The Ising Model - From inspiration to exact solutions

2019

Contents

Introduction

An Ising model consists of an array of points with binary states; each state can interact with others and is random. In this essay, I will justify the model, discuss how relevant quantities are related to its basic properties, and make precise quantitative predictions from it.

1 Physical inspiration for the Ising Model

Materials behave in many different ways when exposed to an external magnetic field. The measurable result of a material interacting with the field is a change in the overall magnetic field strength, meaning a second field is induced in the material; how the induced field varies with external field characterises the behaviour of the material. Reconciling with the rest of physics, we need to describe the induced field as arising from the collective behaviour of atoms; to allow this, we must speak in terms of a quantity which is[1] proportional to the induced magnetic field and associated with the same direction as the field, namely **magnetisation**.

The accepted model is as follows: each of the atoms[2] in a solid, for example, has a magnetic moment (a vector). Each moment interacts with the external magnetic field (a vector field) in a way results in a reorientation of the moment relative to the direction of the field near that atom. The pattern of atomic moments' orientations depends on the material and the external field, with the measurable affect being that a material may gain a net magnetic moment; this is its magnetisation. To rephrase the initial observation, a material acquires a magnetisation when an external field runs through it.

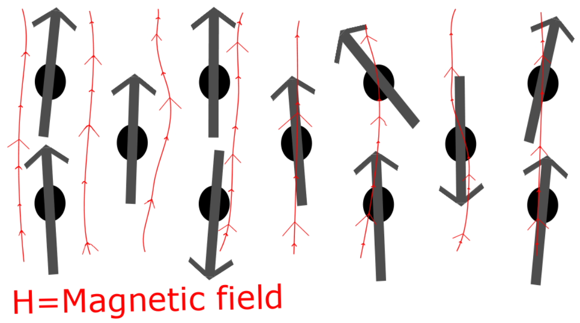

Figure 1: Atoms in a solid, represented by dots, each have a magnetic moment, represented by the arrows superimposed on each dot. The field lines of the external magnetic field **H** are shown in red.

In a **ferromagnetic material**, moments tend to align with the magnetic field, so they add to give a large magnetisation in the presence of a fairly homogeneous external field. A property which is most interesting about this class of materials is the behaviour of the magnetisation M as the external field $H \to 0$ at a certain temperature T.

The following is empirically derived [1, 3], and assumes a homogeneous magnetic field[3]: we see that M is an increasing, odd function of H, with behaviour shown by fig 2(a) at some temperatures. The curve changes as temperature is decreased, with the gradient $\frac{\partial M}{\partial H}$ near $H = 0$ increasing. There is a critical temperature T_c for which, as temperature $T \to T_c^+$, $\frac{\partial M}{\partial H}|_{H=0} \to \infty$, i.e the curve near $H = 0$ aligns ever closer to the M axis. In other words, as temperature falls while still above T_c, weaker magnetic fields induce stronger magnetisations.

Below T_c (fig 2(c)), as $H \to 0^+$ or $H \to 0^-$, the material is observed to retain a finite magnetisation; ferromagnetic materials can have an observable magnetic field without the support of

[1] Roughly and over-simplistically, yet representatively, this is the case

[2] These are not actually atoms in the average ferromagnet, but ions - it is not necessary to go into enough physical detail to be physically accurate here, and to call them atoms is not too misleading to the average young Mathematician

[3] This means the external field is a constant vector field, so we can describe the same phenomenon in terms of a scalar field of constant value H at every point

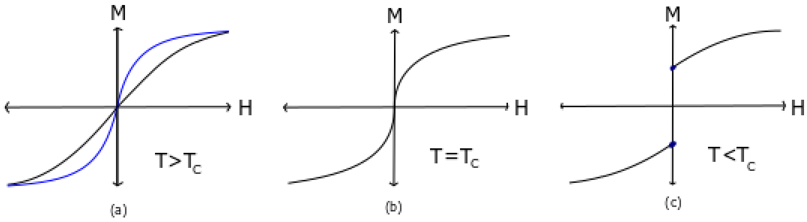

Figure 2: As $T \to T_c^+$ (a), the gradient near H=0 increases. The blue line corresponds to a temperature at T' with $T_c < T' < T$. Sketch (b) shows the limiting case as the gradient near $H = 0$ tends to infinity. On (c), the possible finite values of M at $H = 0$ are shown in blue.

an external field. We refer to the case of finite M with H arbitrarily close to 0 as **spontaneous magnetisation**.

The Ising Model, initially proposed[2] by Wilhelm Lenz in 1920, aims to explain and predict the behaviour of a ferromagnetic material.

2 Justifying the Ising Model

With its purpose in mind, some requirements of the Ising model at the stage of formulation are clear.

2.1 The Ising Model as a lattice of spins

The first requirement is that the model explicitly considers the behaviour of individual atoms, from which we must be able to infer the large scale behaviour of the material. Sadly, this could produce a very complex system, and the purpose of a model is to be an easy-to-analyse proxy for the real entity. A second requirement is that the model couples moments in close vicinity to each other, as in the corresponding physics, but also couples moments to the external field. To retain input from all atoms, the model strips the atoms to particles with simplifications of relevant factors, moment and position.

The Ising Model describes points on a **lattice**, corresponding to points in physical space. Were a formal treatment of lattices necessary, we would describe a *d-dimesional lattice* as a set $\mathfrak{I} \subset \mathbb{R}^d$ forming the abelian group $(\mathfrak{I}, +