified as

$$(\mathbb{Z}^2)^* = \mathbb{Z}^2 + \left(\frac{1}{2}, \frac{1}{2}\right).$$

Given a spin configuration $\sigma$ on $\Lambda$, we draw its dual bonds $\Gamma^*(\sigma)$ on the dual lattice which form a collection of closed contours.

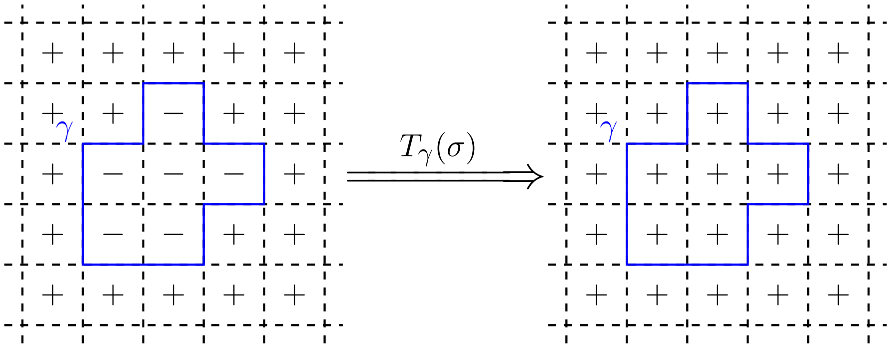


We consider plus boundary condition, so every connected component of minus spin is finite and is surrounded by such a contour. Thus

- minus droplets correspond to closed dual contours
- if $\sigma_0 = -1$, then the origin belongs to a minus component, hence is surrounded by at least one contour.

Let now $\sigma$ be a spin configuration with plus boundary condition, and $\gamma \subset \Gamma^*(\sigma)$ be the outer contour of a minus droplet. Let

$$T_\gamma(\sigma) = \text{flipping all spins of } \sigma \text{ inside } \gamma$$



The Hamiltonian changes under this flip by

$$H_\Lambda(\sigma|+) = H_\Lambda(T_\gamma(\sigma)|+) + 2J|\gamma|$$

where $|\gamma|$ is the number of edges in $\gamma$. In Boltzman weight

$$e^{-\beta H_\Lambda(\sigma|+)} = e^{-2\beta J|\gamma|}e^{-\beta H_\Lambda(T_\gamma(\sigma)|+)}.$$

Let $A_\gamma$ be the set of spin configurations $\sigma$ such that $\gamma$ is present as a contour. Then

$$\sum_{\sigma\in A_\gamma} e^{-\beta H_\Lambda(\sigma|+)} = e^{-2\beta J|\gamma|} \sum_{\sigma\in T_\gamma(A_\gamma)} e^{-\beta H_\Lambda(\sigma|+)} \leq e^{-2\beta J|\gamma|} \sum_{\sigma} e^{-\beta H_\Lambda(\sigma|+)} = e^{-2\beta J|\gamma|} Z_\Lambda^+.$$

This yields the crucial Peierls estimate

$$\mu_\Lambda^+(A_\gamma) \leq e^{-2\beta J|\gamma|}.$$

Consider spin configurations with

$$\sigma_0 = -1$$

under the plus boundary condition. We have

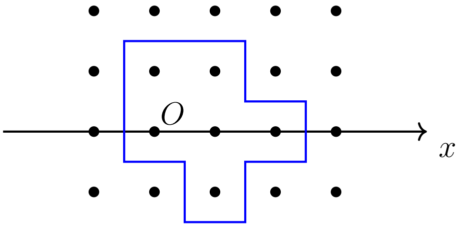
$$\{\sigma_0 = -1\} \subseteq \bigcup_{\gamma \text{ surrounds } 0} A_\gamma.$$

By the above Peierls estimate,

$$\mu_\Lambda^+(\sigma_0 = -1) \leq \sum_{\gamma \text{ surrounds } 0} \mu_\Lambda^+(A_\gamma) \leq \sum_{\gamma \text{ surrounds } 0} e^{-2\beta J|\gamma|}.$$

The problem is reduced to counting contours.

A contour of length $n$ surrounding the origin must intersect the positive x-axis within distance $n$ from the origin.



We can use the intersecting edge as the starting edge. There are at most $Cn$ such choices for some positive constant $C$ (for example, we can take $C = 1$ for a rough estimate). Once the starting edge and the orientation are chosen, at each step there are at most 3 choices for continuation (because one can not immediately backtrack). Thus the number $N_n$ of contours of length $n$ surrounding the origin is bounded by

$$N_n \leq Cn3^n.$$

Using this contour count,

$$\mu_\Lambda^+(\sigma_0 = -1) \leq \sum_{n\geq 4} N_n e^{-2\beta Jn} \leq C \sum_{n\geq 4} n(3e^{-2\beta J})^n.$$

This converges provided

$$3e^{-2\beta J} < 1 \quad \text{i.e.} \quad \beta > \frac{\log 3}{2J}.$$

For such $\beta$, the series sum can be made arbitrarily small as $\beta \to \infty$. In particular, for $\beta$ sufficiently large

$$\mu^+_\Lambda(\sigma_0 = -1) < \frac{1}{2}$$

uniformly in $\Lambda$. Therefore in the thermodynamic limit

$$\mu^+(\sigma_0 = -1) < \frac{1}{2}.$$

By spin-flip symmetry

$$\mu^-(\sigma_0 = +1) < \frac{1}{2} \qquad \text{i.e.} \qquad \mu^-(\sigma_0 = -1) > \frac{1}{2}.$$

This proves $\mu^+ \neq \mu^-$ and shows the existence of a phase transition at low temperature.

In the phase $\mu^+$, the magnetization density

$$m^+_\Lambda(\beta) = \mu^+_\Lambda(\sigma_0 = +1) - \mu^+_\Lambda(\sigma_0 = -1) = 1 - 2\mu^+_\Lambda(\sigma_0 = -1) > 0$$

uniformly in $\Lambda$. Thus in the thermodynamic limit

$$m^+(\beta) > 0.$$

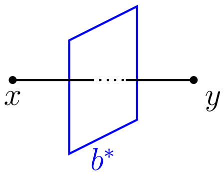
This shows a spontaneous magnetization in the plus phase.

### Peierls estimate in $d = 3$

One key feature of Peierls argument is its effectiveness also in $d \geq 3$. We illustrate the case $d = 3$. Consider a finite region $\Lambda \subset \mathbb{Z}^3$. The dual lattice is

$$(\mathbb{Z}^3)^* = \mathbb{Z}^3 + \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right).$$

Each nearest-neighbor bond $b = \langle x, y \rangle$ of $\mathbb{Z}^3$ corresponds dually to a unique square plaquette $b^*$ in the dual lattice



Given a spin configuration $\sigma$ on $\Lambda$ with plus boundary condition outside $\Lambda$, define the set of disagreement bonds

$$D(\sigma) = \{\langle x, y \rangle \mid \sigma_x \neq \sigma_y\}.$$

For each $\langle x, y\rangle \in D(\sigma)$, attach the dual plaquette. The union of all such dual plaquettes is the interface between plus and minus regions.

Let $C \subset \Lambda$ be a connected component of minus spins:

$$\begin{aligned} \sigma_x &= -1 \qquad \text{for } x \in C \\ \sigma_y &= +1 \qquad \text{for } y \notin C, \ \operatorname{dist}(y, C) = 1. \end{aligned}$$

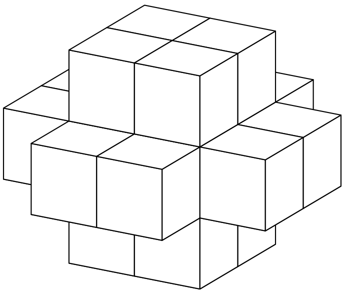
Its edge boundary is

$$\partial_e C = \{\langle x, y\rangle \mid x \in C, y \notin C\}.$$

For each bond in $\partial_e C$, place the corresponding dual plaquette. These plaquettes form a polyhedral surface $S(C)$. A bit of imagination shows that $S(C)$ is a closed surface. This is the 3D analogue of the dual closed contour in $d = 2$. Let

$$|S| = \# \text{ of plaquettes in } S = |\partial_e C|.$$

This $|S|$ is the analogue of contour length in $d = 2$.



S

Let $T_S\sigma$ be the spin configuration obtained by flipping all spins inside $S$. Then

$$H_\Lambda(\sigma|+) = H_\Lambda(T_S\sigma|+) + 2J|S|.$$

This leads to the Peierls estimate

$$\mu_\Lambda^+(A_S) \le e^{-2\beta J|S|}$$

where $A_S$ is the set of configurations $\sigma$ such that $S$ appears as the boundary of a minus droplet.

Let $N_n$ be the number of closed dual surface $S$ of area $|S| = n$ surrounding the origin. Then

$$\mu_\Lambda^+(\sigma_0 = -1) \le \sum_{n \ge 6} N_n e^{-2\beta J n}.$$

Here the minimal closed surface is the boundary of a single cube, which has 6 plaquettes.

We can proceed as in $d = 2$ to estimate $N_n$. We choose the first intersection of $S$ with the positive $x$-axis, and this gives $O(n)$ choices. Start from the anchor plaquette, we build a connected set of $n$ plaquettes by adding one plaquette at a time. Each plaquette has only finitely many neighboring plaquettes sharing an edge, say the number $= M$. In fact $M = 12$ in $d = 3$. Thus

$$N_n \leq CnM^n.$$

Substituting into the previous bound,

$$\mu_\Lambda^+(\sigma_0 = -1) \leq \sum_{n \geq 6} Cn(Me^{-2\beta J})^n.$$

So for sufficiently large $\beta$,

$$\mu_\Lambda^+(\sigma_0 = -1) < \frac{1}{2}$$

uniformly in $\Lambda$. In the thermodynamic limit

$$\mu^+(\sigma_0 = -1) < \frac{1}{2}.$$

This illustrates phase coexistence and thus a phase transition.

*Remark* 4.5 (Peierls argument fails in $d = 1$). In $d = 1$, a minus droplet is just an interval. Its boundary consists of two endpoints, so the energy cost is always $4J$, independent of droplet size. In $d \geq 2$, by contrast, a droplet with large boundary has energy cost proportional to its area, which grows with its size. This is the geometric reason the Peierls argument works.

## 4.10 Lee-Yang circle theorem

The Lee-Yang theory [27, 28] gives a conceptually beautiful explanation of how phase transitions arise in statistical mechanics. The central message is that instead of looking only at the partition function for real values of the external parameters, one analytically continues to complex values and studies the zeros of the partition function on the complex plane. Then a nonanalyticity of the thermodynamic free energy appears when, in the thermodynamic limit, the zeros of finite-volume partition functions accumulate and pinch the physical axis.

For the ferromagnetic Ising model in an external magnetic field, Lee and Yang proved a remarkable theorem: if one writes the partition function as a polynomial in a suitable fugacity variable, then all zeros lie on the unit circle. This implies that phase transitions in the ferromagnetic Ising model can possibly appear only at zero magnetic field $h = 0$.

### The Ising model with complex magnetic field

Consider a finite region $\Lambda \subset \mathbb{Z}^d$ with spins

$$\sigma_x \in \{\pm 1\}\ , \quad x \in \Lambda.$$

We consider a general Hamiltonian of the form

$$H_\Lambda(\sigma) = -\sum_{\langle x,y\rangle} J_{xy}\sigma_x\sigma_y - h\sum_{x\in\Lambda}\sigma_x$$

where we allow different coupling $J_{xy}$ on each nearest-neighbor bond $\langle x, y\rangle$. We require $J_{xy} \geq 0$ for all $\langle x, y\rangle$ in the ferromagnetic case. The partition function reads

$$Z_\Lambda(\beta, h) = \sum_\sigma \exp\left(\beta\sum_{\langle x,y\rangle} J_{xy}\sigma_x\sigma_y + \beta h\sum_x \sigma_x\right).$$

We want to study $Z_\Lambda$ not only for real $h$, but for <u>complex</u> $h$. Introduce the variable

$$z = e^{-2\beta h}.$$

Let

$$N = |\Lambda|\ , \quad n(\sigma) = \text{number of } -1 \text{ spins}.$$

We have

$$\sum_x \sigma_x = N - 2n(\sigma)$$

Thus

$$Z_\Lambda(\beta, h) = e^{\beta h N} P_\Lambda(z)$$

where

$$P_\Lambda(z) = \sum_\sigma \exp\left(\beta\sum_{\langle x,y\rangle} J_{xy}\sigma_x\sigma_y\right) z^{n(\sigma)}$$

is a polynomial of degree $N$ in $z$. Let

$$z_1, z_2, \dots, z_N$$

be zeros of $P_\Lambda(z)$. Then

$$P_\Lambda(z) = C_\Lambda \prod_{j=1}^N (z - z_j).$$

where the factor $C_\Lambda$ only depends on $\Lambda$. Taking logs,

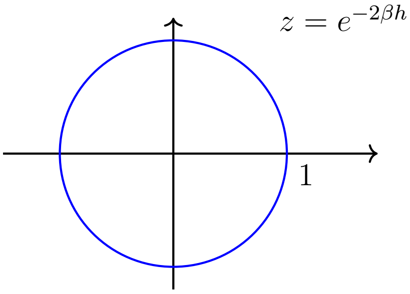
$$\frac{1}{N}\log Z_\Lambda = \beta h + \frac{1}{N}\log C_\Lambda + \frac{1}{N}\sum_{j=1}^N \log(z - z_j).$$

In the thermodynamic limit, the last term will become an integral on the distribution of zeros that capture the non-analyticity. Thus the accumulation/distribution of zeros in the thermodynamic limit tells the information on phase transitions.

Lee and Yang discovered that in the ferromagnetic Ising case, all zeros lie exactly on the unit circle

$$|z_j| = 1 \ , \qquad j = 1, 2, \ldots, N.$$

See Lee-Yang circle theorem below.



Thus in the thermodynamic limit, we expect

$$\frac{1}{N} \log Z_N \longrightarrow -\beta f = \beta h + C(\beta) + \frac{1}{2\pi} \int_0^{2\pi} \log(z - e^{i\theta}) \rho(\theta) d\theta$$

where $\rho(\theta)$ is the angular density of zeros on the unit circle. The magnetization density is given by

$$m(\beta, h) = -\frac{\partial}{\partial h} f = 1 - \frac{z}{\pi} \int_0^{2\pi} \frac{\rho(\theta) d\theta}{z - e^{i\theta}}$$

which is a Cauchy-type transform of the density of Lee-Yang zeros. Since

$$|z = e^{-2\beta h}| = 1 \iff \operatorname{Re} h = 0.$$

For real $h$, the thermodynamic phase transition can only occur at $z = 1$, i.e. $h = 0$. This leads to a concrete mechanism of phase transition:

- if $\rho(\theta)$ stays away from $z = 1$, $m$ stays analytic near $h = 0$.
- if zeros accumulate at $z = 1$, then the integral can develop a singularity, producing spontaneous magnetization and nonanalytic free energy.

### Lee-Yang circle theorem

We now discuss the fundamental property on distributions of zero discovered by Lee and Yang [28]. We discuss extension of Lee-Yang theorem in the form generalized by Asano [1] and Ruelle [40]. See also [41, 26] for further developments and expositions.

Consider a general Ising model on a finite undirected graph $G = (V, E)$. Here

$$V = \{1, 2, \ldots, N\}$$

is the set of vertices and $E$ is the set of edges. Let $J_{ij}$ be the coupling attached to the edge $(i, j) \in E$. Let $h_i$ be the external field attached to each vertex $i \in V$. The partition function is

$$Z_G = \sum_{\sigma\in\{\pm1\}^V} \exp\left(\beta \sum_{(i,j)\in E} J_{ij}\sigma_i\sigma_j + \beta \sum_{i\in V} h_i\sigma_i\right).$$

Let $K_{ij} = \beta J_{ij}$. Introduce variables

$$z_i = e^{-2\beta h_i} \qquad i \in V.$$

Then

$$Z_G = e^{\beta \sum\limits_{i\in V} h_i} P_G(z_1, \ldots, z_N)$$

where the multivariate Ising polynomial is

$$P_G(z_1, \ldots, z_N) = \sum_\sigma \exp\left(\sum_{(i,j)\in E} K_{ij}\sigma_i\sigma_j\right) \prod_{i:\sigma_i=-1} z_i.$$

Zeroes of $Z_G$ are the same as zeroes of $P_G$.

**Theorem 4.6** (Lee-Yang Circle Theorem). *Let $G = (V, E)$ be a finite graph and suppose $K_{ij} \geq 0$ for every edge (i.e. in the ferromagnetic case). Then the polynomial $P_G$ has the following zero-free properties:*

*(1) If $|z_i| < 1$ for every $i \in V$, then $P_G(z_1, \cdots, z_N) \neq 0$.*

*(2) If $|z_i| > 1$ for every $i \in V$, then $P_G(z_1, \cdots, z_N) \neq 0$.*

*In particular, every zero of $P_G(z, z, \cdots, z)$ lies on the unit circle:*

$$|z| = 1.$$

We present a simplified proof based on Asano contraction.

**Definition 4.7** (Asano contraction). Let

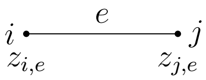
$$\Phi(u, v) = A + Bu + Cv + Duv$$

be affine in each of the two variables. Its Asano contraction is the one-variable polynomial

$$\widetilde{\Phi}(z) = A + Dz.$$

The Asano contraction has the interpretation of gluing half-edges variables into a single vertex variable as follows. For each edge $e = (i, j)$, we write two half-edge variables

$$z_{i,e} \quad \text{and} \quad z_{j,e}.$$



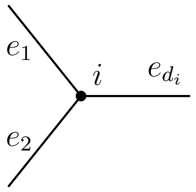
Define the edge factor

$$F_e(z_{i,e}, z_{j,e}) = e^{K_e} + e^{-K_e} z_{i,e} + e^{-K_e} z_{j,e} + e^{K_e} z_{i,e} z_{j,e}$$

and form the product

$$\mathcal{F}(\{z_{i,e}\}) = \prod_{e=(i,j)\in E} F_e(z_{i,e}, z_{j,e})$$

Fix a vertex $i$ of degree $d_i$. Let its incidence half-edge variable be $z_{i,e_1}, \cdots, z_{i,e_{d_i}}$



We now contract them, two at a time, into a single vertex variable $z_i$. At each contraction step, we view the current polynomial as a polynomial affine separately in two of the contracted variables

$$A + Bz_{i,e_a} + Cz_{i,e_b} + Dz_{i,e_a} z_{i,e_b}$$

and replace this expression by via Asano contraction

$$A + Dz_{i,e_a e_b}.$$

Repeating this operation at each vertex contracts

$$z_{i,e_1}, \cdots, z_{i,e_d} \implies z_i = z_{i,e_1 \dots e_d}$$

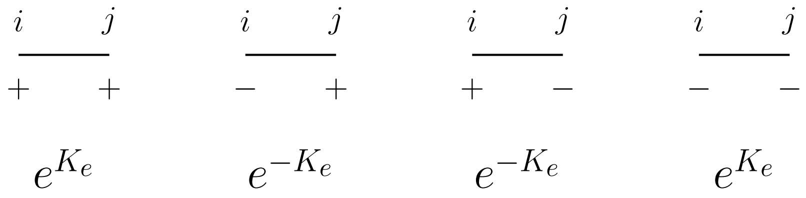
and we obtain a polynomial

$$\mathcal{F}(\{z_{i,e}\}) \implies \widetilde{\mathcal{F}}(z_1, z_2, \dots, z_N).$$

To identify $\widetilde{\mathcal{F}}$, observe that each edge factor contributes one of the four monomials

$$e^{K_e}, \quad e^{-K_e} z_{i,e}, \quad e^{-K_e} z_{j,e}, \quad e^{K_e} z_{i,e} z_{j,e}$$

These correspond to the four local spin configuration



This is exactly $e^{K_e \sigma_i \sigma_j}$.

The vertex contractions enforce the condition that all half-edges incident to a vertex choose the same spin state at that vertex. Thus the contracted monomial associated to a global spin configuration $\sigma$ is

$$\prod_{i:\sigma_i=-1} z_i$$

and its coefficient is

$$\exp\left(\sum_{(i,j)\in E} K_{ij}\sigma_i\sigma_j\right).$$

Consequently

$$\widetilde{\mathcal{F}}(z_1,\dots,z_N) = P_G(z_1,\dots,z_N).$$

Thus we find

$$\mathcal{F}(\{z_{i,e}\}) \xRightarrow[\text{Asano contraction}]{\text{Successive}} P_G(z_1,\dots,z_N)$$

**Lemma 4.8.** *Let* $\Phi(u,v) = A + Bu + Cv + Duv$. *Assume*

$$\Phi(u,v) \neq 0 \qquad \textit{whenever} \quad |u|<1,\ |v|<1.$$

*Then its Asano contraction* $\widetilde{\Phi}(z) = A + Dz$ *satisfies*

$$\widetilde{\Phi}(z) \neq 0 \qquad \textit{for} \quad |z|<1.$$

*Proof:* Let us write

$$\Phi(u,v) = (A+Cv) + (B+Dv)u.$$

By assumption,

$$|v|<1 \quad \text{imiplies} \quad |A+Cv| \geq |B+Dv|.$$

Take $v = \rho e^{i\theta}$, $0<\rho<1$, and average over $\theta$, we get

$$\frac{1}{2\pi}\int_0^{2\pi} |A + C\rho e^{i\theta}|^2 d\theta \geq \frac{1}{2\pi}\int_0^{2\pi} |B + D\rho e^{i\theta}|^2 d\theta$$

$$\Longrightarrow \quad |A|^2 + \rho^2|C|^2 \geq |B|^2 + \rho^2|D|^2 \qquad \forall 0<\rho<1$$

Switch the role of $u$ and $v$,

$$\Longrightarrow \quad |A|^2 + \rho^2|B|^2 \geq |C|^2 + \rho^2|D|^2 \qquad \forall 0<\rho<1$$

Adding the above two inequalities and let $\rho \to 1^-$, we get

$$|A| \geq |D|.$$

Note that $A = \Phi(0,0) \neq 0$. The lemma follows. $\square$

**Lemma 4.9.** *Let* $\Phi(u,v) = e^K + e^{-K}u + e^{-K}v + e^K uv$ *where* $K \geq 0$. *Then*

$$\Phi(u,v) \neq 0 \quad \textit{whenever} \quad |u| < 1, |v| < 1.$$

*Proof:* If $\Phi(u,v) = 0$, then

$$v = -\frac{1+tu}{t+u} \qquad \text{where} \quad t = e^{-K}.$$

Assume $|u| < 1$,

$$|1+tu|^2 - |t+u|^2 = 1 + t^2|u|^2 - t^2 - |u|^2 = (1-t^2)(1-|u|^2).$$

If $K > 0$ so $0 < t < 1$, then

$$|1+tu| > |t+u| \implies |v| > 1.$$

If $K = 0$ so $t = 1$, then $v = -1$. $\square$

<u>Proof of Theorem 4.6</u>: By Lemma 4.9,

$$\mathcal{F}(\{z_{i,e}\}) = \prod_{e=(i,j)\in E} F_e(z_{i,e}, z_{j,e})$$

satisfies

$$\mathcal{F}(\{z_{i,e}\}) \neq 0 \qquad \text{whenever all} \quad |z_{i,e}| < 1.$$

Since

$$\mathcal{F}(\{z_{i,e}\}) \xrightarrow[\text{Asano contraction}]{\text{Successive}} P_G(z_1, \dots, z_N)$$

By Lemma 4.8, we find

$$P_G(z_1, \dots, z_N) \neq 0 \qquad \text{whenever} \quad |z_1| < 1, \dots, |z_N| < 1.$$

Flipping the spin $\sigma_i \to -\sigma_i$, we have

$$P_G(z_1, \dots, z_N) = \left(\prod_{i=1}^N z_i\right) P_G(z_1^{-1}, \dots, z_N^{-1}).$$

If $|z_i| > 1$ for all $i$, then $|z_i^{-1}| < 1$ for all $i$. By the above proven case, we have

$$P_G(z_1, \dots, z_N) \neq 0 \qquad \text{whenever} \quad |z_1| > 1, \dots, |z_N| > 1.$$

$\square$

# 5 The Six-vertex Model

## Contents

## 5.1 Ice rule

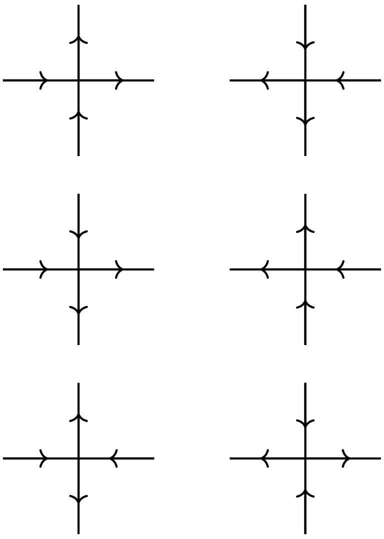
The six-vertex model is a statistical mechanical model on a square lattice. Each edge carries an arrow:

$$\longrightarrow \quad \longleftarrow \qquad \Big\uparrow \quad \Big\downarrow$$

At each vertex, four arrows meet. The rule for the allowed configuration is:

$$\text{at each vertex: } \boxed{\text{two arrows in and two arrows out}}$$

This is called the ice rule, because it comes from the hydrogen-bond arrangement in square ice. There are exactly six allowed local configurations.



weight $w = a$

weight $w = b$

weight $w = c$

They are grouped into three pairs, with Boltzman weights

$$a \ , \ b \ , \ c \ .$$

The two configurations of each pair above are related to each other by flipping the arrow directions. Such weight assignment of local configuration is under the assumption of no external field. In such case the Boltzmann weights should be invariant when all the polarizations are reversed. In general, we would have six different weights.

A global configuration $C$ is an assignment of arrows to every edge in the region $\Lambda$ under consideration, satisfying the ice rule at every vertex. Its statistical weight is

$$W(C) = \prod_{v \in V(C)} w(v)$$

where $V(C)$ is the set of vertices in $C$. The partition function of this model is

$$Z = \sum_C W(C) = \sum_C \prod_{v \in V(C)} w(v).$$

We consider a square lattice with $M$ rows and $N$ columns. We set the periodic boundary conditions in both directions. The partition function can be written as

$$Z_{M,N}(a,b,c) = \sum_C a^{N_a(C)} b^{N_b(C)} c^{N_c(C)}$$

where

$$N_a(C), \quad N_b(C), \quad N_c(C)$$

count how many vertices of each type appear. The thermodynamic free energy per site is

$$f = -\frac{1}{\beta} \lim_{M,N\to\infty} \frac{1}{MN} \log Z_{M,N}.$$

The exact computation of this free energy is one of the central achievements of solvable lattice models.

## 5.2 Transfer matrix

Let a row configuration of vertical arrows be

$$\sigma = (\sigma_1, \sigma_2, \cdots, \sigma_N) \qquad \sigma_i = \pm 1.$$

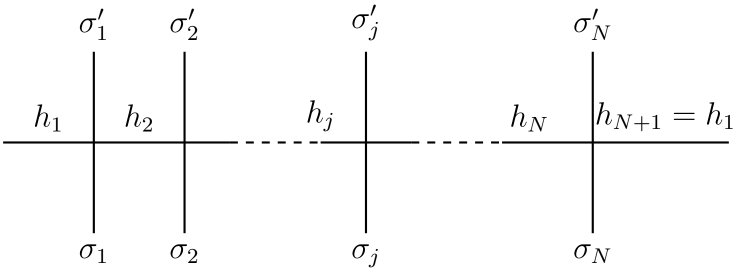
For example,

$$\uparrow \quad \uparrow \quad \downarrow \quad \uparrow \quad \downarrow \quad \downarrow$$

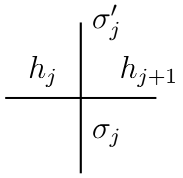
$$\sigma = (+,+,-,+,-,-)$$

We define the transfer matrix $T_\sigma^{\sigma'}$ by

$$T_\sigma^{\sigma'} = \sum_{h_1,\cdots,h_N} \prod_{j=1}^N R_{h_j \sigma_j}^{h_{j+1}\sigma'_j}$$



where

$$R^{h_{j+1}\sigma'_j}_{h_j\sigma_j} = \text{Weight of}$$  .

Under the periodic boundary conditions, the partition function $Z_{M,N}$ is given by

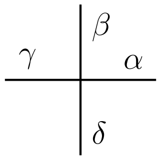
$$Z_{M,N} = \mathrm{Tr}\ T^M.$$

In the thermodynamic limit $M, N \to \infty$, the free energy is given by

$$-\beta f = \lim_{N\to\infty} \log \Lambda_{\max}$$

where $\Lambda_{\max}$ is the largest eigenvalue of the transfer matrix $T$.

At each vertex, the R-matrix defined above

$$R^{\alpha\beta}_{\gamma\delta} = \text{weight of}$$ 

defines a linear operator

$$R \in \mathrm{End}(\mathbb{C}^2 \otimes \mathbb{C}^2)$$

on basis by

$$R^{\alpha\beta}_{\gamma\delta} = \langle \alpha\beta | R | \gamma\delta \rangle.$$

Explicitly,

$$\begin{aligned} R^{++}_{++} &= R^{--}_{--} = a & \qquad + :\ \ \uparrow\ \ \rightarrow \\ R^{+-}_{+-} &= R^{-+}_{-+} = b & \\ R^{-+}_{+-} &= R^{+-}_{-+} = c & \qquad - :\ \ \downarrow\ \ \leftarrow \end{aligned}$$

where $+$ represents $\uparrow, \rightarrow$ and $-$ represents $\downarrow, \leftarrow$. In terms of the ordered basis

$$|{+}{+}\rangle, \quad |{+}{-}\rangle, \quad |{-}{+}\rangle, \quad |{-}{-}\rangle$$

the $R$-matrix takes the form

$$R = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ 0 & 0 & 0 & a \end{pmatrix}.$$

We introduce the spectral parameter $u$ such that the weights $a(u), b(u), c(u)$ depend on $u$. The corresponding six-vertex $R$-matrix becomes

$$R(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & a(u) \end{pmatrix}.$$

## 5.3 The Yang-Baxter equation

Let $V$ be a vector space. In general, an R-matrix is a linear map

$$R(u) \colon V \otimes V \to V \otimes V$$

parametrized by a complex variable $u$, called the spectral parameter. Let $R_{ij}$ denote the operator $R$ acting on the $i$th and $j$th tensor factors of

$$V \otimes V \otimes V.$$

Thus

$$R_{12}(u) = R(u) \otimes I, \quad R_{23}(u) = I \otimes R(u),$$

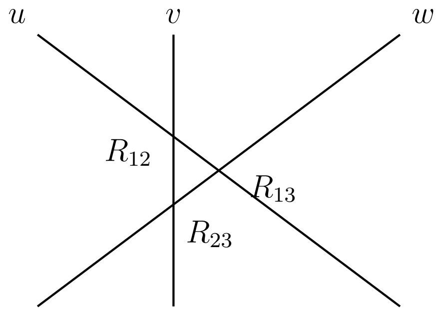
and $R_{13}(u)$ acts on the first and third tensor factors.

**Definition 5.1.** The $R$-matrix is said to satisfy the Yang-Baxter equation (YBE) if

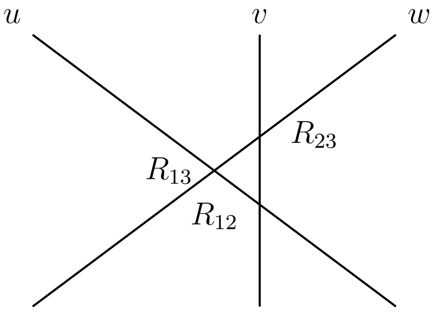
$$R_{12}(u-v)R_{13}(u-w)R_{23}(v-w) = R_{23}(v-w)R_{13}(u-w)R_{12}(u-v).$$

This is an identity of operators on $V \otimes V \otimes V$.

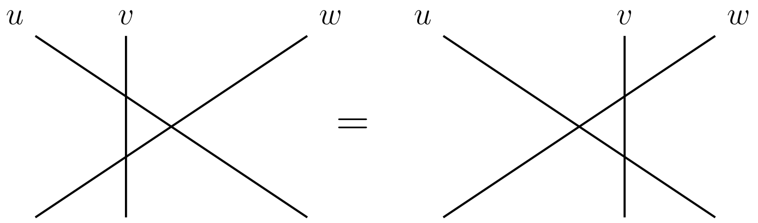
Graphically, the LHS is



the RHS is



Thus the Yang-Baxter equation says



One key fact about Yang-Baxter equation is on integrability. Let us consider

$$\mathcal{H}_N = V_1 \otimes V_2 \otimes \cdots \otimes V_N$$

where $V_i$'s are copies of $V$. We also introduce an auxiliary space $V_a = V$ indexed by $a$. Define the monodromy matrix

$$M_a(u) = R_{aN}(u) R_{a,N-1}(u) \cdots R_{a1}(u)$$

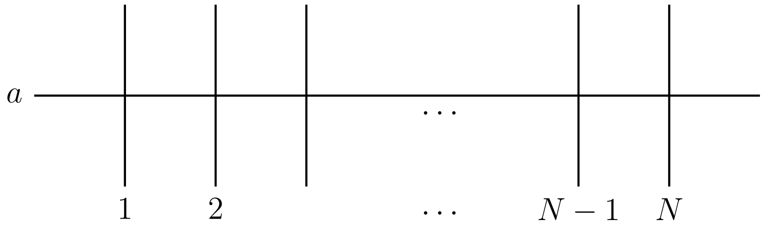
where

$$R_{aj}(u) : V_a \otimes V_j \to V_a \otimes V_j.$$

Thus

$$M_a(u) : V_a \otimes \mathcal{H}_N \to V_a \otimes \mathcal{H}_N.$$

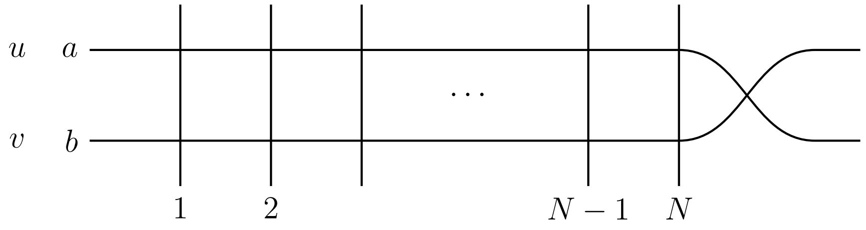
Graphically $M_a$ is represented by



**Proposition 5.2** (RTT relation)**.** *The Yang-Baxter equation implies*

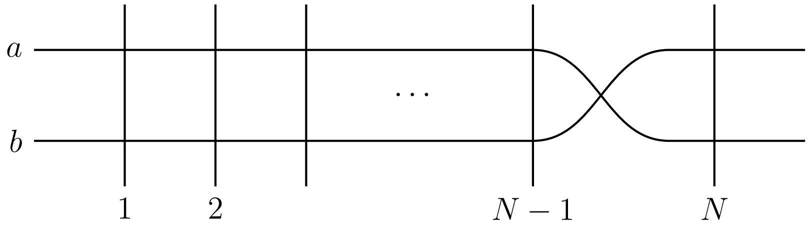
$$R_{ab}(u-v) M_a(u) M_b(v) = M_b(v) M_a(u) R_{ab}(u-v)$$

*Proof:* Graphically, the LHS of RTT relation is



We can apply Yang-Baxter equation successively

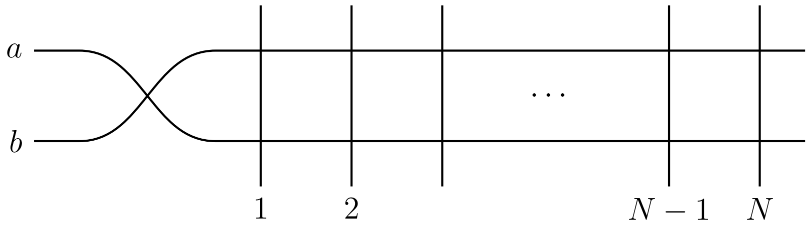
||



||

⋮

||



This is precisely the RHS of RTT relation. ☐

The transfer matrix is obtained by

$$T(u) = \mathrm{Tr}_a M_a(u) : \mathcal{H} \to \mathcal{H}$$

where $\mathrm{Tr}_a$ is the trace over the auxiliary $V_a$.

**Proposition 5.3.** *Assume $R_{ab}(u-v)$ is invertible. Then*

$$[T(u), T(v)] = 0 \qquad \forall u, v$$

*Proof:* Let $\mathrm{Tr}_{a,b}$ denote the trace over both auxiliary spaces $V_a$ and $V_b$. By the RTT relation

$$\begin{aligned}
\mathrm{Tr}_{a,b}(M_a(u)M_b(v)) &= \mathrm{Tr}_{a,b}(R_{ab}(u-v)M_a(u)M_b(v)R_{ab}^{-1}(u-v)) \\
&= \mathrm{Tr}_{a,b}(M_b(v)M_a(u)R_{ab}(u-v)R_{ab}^{-1}(u-v)) \\
&= \mathrm{Tr}_{a,b}(M_b(v)M_a(u)).
\end{aligned}$$

Thus

$$T(u)T(v) = \mathrm{Tr}_{a,b}(M_a(u)M_b(v)) = \mathrm{Tr}_{a,b}(M_b(v)M_a(u)) = T(v)T(u).$$

□

*Remark* 5.4. The essential mechanism is:

$$\mathrm{RTT} \implies M_a(u)M_b(v) \text{ is conjugate to } M_b(v)M_a(u)$$

In most situations we are interested, we can expand the transfer matrix

$$T(u) = \sum_{n=0}^{\infty} I_n u^n.$$

Because the entire generating functions commute

$$[T(u), T(v)] = 0 \qquad \forall u, v,$$

the coefficients also commute

$$[I_m, I_n] = 0.$$

Thus one obtains infinitely many commuting operators

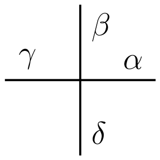
$$I_0, I_1, I_2, \cdots$$

The Hamiltonian belongs to this family, so these are conserved quantities.

## 5.4 Six-vertex solutions of YBE

Now let us analyze the R-matrix in the six-vertex model.

$$R^{\alpha\beta}_{\gamma\delta} = \text{weight of}$$



In component,

$$\begin{aligned} R^{++}_{++} &= R^{--}_{--} = a(u) \\ R^{+-}_{+-} &= R^{-+}_{-+} = b(u) \\ R^{-+}_{+-} &= R^{+-}_{-+} = c(u) \end{aligned}$$

Set

$$x = u - v \ , \quad y = v - w \ , \quad x + y = u - w$$

The YBE is

$$R_{12}(x)R_{13}(x+y)R_{23}(y) = R_{23}(y)R_{13}(x+y)R_{12}(x).$$

We look for $a(u), b(u), c(u)$ that solves this equation.

observe that the six-vertex R-matrix preserves the number of + signs. The check of YBE splits into four sectors:

$$0\ ,\ 1\ ,\ 2\ ,\ 3 \qquad \text{plus signs.}$$

The sectors with 0 and 3 plus signs are one-dimensional:

$$|- - -\rangle\ , \qquad |+ + +\rangle.$$

On these states the YBE is immediate, since both sides are multiplication by

$$a(x)a(x+y)a(y).$$

The nontrivial check is in the sectors

$$W_1 = \mathrm{Span}\{|+ - -\rangle, |- + -\rangle, |- - +\rangle\}$$
$$W_2 = \mathrm{Span}\{|- + +\rangle, |+ - +\rangle, |+ + -\rangle\}$$

By spin reversal symmetry, it is enough to check one of them, say $W_1$. On the ordered basis

$$|+ - -\rangle, \quad |- + -\rangle, \quad |- - +\rangle$$

the three two-site R-matrix act as follows:

$$R_{12} = \begin{pmatrix} b & c & 0 \\ c & b & 0 \\ 0 & 0 & a \end{pmatrix} \qquad R_{23} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & c \\ 0 & c & b \end{pmatrix} \qquad R_{13} = \begin{pmatrix} b & 0 & c \\ 0 & a & 0 \\ c & 0 & b \end{pmatrix}$$

**Proposition 5.5.** *Let*

$$a(u) = \sinh(u+\eta)\ , \quad b(u) = \sinh u\ , \quad c(u) = \sinh \eta$$

*where* $c = \sinh \eta$ *is independent of the spectral parameter* $u$*. Then the six-vertex R-matrix*

$$R(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & a(u) \end{pmatrix}$$

*satisfies the Yang-Baxter equation.*

*Proof:* For brevity write

$$a_u = a(u)\ , \quad b_u = b(u)\ , \quad c = c(u).$$

We check the YBE on the sector $W_1$ described above. The LHS of YBE acting on $W_1$ is

$$R_{12}(x)R_{13}(x+y)R_{23}(y) = \begin{pmatrix} b_x & c & 0 \\ c & b_x & 0 \\ 0 & 0 & a_x \end{pmatrix} \begin{pmatrix} b_{x+y} & 0 & c \\ 0 & a_{x+y} & 0 \\ c & 0 & b_{x+y} \end{pmatrix} \begin{pmatrix} a_y & 0 & 0 \\ 0 & b_y & c \\ 0 & c & b_y \end{pmatrix}$$

$$= \begin{pmatrix} a_y b_x b_{x+y} & c a_{x+y} b_y + c^2 b_x & c^2 a_{x+y} + c b_x b_y \\ c a_y b_{x+y} & a_{x+y} b_x b_y + c^3 & c a_{x+y} b_x + c^2 b_y \\ c a_x a_y & c a_x b_{x+y} & a_x b_{x+y} b_y \end{pmatrix}$$

The RHS of YBE acting on $W_1$ is

$$R_{23}(y)R_{13}(x+y)R_{12}(x) = \begin{pmatrix} a_y & 0 & 0 \\ 0 & b_y & c \\ 0 & c & b_y \end{pmatrix} \begin{pmatrix} b_{x+y} & 0 & c \\ 0 & a_{x+y} & 0 \\ c & 0 & b_{x+y} \end{pmatrix} \begin{pmatrix} b_x & c & 0 \\ c & b_x & 0 \\ 0 & 0 & a_x \end{pmatrix}$$

$$= \begin{pmatrix} a_y b_x b_{x+y} & c a_y b_{x+y} & c a_x a_y \\ c a_{x+y} b_y + c^2 b_x & a_{x+y} b_x b_y + c^3 & c a_x b_{x+y} \\ c^2 a_{x+y} + c b_x b_y & c a_{x+y} b_x + c^2 b_y & a_x b_{x+y} b_y \end{pmatrix}$$

Thus YBE is reduced to the following identities:

$$\begin{aligned} a_{x+y} b_y + c b_x &= a_y b_{x+y} \\ c a_{x+y} + b_x b_y &= a_x a_y \\ a_{x+y} b_x + c b_y &= a_x b_{x+y} \end{aligned}$$

Now substitute

$$a_u = \sinh(u+\eta)\ , \quad b_u = \sinh u\ , \quad c = \sinh\eta.$$

Thee above identities becomes elementary hyperbolic ones. □

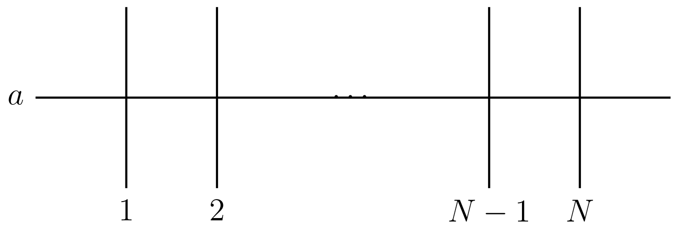
## 5.5 Algebraic Bethe ansatz

The original Bethe ansatz is a clever guess for many-body wavefunctions. We discuss Bethe ansatz for the Six-vertex model to diagonalize the transfer matrix.

We start with the algebraic Bethe ansatz. Recall the monodromy matrix

$$M_a(u) : V_a \otimes \mathcal{H}_N \to V_a \otimes \mathcal{H}_N$$

$$M_a(u) = R_{aN}(u) R_{a,N-1}(u) \cdots R_{a1}(u)$$



Here $\mathcal{H}_N = (\mathbb{C}^2)^{\otimes N}$ and $V_a = \mathbb{C}^2$. As a $2 \times 2$ matrix in the auxiliary space,

$$M_a(u) = \begin{pmatrix} A(u) & B(u) \\ C(u) & D(u) \end{pmatrix}_a$$

where $A(u), B(u), C(u), D(u)$ are operators on $\mathcal{H}_N$. The RTT relation

$$R_{ab}(u-v)M_a(u)M_b(v) = M_b(v)M_a(u)R_{ab}(u-v)$$

implies algebraic relations among $A, B, C, D$. In the ordered basis

$$|+_a +_b\rangle, \quad |+_a -_b\rangle, \quad |-_a +_b\rangle, \quad |-_a -_b\rangle$$

we have

$$R(u) = \begin{pmatrix} a(u) & 0 & 0 & 0 \\ 0 & b(u) & c(u) & 0 \\ 0 & c(u) & b(u) & 0 \\ 0 & 0 & 0 & a(u) \end{pmatrix}$$

where $a(u) = \sinh(u+\eta)$ , $b(u) = \sinh u$ , $c(u) = \sinh \eta$ for six-vertex model. And

$$M_a(u)M_b(v) = \begin{pmatrix} A(u)A(v) & A(u)B(v) & B(u)A(v) & B(u)B(v) \\ A(u)C(v) & A(u)D(v) & B(u)C(v) & B(u)D(v) \\ C(u)A(v) & C(u)B(v) & D(u)A(v) & D(u)B(v) \\ C(u)C(v) & C(u)D(v) & D(u)C(v) & D(u)D(v) \end{pmatrix}$$

$$M_b(v)M_a(u) = \begin{pmatrix} A(v)A(u) & B(v)A(u) & A(v)B(u) & B(v)B(u) \\ C(v)A(u) & D(v)A(u) & C(v)B(u) & D(v)B(u) \\ A(v)C(u) & B(v)C(u) & A(v)D(u) & B(v)D(u) \\ C(v)C(u) & D(v)C(u) & C(v)D(u) & D(v)D(u) \end{pmatrix}$$

The algebraic relations we need come from the $(1,4)$-entry, $(1,2)$-entry and $(3,4)$-entry of the RTT relation. The (1,4)-entry gives

$$B(u)B(v) = B(v)B(u).$$

The $(1,2)$-entry gives

$$a(u-v)A(u)B(v) = b(u-v)B(v)A(u) + c(u-v)A(v)B(u)$$

The $(3,4)$-entry gives

$$c(u-v)B(u)D(v)+b(u-v)D(u)B(v)=a(u-v)B(v)D(u)$$

It turns out to be convenient to work with another equivalent form for the $(1,2)$-entry identity. Swap $u$ and $v$, we get

$$a(v-u)A(v)B(u)=b(v-u)B(u)A(v)+c(v-u)A(u)B(v).$$

Since $c$ is a constant,

$$A(v)B(u)=\frac{b(v-u)}{a(v-u)}B(u)A(v)+\frac{c}{a(v-u)}A(u)B(v).$$

Substitute this into the previous (1,2)-entry equation

$$\left[a(u-v)-\frac{c^2}{a(v-u)}\right]A(u)B(v)=b(u-v)B(v)A(u)+\frac{cb(v-u)}{a(v-u)}B(u)A(v).$$

Using the six-vertex identity

$$a(u-v)a(v-u)-c^2=b(u-v)b(v-u)$$

we get

$$A(u)B(v)=\frac{a(v-u)}{b(v-u)}B(v)A(u)-\frac{c}{b(v-u)}B(u)A(v).$$

The above (3,4)-entry gives

$$D(u)B(v)=\frac{a(u-v)}{b(u-v)}B(v)D(u)-\frac{c}{b(u-v)}B(u)D(v).$$

Thus we find the following identities from the RTT relation

$$[B(u),B(v)]=0$$

$$\text{(AB-relation)}\qquad A(u)B(v)=f(v-u)B(v)A(u)-g(v-u)B(u)A(v)$$

$$\text{(DB-relation)}\qquad D(u)B(v)=f(u-v)B(v)D(u)-g(u-v)B(u)D(v)$$

where

$$f(x)=\frac{a(x)}{b(x)}\ ,\quad g(x)=\frac{c(x)}{b(x)}.$$

In particular, we can simultaneously diagonalize $B(u)$. The transfer matrix is

$$T(u)=\mathrm{Tr}_a M(u)=A(u)+D(u).$$

**Bethe vector and eigenvalue**

Let us first consider

$$|0\rangle = |+ + \cdots +\rangle \in \mathcal{H}_N.$$

Because of the triangular action of $R(u)$ on $+$ spin, one obtains

$$\begin{aligned} A(u)|0\rangle &= a(u)^N|0\rangle \\ D(u)|0\rangle &= b(u)^N|0\rangle \\ C(u)|0\rangle &= 0 \end{aligned}$$

Thus $|0\rangle$ can be viewed as a highest-weight vector. Consider the one-particle state

$$|\lambda\rangle = B(\lambda)|0\rangle.$$

By the AB-relation,

$$\begin{aligned} A(u)B(\lambda)|0\rangle &= f(\lambda - u)B(\lambda)A(u)|0\rangle - g(\lambda - u)B(u)A(\lambda)|0\rangle \\ &= a(u)^N f(\lambda - u)B(\lambda)|0\rangle - a(\lambda)^N g(\lambda - u)B(u)|0\rangle. \end{aligned}$$

Similarly

$$D(u)B(\lambda)|0\rangle = b(u)^N f(u - \lambda)B(\lambda)|0\rangle - b(\lambda)^N g(u - \lambda)B(u)|0\rangle.$$

Therefore the transfer matrix acts as

$$T(u)B(\lambda)|0\rangle = \Lambda(u;\lambda)B(\lambda)|0\rangle + \mathcal{U}(u;\lambda)B(u)|0\rangle$$

where

$$\begin{aligned} \Lambda(u;\lambda) &= a(u)^N f(\lambda - u) + b(u)^N f(u - \lambda) \\ \mathcal{U}(u;\lambda) &= -a(\lambda)^N g(\lambda - u) - b(\lambda)^N g(u - \lambda). \end{aligned}$$

If $\mathcal{U}(u;\lambda) = 0$, then $B(\lambda)|0\rangle$ is an eigenvector of the transfer matrix. This condition is the one-particle version of Bethe equation:

$$\mathcal{U}(u;\lambda) = 0.$$

Since $g(\lambda - u) = -g(u - \lambda)$, this condition becomes

$$\left(\frac{a(\lambda)}{b(\lambda)}\right)^N = 1.$$

In general, consider

$$|\lambda_1, \cdots, \lambda_m\rangle = B(\lambda_1)\cdots B(\lambda_m)|0\rangle.$$

We can use the AB and DB relations to compute how $T(u) = A(u) + D(u)$ acts on $|\lambda_1, \cdots, \lambda_m\rangle$.

**Proposition 5.6.** *If* $\lambda_1, \cdots, \lambda_m$ *satisfy the following* *Bethe equation*

$$\left[\frac{a(\lambda_j)}{b(\lambda_j)}\right]^N = \prod_{k\neq j} \frac{f(\lambda_j-\lambda_k)}{f(\lambda_k-\lambda_j)} \qquad \forall j.$$

*Then* $|\lambda_1, \cdots, \lambda_m\rangle$ *is an eigenvector of the transfer matrix matrix* $T(u)$ *with eigenvalue*

$$a(u)^N \prod_{j=1}^m f(\lambda_j - u) + b(u)^N \prod_{j=1}^m f(u-\lambda_j).$$

*Proof:* Iterate the above one-particle computation. Exercise. ☐

In the trigonometric parametrization of the six-vertex model, the Bethe equation becomes

$$\left[\frac{\sinh(\lambda_j+\eta)}{\sinh(\lambda_j)}\right]^N = \prod_{k\neq j} \frac{\sinh(\lambda_j-\lambda_k+\eta)}{\sinh(\lambda_j-\lambda_k-\eta)}$$

The corresponding eigenvalue is

$$a(u)^N \prod_{j=1}^m \frac{\sinh(\lambda_j-u+\eta)}{\sinh(\lambda_j-u)} + b(u)^N \prod_{j=1}^m \frac{\sinh(u-\lambda_j+\eta)}{\sinh(u-\lambda_j)}.$$

### Thermodynamic free energy

The thermodynamic free energy is

$$f(u) = -\frac{1}{\beta}\lim_{N\to\infty}\frac{1}{N}\log\Lambda_{\max}(u)$$

where $\Lambda_{\max}(u)$ is the largest eigenvalue of $T(u)$. This can be computed from Bethe ansatz as follows. Define

$$p(\lambda) = -i\log\frac{\sinh(\lambda+\eta)}{\sinh\lambda}$$
$$\theta(\lambda) = -i\log\frac{\sinh(\lambda+\eta)}{\sinh(\lambda-\eta)}$$

Taking the log of the Bethe equation, one obtains

$$Np(\lambda_j) = 2\pi I_j + \sum_{k\neq j}\theta(\lambda_j-\lambda_k)$$

where $I_j$ are called Bethe quantum numbers.

Define the density in the thermodynamic limit

$$\rho(\lambda) = \lim_{N\to\infty}\frac{1}{N}\sum_j \delta(\lambda-\lambda_j)$$

so that

$$\frac{1}{N}\sum_{j=1} F(\lambda_j) \to \int F(\lambda)\rho(\lambda)d\lambda.$$

Denote

$$G(\lambda) = \frac{1}{2\pi}\left[Np(\lambda) - \sum_k \theta(\lambda - \lambda_k)\right].$$

The logarithmic Bethe ansatz says

$$G(\lambda_j) = I_j - \frac{1}{2}.$$

Differentiate $G(\lambda)$ gives

$$G'(\lambda) = \frac{1}{2\pi}\left[Np'(\lambda) - \sum_k \theta'(\lambda - \lambda_k)\right].$$

The number of Bethe roots in a small interval $[\lambda, \lambda + d\lambda]$ is

$$N\rho(\lambda)d\lambda.$$

The number $I_j$ increases by one from one root to the next, thus

$$N\rho(\lambda) = G'(\lambda).$$

In the limit $N \to \infty$,

$$2\pi\rho(\lambda) = p'(\lambda) - \int \theta'(\lambda - \mu)\rho(\mu)d\mu.$$

This is the integral equation for the Bethe root density.

Take the Fourier transform

$$\hat{\varphi}(\omega) = \int_{-\infty}^{\infty} e^{i\omega\lambda}\varphi(\lambda)d\lambda.$$

The above integral equation becomes

$$2\pi\hat{\rho}(\omega) = \widehat{p'}(\omega) - \widehat{\theta'}(\omega)\hat{\rho}(\omega)$$

Inverting the Fourier transform, one solves

$$\rho(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\lambda}\frac{\widehat{p'}(\omega)}{2\pi + \widehat{\theta'}(\omega)}d\omega.$$

Now recall the Bethe eigenvalue

$$\Lambda(u) = a(u)^N \prod_{j=1}^{m} \frac{\sinh(\lambda_j - u + \eta)}{\sinh(\lambda_j - u)} + b(u)^N \prod_{j=1}^{m} \frac{\sinh(u - \lambda_j + \eta)}{\sinh(u - \lambda_j)}.$$

Define

$$\Phi_A(u) = \log a(u) + \int \log\frac{\sinh(\lambda - u + \eta)}{\sinh(\lambda - u)}\rho(\lambda)d\lambda$$

$$\Phi_D(u) = \log b(u) + \int \log\frac{\sinh(u - \lambda + \eta)}{\sinh(u - \lambda)}\rho(\lambda)d\lambda$$

Then

$$\Lambda(u) \approx e^{N\Phi_A(u)} + e^{N\Phi_D(u)} \qquad N \to \infty.$$

Consequently

$$f(u) = -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} \log \Lambda(u) = -\frac{1}{\beta} \max\{\Phi_A(u), \Phi_D(u)\}.$$

The phase diagram is organized by the six-vertex parameter

$$\Delta = \frac{a^2 + b^2 - c^2}{2ab}.$$

Plug in the hyperbolic parametrization

$$a(u) = \sinh(u+\eta), \quad b(u) = \sinh u, \quad c(u) = \sinh \eta$$

one finds

$$\Delta = \cosh \eta.$$

- ferroelectric phase: $\Delta > 1$
- anti-ferroelectric phase: $\Delta < -1$
- disordered phase: $-1 < \Delta < 1$.

The ferroelectric and anti-ferroelectric phases are gapful (massive), while the disordered phase (also called critical phase) is gapless (massless).

Let us consider the critical regime

$$-1 < \Delta < 1.$$

We write $\eta = i\gamma$ and

$$\Delta = \cosh \eta = \cos \gamma \qquad 0 < \gamma < \pi.$$

It is convenient to consider the shift

$$\lambda \to \lambda - \frac{\eta}{2} = \lambda - \frac{i}{2}\gamma$$

and denote

$$\tilde{p}(\lambda) = p\left(\lambda - \frac{i}{2}\gamma\right) = -i \log \frac{\sinh(\lambda + \frac{i}{2}\gamma)}{\sinh(\lambda - \frac{i}{2}\gamma)}$$
$$\tilde{\rho}(\lambda) = \rho\left(\lambda - \frac{i}{2}\gamma\right).$$

The root density integral equation is

$$2\pi\tilde{\rho}(\lambda) = \tilde{p}'(\lambda) - \int \theta'(\lambda - \mu)\tilde{\rho}(\mu) d\mu.$$

In this symmetric form, we can write

$$\frac{1}{2\pi}\tilde{p}'(\lambda) = a_1(\lambda) \quad , \quad \frac{1}{2\pi}\theta'(\lambda) = a_2(\lambda)$$

where

$$a_n(\lambda) = \frac{1}{2\pi}\frac{2\sin(n\gamma)}{\cosh(2\lambda) - \cos(n\gamma)}.$$

The Fourier transform of $a_n$'s are

$$\hat{a}_n(\omega) = \frac{\sinh\left(\frac{\pi - n\gamma}{2}\omega\right)}{\sinh\left(\frac{\pi}{2}\omega\right)}$$

for $0 < n\gamma < \pi$. The root density equation gives

$$\hat{\tilde{\rho}}(\omega) = \frac{\hat{a}_1(\omega)}{1 + \hat{a}_2(\omega)} = \frac{\sinh\left(\frac{\pi-\gamma}{2}\omega\right)}{\sinh\left(\frac{\pi}{2}\omega\right) + \sinh\left(\frac{\pi-2\gamma}{2}\omega\right)} = \frac{1}{2\cosh(\frac{\gamma\omega}{2})}.$$

Thus

$$\tilde{\rho}(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\lambda}\frac{1}{2\cosh(\gamma\omega/2)}\, d\omega = \frac{1}{2\gamma}\frac{1}{\cosh(\pi\lambda/\gamma)}.$$

This gives the root density in the critical regime

$$\rho\left(\lambda - \frac{i}{2}\gamma\right) = \frac{1}{2\gamma}\frac{1}{\cosh(\pi\lambda/\gamma)}.$$

## 5.6 Coordinate Bethe ansatz

We briefly discuss the traditional coordinate Bethe ansatz approach. The reference state is again

$$|0\rangle = |+ + \cdots +\rangle$$

and let $|x\rangle$ denote the state

$$|x\rangle = \left| + \cdots + \underset{\substack{\uparrow \\ \text{position } x}}{-} + + \cdots + \right\rangle.$$

A one-particle state is a superposition

$$|\psi\rangle = \sum_{x=1}^{N} \psi(x)|x\rangle.$$

Translation invariance suggests

$$\psi(x) = e^{ikx}$$

and periodic boundary condition implies

$$e^{ikN} = 1$$

which is the one-particle Bethe equation.

In general, we consider

$$|x_1, \cdots, x_m\rangle \qquad (1 \leq x_1 < \cdots < x_m \leq N)$$

where $x_j$ are positions of $-$ spins. For example of two particles in the ordered region $x_1 < x_2$, one takes the wave function

$$\psi(x_1, x_2) = A_{12} e^{i(k_1 x_1 + k_2 x_2)} + A_{21} e^{i(k_2 x_1 + k_1 x_2)}.$$

The first term corresponds to particle 1 carrying momentum $k_1$, particle 2 carrying momentum $k_2$. The second term corresponds to exchanged momentum. The state

$$\sum_{x_1 < x_2} \psi(x_1, x_2) |x_1, x_2\rangle$$

is an eigenstate if

$$A_{21} = S(\lambda_1, \lambda_2) A_{12}$$

where the scattering amplitude $S$ is

$$S(\lambda_1, \lambda_2) = \frac{\sinh(\lambda_1 - \lambda_2 + \eta)}{\sinh(\lambda_1 - \lambda_2 - \eta)}.$$

The parameters $\lambda_i$ is related to the momentum $k_i$ by

$$e^{ik} = \frac{a(u)}{b(u)} = \frac{\sinh(\lambda + \eta)}{\sinh \lambda}, \quad \Delta = \cosh \eta.$$

For $M$ particles, the coordinate Bethe ansatz assumes an eigenvector of the form

$$\sum_{1 \leq x_1 < x_2 < \cdots < x_M \leq N} \psi(x_1, \cdots, x_M) |x_1, \cdots, x_M\rangle$$

with wave function

$$\psi(x_1, \cdots, x_M) = \sum_{\sigma \in S_M} A_\sigma \prod_{j=1}^{M} e^{i k_{\sigma(j)} x_j}.$$

All amplitudes are determined by the scattering relation

$$A_{\cdots ji \cdots} = S(\lambda_i, \lambda_j) A_{\cdots ij \cdots}.$$

Consistency of different decompositions of the same permutation is precisely the factorization principle behind the Yang-Baxter equation

$$S_{12} S_{13} S_{23} = S_{23} S_{13} S_{12}.$$

Now send the $j$-th particle once around the periodic chain. It picks up a free phase $e^{ik_j N}$ and scatters with every other magnon. Periodicity gives

$$e^{ik_j N} = \prod_{k \neq j} S(\lambda_j - \lambda_k)$$

which is exactly the Bethe equation.

$$\left[\frac{\sinh(\lambda_j+\eta)}{\sinh(\lambda_j)}\right]^N=\prod_{k\neq j}\frac{\sinh(\lambda_j-\lambda_k+\eta)}{\sinh(\lambda_j-\lambda_k-\eta)}.$$

# 6 The Hubbard Model

## Contents

The Ising model and six-vertex model are classical lattice models. Their mathematical structure is already rich: transfer matrix, dualities, exact free energies, Yang-Baxter equations, and Bethe ansatz. We next consider the Hubbard model which introduces quantum fermions, spin, charge, and local interaction in the simplest possible way. This model was introduced independently in related form by Hubbard, Kanamori, and Gutzwiller in the early 1960s. The basic Hamiltonian is

$$H = -t \sum_{\langle i,j\rangle,\sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) + U \sum_i n_{i\uparrow} n_{i\downarrow}.$$

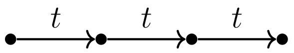
Here t is the nearest-neighbor hopping amplitude, and $U$ is the on-site interaction. The operators $c^\dagger_{i\sigma}$ creates an electron of spin $\sigma \in \{\uparrow, \downarrow\}$ at lattice site $i$, while $n_{i\sigma} = c^\dagger_{i\sigma} c_{i\sigma}$ is the number operator. This model is important for several reasons such as:

- It is the simplest lattice model of interacting electrons.
- At half filling and large positive $U$, it gives a Mott insulator and antiferromagnetism.
- In one dimension it is exactly solvable by the nested Bethe ansatz.
- For negative $U$, it gives a clean lattice model of Cooper pairing and BCS superconductivity.

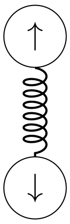
Basically,

$$\text{Hubbard model} = \text{Kinetic hopping} + \text{local Coulomb repulsion}$$

The competition between the two terms is the essential physics. The hopping term tries to delocalize electrons and form bands. The interaction term penalizes double occupation and tends to localize charge.



hopping delocalize electrons



on-site cost U.

## 6.1 Fermion algebra and the Hubbard Hamiltonian

Fermionic creation and annihilation operators satisfy the anti-commutator relations

$$\{c_{i\sigma}, c^\dagger_{j\sigma'}\} = \delta_{ij}\delta_{\sigma\sigma'}, \quad \{c_{i\sigma}, c_{j\sigma'}\} = \{c^\dagger_{i\sigma}, c^\dagger_{j\sigma'}\} = 0.$$

The local Hilbert space at one site is four dimensional with basis:

$$|0\rangle, \quad |{\uparrow}\rangle = c^\dagger_\uparrow|0\rangle, \quad |{\downarrow}\rangle = c^\dagger_\downarrow|0\rangle, \quad |{\uparrow\downarrow}\rangle = c^\dagger_\uparrow c^\dagger_\downarrow|0\rangle.$$

Let $\mathcal{H}_i$ denote such Hilbert space at site $i$. The full Hilbert space is the tensor product

$$\bigotimes_i \mathcal{H}_i.$$

The number operator at site $i$ is

$$n_{i\sigma} = c^\dagger_{i\sigma} c_{i\sigma}.$$

$n_{i\uparrow}$ counts the number of ↑ at site $i$, and $n_{i\downarrow}$ counts the number of ↓ at site $i$.

**Definition 6.1.** Let $\Lambda$ be a finite graph with vertex set also denoted by $\Lambda$. We write $\langle i, j\rangle$ for un-ordered nearest-neighbor edges. The Hubbard Hamiltonian on $\Lambda$ is

$$H = -t \sum_{\langle i,j\rangle,\sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) + U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_{i,\sigma} n_{i\sigma}.$$

Here $t > 0$ is hopping, $U$ is the on-site interaction, and $\mu$ is the chemical potential.

- $U > 0$: double occupation $|{\uparrow\downarrow}\rangle$ is penalized.
- $U < 0$: double occupation is favored and the model supports pairing.

The chemical potential is coupled to the total particle number

$$N = \sum_{i,\sigma} n_{i\sigma} = \sum_i n_{i\uparrow} + n_{i\downarrow}.$$

The particle number is conserved for the Hubbard Hamiltonian,

$$[H, N] = 0.$$

Define the spin operator at site $i$

$$\vec{S}_i = \frac{1}{2} \sum_{\alpha,\beta} c^\dagger_{i\alpha} \vec{\sigma}_{\alpha\beta} c_{i\beta}$$

where $\vec{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ are Pauli matrices. They satisfy su(2) commutation relations

$$[S^a_i, S^b_i] = \frac{i}{2}\epsilon^{abc} S^c_i.$$

Denote the total spin

$$\vec{S} = \sum_i \vec{S}_i$$

which satisfies the same su(2) commutation relation.

**Proposition 6.2** (Spin SU(2) symmetry)**.** *The Hubbard Hamiltonian satisfies*

$$[H, S^a] = 0, \qquad a = x, y, z.$$

**Example 6.3.** *Consider two sites and two electrons ($N = 2$). We omit the chemical potential ($\mu = 0$). The two electron states decompose into* SU(2) *multiplets. The spin-triplet states*

$$|\uparrow,\uparrow\rangle, \quad \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle), \quad |\downarrow,\downarrow\rangle$$

*have energy $H = 0$. The singlet*

$$|S\rangle = \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle)$$

*mixes with*

$$|D\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow,0\rangle + |0,\uparrow\downarrow\rangle).$$

*In the basis $|S\rangle, |D\rangle$, the Hamiltonian is*

$$H = \begin{pmatrix} 0 & -2t \\ -2t & u \end{pmatrix}.$$

*The lowest eigenvalue is*

$$E_{\text{min}} = \frac{u - \sqrt{u^2 + 16t^2}}{2}.$$

*For $u \gg t$,*

$$E_{\text{min}} = -\frac{4t^2}{u} + O(t^4/u^3).$$

*So the singlet energy is lower by $-\frac{4t^2}{u}$ than triplet. This is the simplest derivation of anti-ferromagnetic exchange.*

## 6.2 Mott physics

Set $t = 0$ and consider the atomic limit

$$H_{\text{at}} = U \sum_i n_{i\uparrow} n_{i\downarrow} - \mu \sum_i (n_{i\uparrow} + n_{i\downarrow}).$$

The one-site energies are

| State | particle number | energy |
|---|---|---|
| $\lvert 0\rangle$ | 0 | 0 |
| $\lvert\uparrow\rangle, \lvert\downarrow\rangle$ | 1 | $-\mu$ |
| $\lvert\uparrow\downarrow\rangle$ | 2 | $U - 2\mu$ |

**Definition 6.4** (Half filling)**.** For a lattice with $L$ sites, half filling means

$$N = L$$

Equivalently, the average number of electrons per site is one.

At half filling one usually takes $\mu = U/2$ (see Section 6.5). Then

$$E(0) = 0, \quad E(1) = -U/2, \quad E(2) = 0.$$

For $U > 0$, the singly occupied states are lowest. Thus at half filling and $t = 0$, the ground state manifold consists of all spin configurations with one electron per site. Charge is frozen, but spin remains free.

The elementary charge excitation creates a doubly occupied site. In the atomic limit this costs energy of order $U$. Therefore, even though band theory predicts a metal at half filling, interactions produce an insulating state. This is the Mott mechanism.

## 6.3 The t-J model

One key physical regime of Hubbard model is

$$U \gg t.$$

In this regime double occupation costs a large energy $U$. Therefore at low energies the system avoids doubly occupied sites. Nevertheless, electrons can still move when holes are present, and virtual hopping processes generate spin exchange. We derive the effective low-energy model in this regime, which is the t-J model.

To describe the exclusion of double occupancy, we introduce the Gutzwiller projection

$$P = \prod_i (1 - n_{i\uparrow} n_{i\downarrow}).$$

The projected Hilbert space at each site $i$ is

$$P\mathcal{H}_i = \mathrm{Span}\{|0\rangle_i, |\uparrow\rangle_i, |\downarrow\rangle_i\}.$$

The projected electron operators are

$$\begin{aligned} \tilde{c}_{i\sigma} &= P c_{i\sigma} P = c_{i\sigma}(1 - n_{i,-\sigma}) \\ \tilde{c}^\dagger_{i\sigma} &= P c^\dagger_{i\sigma} P = c^\dagger_{i\sigma}(1 - n_{i,-\sigma}) \end{aligned}$$

They annihilate any process that would create double occupancy.

Let us write the Hubbard Hamiltonian as

$$H_{\mathrm{Hub}} = T + H_U$$

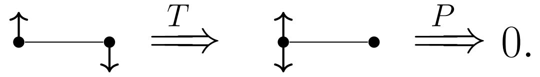
where

$$T = -t \sum_{\langle i,j\rangle,\sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) \ , \quad H_U = U \sum_i n_{i\uparrow} n_{i\downarrow}.$$

The projected hopping term is

$$\tilde{T} = PTP = -t \sum_{\langle i,j\rangle,\sigma} (\tilde{c}^\dagger_{i\sigma} \tilde{c}_{j\sigma} + \tilde{c}^\dagger_{j\sigma} \tilde{c}_{i\sigma}).$$

This term vanishes at exact half filling, where every site is singly occupied, but is nonzero when holes are introduced.



The projected on-site repulsion vanishes

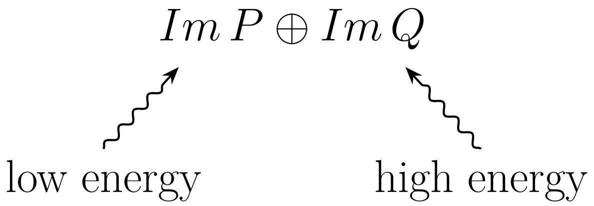
$$PH_U P = 0.$$

Let us denote

$$Q = 1 - P.$$

The operator $Q$ projects to states with at least one double occupancy, and therefore has energy of order $U$. So we can invert $H_U$ on $\mathrm{Im}(Q)$.

We decompose the Hilbert space as



The Hubbard Hamiltonian takes the block form

$$H_{\mathrm{Hub}} = \begin{pmatrix} PH_{\mathrm{Hub}}P & PH_{\mathrm{Hub}}Q \\ QH_{\mathrm{Hub}}P & QH_{\mathrm{Hub}}Q \end{pmatrix} = \begin{pmatrix} PTP & PTQ \\ QTP & QTQ + QH_U Q \end{pmatrix}.$$

The low energy effective Hamiltonian is an operator

$$H_{\mathrm{eff}} : \mathrm{Im}\ P \to \mathrm{Im}\ P$$

that reproduces the same spectrum in the low energy regime

$$\mathrm{Spec}(H_{\mathrm{eff}}) = \text{low energy } \mathrm{Spec}(H_{\mathrm{Hub}}).$$

This can be described by the resolvent relation

$$(z - H_{\mathrm{eff}})^{-1} = P(z - H_{\mathrm{Hub}})^{-1} P \quad \text{at low energy}$$

which leads to

$$\begin{aligned}H_{\text{eff}}(z) &= PH_{\text{Hub}}P + PH_{\text{Hub}}Q(z - QH_{\text{Hub}}Q)^{-1}QH_{\text{Hub}}P \\ &= PTP + PTQ(z - QH_UQ - QTQ)^{-1}QTP.\end{aligned}$$

For $U \gg t$, low energy satisfies $z \ll U$, so

$$(z - QH_UQ - QTQ)^{-1} = -(QH_UQ)^{-1} + O(t/U^2).$$

Therefore the effective Hamiltonian through order $t^2/U$ is

$$H_{\text{eff}} = PTP - PTQ\frac{1}{H_U}QTP + O(t^3/U^2).$$

If the intermediate states have exactly one double occupancy, then $QH_UQ = UQ$. So

$$H_{\text{eff}} = PTP - \frac{1}{U}PTQTP + O(t^3/U^2).$$

**Proposition 6.5.** *On a nearest-neighbor bond* $\langle i, j\rangle$,

$$PT_{ij}QT_{ij}P = -4t^2\left(\vec{S}_i \cdot \vec{S}_j - \frac{1}{4}n_in_j\right)$$

*where* $T_{ij} = -t\sum_\sigma(c^\dagger_{i\sigma}c_{j\sigma} + c^\dagger_{j\sigma}c_{i\sigma})$.

*Proof:* Let us denote

$$\Phi_{ij} = PT_{ij}QT_{ij}P.$$

If one of the two sites is empty, hopping $T_{ij}$ does not create double occupancy. Hence $Q$ kills the result. Thus the only nonzero action of $\Phi_{ij}$ is on the two-electron sector

$$|\uparrow,\uparrow\rangle, \quad |\uparrow,\downarrow\rangle, \quad |\downarrow,\uparrow\rangle, \quad |\downarrow,\downarrow\rangle.$$

The SU(2) triplet states are

$$|t_+\rangle = |\uparrow,\uparrow\rangle, \quad |t_-\rangle = |\downarrow,\downarrow\rangle, \quad |t_0\rangle = \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle).$$

For $|t_\pm\rangle$, $T_{ij}|t_\pm\rangle = 0$ by Pauli exclusion. For $|t_0\rangle$, the two virtual hopping cancel by signs

$$PT_{ij}QT_{ij}P|t_0\rangle = 0.$$

Thus $\Phi_{ij} = 0$ on triplet states.

The SU(2) singlet state is

$$|s\rangle = \frac{1}{\sqrt{2}}(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle).$$

Direct computation of the virtual hopping gives

$$\Phi_{ij}|s\rangle = 4t^2|s\rangle.$$

On the other hand

$$\vec{S}_i \cdot \vec{S}_j = \begin{cases} -3/4 & \text{Singlet} \\ 1/4 & \text{triplet} \end{cases}$$

The proposition follows. □

By Proposition 6.5, we find

$$H_{\text{eff}} = H_{t,J} + O(t^3/U^2)$$

where

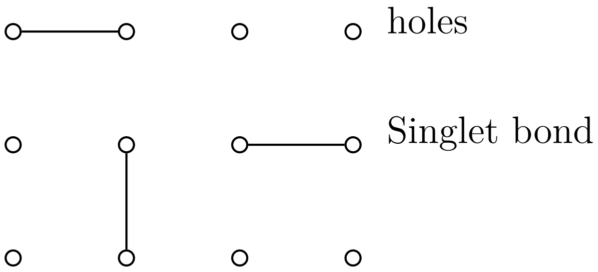
$$H_{t,J} = -t \sum_{\langle i,j\rangle,\sigma} P(c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma})P + J \sum_{\langle i,j\rangle} \left(\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j\right), \quad J = \frac{4t^2}{U}$$

is the Hamiltonian of the t-J model.

Note that the anti-ferromagnetic exchange $J = 4t^2/U$ is not put into the Hubbard model by hand. It is generated dynamically by virtual charge fluctuations. At half-filling, each site is occupied by a single electron. The hopping term vanishes due to forbidden of double occupancy. The t-J Hamiltonian becomes the Heisenberg model

$$J \sum_{\langle i,j\rangle} \left(\vec{S}_i \cdot \vec{S}_j - \frac{1}{4}\right), \quad J = \frac{4t^2}{U}.$$

| Regime | Low-energy model | Physics |
|---|---|---|
| $U \gg t$, half-filling | Heisenberg model | anti-ferromagnetism |
| $U \gg t$, finite hopping | t-J model | holes in spin background |



## 6.4 RVB

We have seen that at half filling and strong coupling $U \gg t$, the low energy subspace has one electron per site, and virtual hopping produces anti-ferromagnetic superexchange

$$J \sum_{\langle i,j\rangle} \left(\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j\right), \quad J = \frac{4t^2}{U}.$$

Upon doping the Mott insulator, the low energy effective theory is the t-J model

$$H_{t,J} = -t \sum_{\langle i,j \rangle,\sigma} P(c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma})P + J \sum_{\langle i,j \rangle} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j \right)$$

where $P$ is the Gutzwiller projector

$$P = \prod_i (1 - n_{i\uparrow} n_{i\downarrow}).$$

The anti-ferromagnetic term has an especially important form in the singlet channel. Define the nearest-neighbor singlet creation operator

$$B^\dagger_{ij} = \frac{1}{\sqrt{2}} (c^\dagger_{i\uparrow} c^\dagger_{j\downarrow} - c^\dagger_{i\downarrow} c^\dagger_{j\uparrow}).$$

The corresponding singlet is

$$|ij\rangle = B^\dagger_{ij} |0\rangle = \frac{1}{\sqrt{2}} (|{\uparrow_i \downarrow_j}\rangle - |{\downarrow_i \uparrow_j}\rangle).$$

**Proposition 6.6** (Exchange as a singlet attraction)**.** *On the P-projected Hilbert space,*

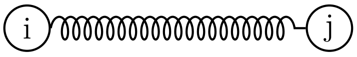
$$\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j = -B^\dagger_{ij} B_{ij}.$$

*Proof:* $\vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j$ has eigenvalue $-1$ on the spin singlet and 0 on the triplets. □

Thus the exchange term can be written as

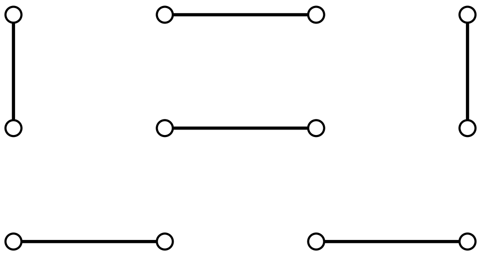
$$J \sum_{\langle i,j \rangle} \left( \vec{S}_i \cdot \vec{S}_j - \frac{1}{4} n_i n_j \right) = -J \sum_{\langle i,j \rangle} B^\dagger_{ij} B_{ij}.$$

The anti-ferromagnetic coupling can be read as an attraction in the nearest-neighbor spin-singlet bond.



$$\frac{1}{\sqrt{2}} (|{\uparrow_i \downarrow_j}\rangle - |{\downarrow_i \uparrow_j}\rangle).$$

A valence-bond state is a product of such singlet bonds pairing sites in a lattice.



Let $\Lambda$ be a set of $N$ sites with $N$ even. A valence-bond covering $D$ is a partition of $\Lambda$ into unordered pairs $(i,j)$. The corresponding valence-bond state is

$$|D\rangle = \prod_{(i,j)\in D} B_{ij}^{\dagger}|0\rangle.$$

The resonating valence bond (RVB), introduced by Anderson in the context of Cuprates, is a quantum superposition of valence-bond covering:

$$|\Psi_{\text{RVB}}\rangle = \sum_{D} A(D)|D\rangle.$$

The phrase "resonating valence bond" means that the ground state is not one fixed pattern of dimers but a coherent superposition of many singlet coverings. RVB theory was proposed as a way in which a quantum anti-ferromagnet may avoid conventional magnetic order. The system fluctuates among singlet pairings. In some setting this gives a quantum spin liquid.

## 6.5 Mean-field theory

The Hubbard model is an interacting many-body problem that is generally hard to solve in dimension $d \geq 2$. Mean-field theory is a systematic approximation in which an interacting many-body problem is replaced by an effective one-particle problem in a self-consistent background field. The general idea of mean field theory is as follows.

Suppose we have two operators $A$ and $B$. Write

$$A = \langle A\rangle + \delta A\ , \quad B = \langle B\rangle + \delta B.$$

Then

$$AB = \langle A\rangle \langle B\rangle + \langle A\rangle \delta B + \langle B\rangle \delta A + \delta A \delta B.$$

Mean-field theory neglects the quadratic fluctuation term

$$\delta A \delta B \approx 0.$$

Hence

$$AB \approx A\langle B\rangle + \langle A\rangle B - \langle A\rangle \langle B\rangle .$$

The expectation values appearing in the approximate Hamiltonian are not arbitrary external parameters. They must be computed in the ground state, or thermal state, of the same approximate Hamiltonian. This leads to certain consistency condition.

In the Hubbard model

$$H = -t \sum_{\langle i,j\rangle,\sigma} (c_{i\sigma}^{\dagger} c_{j\sigma} + c_{j\sigma}^{\dagger} c_{i\sigma}) + U \sum_{i} n_{i\uparrow} n_{i\downarrow} - \mu \sum_{i} n_i$$

where

$$n_{i\sigma} = c^\dagger_{i\sigma} c_{i\sigma} \ , \quad n_i = n_{i\uparrow} + n_{i\downarrow}.$$

The difficult part is the quartic interaction

$$U n_{i\uparrow} n_{i\downarrow}.$$

Define

$$\rho_{i\uparrow} = \langle n_{i\uparrow} \rangle \ , \quad \rho_{i\downarrow} = \langle n_{i\downarrow} \rangle$$

Then the mean field approximation is

$$n_{i\uparrow} n_{i\downarrow} \approx n_{i\uparrow} \langle n_{i\downarrow} \rangle + \langle n_{i\uparrow} \rangle n_{i\downarrow} - \langle n_{i\uparrow} \rangle \langle n_{i\downarrow} \rangle \, .$$

The corresponding mean-field Hamiltonian is

$$H_{\mathrm{MF}} = -t \sum_{\langle i,j \rangle, \sigma} (c^\dagger_{i\sigma} c_{j\sigma} + c^\dagger_{j\sigma} c_{i\sigma}) + U \sum_i (\rho_{i\downarrow} n_{i\uparrow} + \rho_{i\uparrow} n_{i\downarrow} - \rho_{i\uparrow} \rho_{i\downarrow}) - \mu \sum_i n_i.$$

This Hamiltonian is now quadratic in the fermion operators.

It is often clearer to introduce local charge density and local magnetization:

$$\rho_i = \langle n_i \rangle \ , \quad m_i = \langle n_{i\uparrow} - n_{i\downarrow} \rangle$$

ao

$$\rho_{i\uparrow} = \frac{\rho_i + m_i}{2} \ , \quad \rho_{i\downarrow} = \frac{\rho_i - m_i}{2} \ .$$

Then

$$n_{i\uparrow} n_{i\downarrow} \approx \frac{\rho_i - m_i}{2} n_{i\uparrow} + \frac{\rho_i + m_i}{2} n_{i\downarrow} - \frac{\rho_i^2 - m_i^2}{4} = \frac{\rho_i}{2} n_i - \frac{m_i}{2} (n_{i\uparrow} - n_{i\downarrow}) - \frac{1}{4} (\rho_i^2 - m_i^2).$$

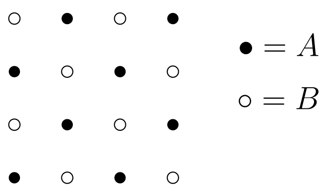
Since $S_i^z = \frac{1}{2}(n_{i\uparrow} - n_{i\downarrow})$, this can be written as

$$n_{i\uparrow} n_{i\downarrow} \approx \frac{\rho_i}{2} n_i - m_i S_i^z - \frac{1}{4} (\rho_i^2 - m_i^2).$$

This interaction creates an effective site-dependent magnetic field proportional to $U m_i$.

### Anti-ferromagnetic mean-field at half filling

The most important mean-field solution of the repulsive Hubbard model on the square lattice is anti-ferromagnetic order at half filling. The square lattice is bipartite: it decomposes into sublattices A and B such that nearest-neighbors of A lie in B and vice versa.



At half filling, we have $\rho_i = 1$. The system naturally tends toward anti-ferromagnetic order. We take the simplest anti-ferromagnetic pattern:

$$m_i = \begin{cases} +m & i \in A \\ -m & i \in B \end{cases}$$

Let $R_i = (x_i, y_i) \in \mathbb{Z}^2$ denote the position of i-th site. Let $Q = (\pi, \pi)$. Then we have

$$m_i = e^{iQ\cdot R_i} m.$$

At half-filling, there is a special particle-hole symmetry that exchanges

$$\text{electrons} \longleftrightarrow \text{holes}$$

Here a hole means absence of an electron. On the bipartite square lattice, the particle-hole transformation is

$$c_{i\sigma} \to (-1)^i c_{i\sigma}^\dagger$$

where

$$(-1)^i = e^{i\vec{Q}\cdot\vec{R}_i} = \begin{cases} +1 & i \in A \\ -1 & i \in B \end{cases}.$$

The number operator transforms as

$$n_{i\sigma} \to 1 - n_{i\sigma}$$

So electrons become holes. The interaction term changes

$$U n_{i\uparrow} n_{i\downarrow} \to U n_{i\uparrow} n_{i\downarrow} - U n_i + U$$

and the chemical potential term changes

$$-\mu n_i \to -2\mu + \mu n_i$$

The hopping term is invariant

$$H_t \to H_t.$$

Combining the above transformed terms, the Hamiltonian becomes invariant if

$$\mu = \frac{1}{2} U.$$

This is the special half-filled point. At this point, the mean-field Hamiltonian becomes

$$\begin{aligned} H_{\mathrm{MF}} &= -t \sum_{\langle i,j\rangle,\sigma} (c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}) - \mu \sum_i n_i + U \sum_i \left( \frac{\rho_i}{2} n_i - \frac{m_i}{2}(n_{i\uparrow} - n_{i\downarrow}) - \frac{1}{4}(\rho_i^2 - m_i^2) \right) \\ &= -t \sum_{\langle i,j\rangle,\sigma} (c_{i\sigma}^\dagger c_{j\sigma} + c_{j\sigma}^\dagger c_{i\sigma}) - \frac{U}{2} m \sum_i (-1)^i (n_{i\uparrow} - n_{i\downarrow}) + N\frac{Um^2}{4}. \end{aligned}$$

This mean field Hamiltonian can be diagonalized in the momentum space. Consider the Fourier transform

$$c_{i\sigma} = \frac{1}{\sqrt{N}} \sum_k e^{ik\cdot R_i} \hat{c}_{k\sigma}, \quad c_{i\sigma}^\dagger = \frac{1}{\sqrt{N}} \sum_k e^{-ik\cdot R_i} \hat{c}_{k\sigma}^\dagger \ .$$

The hopping term can be written as

$$\sum_{k,\sigma} \varepsilon_k \hat{c}_{k\sigma}^\dagger \hat{c}_{k\sigma} \ , \qquad \varepsilon_k = -2t(\cos k_x + \cos k_y).$$

The mean field interaction

$$\sum_i (-1)^i (n_{i\uparrow} - n_{i\downarrow}) = \sum_{i,\sigma} \sigma e^{iQ\cdot R_i} c_{i\sigma}^\dagger c_{i\sigma} = \sum_{k,\sigma} \sigma \hat{c}_{k+Q,\sigma}^\dagger \hat{c}_{k\sigma}.$$

Thus

$$H_{\mathrm{MF}} = \sum_{k,\sigma} \varepsilon_k \hat{c}_{k\sigma}^\dagger \hat{c}_{k\sigma} - \Delta \sum_{k,\sigma} \sigma \hat{c}_{k+Q,\sigma}^\dagger \hat{c}_{k\sigma} + N\frac{\Delta^2}{U} \ , \qquad \Delta = \frac{Um}{2}.$$

The sum $k$ is over the dual site in the Brillouin zone

$$-\pi < k_x \leq \pi \ , \quad -\pi < k_y \leq \pi.$$

Because $k$ is coupled with $k+Q$ above, we let RBZ be the reduced Brillouin zone under the identification $k \sim k+Q$. Thus

$$H_{\mathrm{MF}} = \sum_{k\in \mathrm{RBZ},\sigma} \begin{pmatrix} \hat{c}_{k\sigma}^\dagger & \hat{c}_{k+Q,\sigma}^\dagger \end{pmatrix} \begin{pmatrix} \varepsilon_k & -\sigma\Delta \\ -\sigma\Delta & \varepsilon_{k+Q} \end{pmatrix} \begin{pmatrix} \hat{c}_{k\sigma} \\ \hat{c}_{k+Q,\sigma} \end{pmatrix} + N\frac{\Delta^2}{U}.$$

For the nearest-neighbor square lattice

$$\varepsilon_{k+Q} = -\varepsilon_k$$

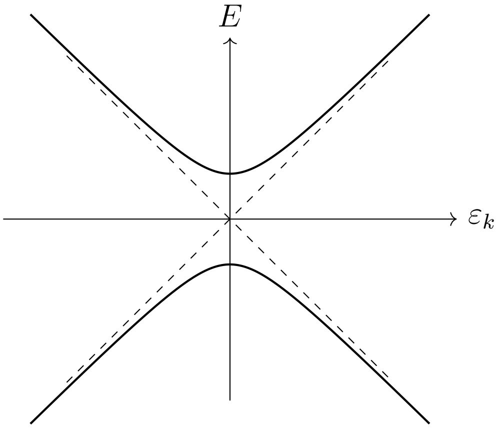
Hence

$$\begin{pmatrix} \varepsilon_k & -\sigma\Delta \\ -\sigma\Delta & \varepsilon_{k+Q} \end{pmatrix} = \begin{pmatrix} \varepsilon_k & -\sigma\Delta \\ -\sigma\Delta & -\varepsilon_k \end{pmatrix}.$$

whose eigenvalues are

$$\pm E_k = \pm\sqrt{\varepsilon_k^2 + \Delta^2}.$$

Thus the Fermi Surface becomes gapped, and anti-ferromagnetic order creates an insulating state. At half-filling and zero temperature, the lower band $-E_k$ is filled and the upper band $+E_k$ is empty.



Now we consider the self-consistency. The order parameter is

$$m = \frac{1}{N}\sum_i (-1)^i \langle n_{i\uparrow} - n_{i\downarrow}\rangle = \frac{1}{N}\sum_{k,\sigma} \sigma \left\langle \hat{c}^\dagger_{k+Q,\sigma}\hat{c}_{k\sigma}\right\rangle .$$

From the 2x2 Hamiltonian

$$\begin{pmatrix} \varepsilon_k & -\sigma\Delta \\ -\sigma\Delta & -\varepsilon_k \end{pmatrix}$$

we find

$$\left\langle \hat{c}^\dagger_{k+Q,\sigma}\hat{c}_{k\sigma}\right\rangle = \frac{\sigma\Delta}{2E_k}\left[f(-E_k) - f(E_k)\right]$$

where

$$f(E) = \frac{1}{e^{\beta E}+1}$$

is the Fermi function. At $T \to 0$, $\beta \to \infty$

$$\left\langle \hat{c}^\dagger_{k+Q,\sigma}\hat{c}_{k\sigma}\right\rangle \to \frac{\sigma\Delta}{2E_k}.$$

Thus at zero temperature

$$m = \frac{1}{N}\sum_{k\in\mathrm{RBZ},\sigma}\frac{\sigma\Delta}{E_k} = \frac{2\Delta}{N}\sum_{k\in\mathrm{RBZ}}\frac{1}{E_k}$$

But $\Delta = \frac{Um}{2}$. For $m \neq 0$, we find

$$1 = \frac{U}{N}\sum_{k\in\mathrm{RBZ}}\frac{1}{\sqrt{\varepsilon_k^2+\Delta^2}}.$$

This is the Hubbard anti-ferromagnetic mean-field gap equation. This equation is directly analogous to the BCS gap equation, but the gap is an anti-ferromagnetic particle-hole gap rather than a superconducting particle-particle gap.

# Bibliography

[1] Asano, Taro. *Theorems on the partition functions of the Heisenberg ferromagnets.* Journal of the Physical Society of Japan 29.2 (1970): 350-359.

[2] Auerbach, Assa. *Interacting electrons and quantum magnetism.* Springer Science & Business Media, 2012.

[3] Baxter, Rodney J. *Exactly solved models in statistical mechanics.* Integrable systems in statistical mechanics. 1985. 5-63.

[4] Benettin, Giancarlo, Giovanni Gallavotti, G. Jona-Lasinio, and A. L. Stella. *On the Onsager-Yang-Value of the spontaneous magnetization.* Communications in Mathematical Physics 30, no. 1 (1973): 45-54.

[5] Bhattacharjee, Somendra M., and Avinash Khare. *Fifty years of the exact solution of the two-dimensional Ising model by Onsager.* Current science 69.10 (1995): 816-821.

[6] Callen, Herbert B. *Thermodynamics and an Introduction to Thermostatistics.* John Wiley & Sons 2 (1993).

[7] Cardy, John. *Scaling and renormalization in statistical physics.* Vol. 5. Cambridge university press, 1996.

[8] Chari, Vyjayanthi, and Andrew N. Pressley. *A guide to quantum groups.* Cambridge university press, 1995.

[9] Deguchi, Tetsuo. *Introduction to solvable lattice models in statistical and mathematical physics.* Classical and Quantum Nonlinear Integrable Systems. CRC Press, 2019. 107-146.

[10] Ellis, Richard S. *Entropy, large deviations, and statistical mechanics.* Springer Science & Business Media, 2012.

[11] Faddeev, L. D. *How algebraic Bethe ansatz works for integrable model.* arXiv preprint hep-th/9605187 (1996).

[12] Fetter, Alexander L., and John Dirk Walecka. *Quantum theory of many-particle systems.* Courier Corporation, 2012.

[13] Fisher, Michael E. *The free energy of a macroscopic system.* Archive for Rational Mechanics and Analysis 17.5 (1964): 377-410.

[14] Friedli, Sacha, and Yvan Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction.* Cambridge University Press, 2017.

[15] Gibbs, Josiah Willard. *Elementary principles in statistical mechanics: developed with especial reference to the rational foundations of thermodynamics.* C. Scribner's sons, 1902.

[16] Goldenfeld, Nigel. *Lectures on phase transitions and the renormalization group.* CRC Press, 2018.

[17] Griffiths, Robert B. *Rigorous results and theorems.* Phase transitions and critical phenomena 1 (1972): 7-109.

[18] Huang, Kerson. *Introduction to statistical physics.* Chapman and Hall/CRC, 2009.

[19] Jaynes, Edwin T. *Information theory and statistical mechanics.* Physical review 106.4 (1957): 620.

[20] Jaynes, Edwin T. *Information theory and statistical mechanics. II.* Physical review 108.2 (1957): 171.

[21] Jaynes, Edwin T. *Probability theory: The logic of science.* Cambridge university press, 2003.

[22] Jimbo, Michio, and Tetsuji Miwa. *Algebraic analysis of solvable lattice models.* Vol. 85. American Mathematical Soc., 1994.

[23] Kaufman, Bruria. *Crystal statistics. II. Partition function evaluated by spinor analysis.* Physical Review 76.8 (1949): 1232.

[24] Kaufman, Bruria, and Lars Onsager. *Crystal statistics. III. Short-range order in a binary Ising lattice.* Physical Review 76.8 (1949): 1244.

[25] Lebowitz, Joel L., and Anders Martin-Löf. *On the uniqueness of the equilibrium state for Ising spin systems.* Communications in Mathematical Physics 25.4 (1972): 276-282.

[26] Lebowitz, Joel L., David Ruelle, and Eugene R. Speer. *Location of the Lee-Yang zeros and absence of phase transitions in some Ising spin systems.* Journal of mathematical physics 53.9 (2012).

[27] Lee, Tsung-Dao, and Chen-Ning Yang. *Statistical theory of equations of state and phase transitions. I. Theory of condensation.* Physical Review 87.3 (1952): 404.

[28] Lee, Tsung-Dao, and Chen-Ning Yang. *Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model.* Physical Review 87.3 (1952): 410.

[29] Li, Si. *Classical Mechanics and Geometry.* Boston, MA: International Press of Boston, 2023.

[30] Li, Si. *Electromagnetism and Geometry.* Boston, MA: International Press of Boston, 2023.

[31] Li, Si. *Quantum Mechanics and Geometry.* https://sili-math.github.io.

[32] McCoy, Barry M. Advanced statistical mechanics. Vol. 146. OUP Oxford, 2009.

[33] McCoy, Barry M., and Tai Tsun Wu. *The two-dimensional Ising model.* Courier Corporation, 2014.

[34] Montroll, Elliott W., Renfrey B. Potts, and John C. Ward. *Correlations and spontaneous magnetization of the two-dimensional ising model.* Journal of Mathematical Physics 4.2 (1963): 308-322.

[35] Mrugala, R. *Geometrical methods in thermodynamics.* Thermodynamics of energy conversion and transport 257 (2000).

[36] Onsager, Lars. *Crystal statistics. I. A two-dimensional model with an order-disorder transition.* Physical review 65.3-4 (1944): 117.

[37] Peierls, Rudolf. *On Ising's model of ferromagnetism.* Mathematical proceedings of the cambridge philosophical society. Vol. 32. No. 3. Cambridge University Press, 1936.

[38] Retore, Ana L. *Introduction to classical and quantum integrability.* Journal of Physics A: Mathematical and Theoretical 55.17 (2022): 173001.

[39] Ruelle, David. *Statistical Mechanics (Rigorous Results)* (1969). W.A. Benjamin, Inc., Amsterdam .

[40] Ruelle, David. *Extension of the Lee-Yang circle theorem.* Physical Review Letters 26.6 (1971): 303.

[41] Ruelle, David. *Zeros of graph-counting polynomials.* Communications in mathematical physics 200.1 (1999): 43-56.

[42] Ruppeiner, George. *Riemannian geometry in thermodynamic fluctuation theory.* Reviews of Modern Physics 67.3 (1995): 605.

[43] Šamaj, Ladislav, and Zoltán Bajnok. *Introduction to the statistical physics of integrable many-body systems.* Cambridge University Press, 2013.

[44] Shannon, Claude Elwood. *A mathematical theory of communication.* The Bell system technical journal 27.3 (1948): 379-423.

[45] Simon, Barry. *The Statistical Mechanics of Lattice Gases, Volume I.* Princeton University Press, 2014.

[46] Yang, Chen Ning. *The spontaneous magnetization of a two-dimensional Ising model.* Physical Review 85.5 (1952): 808.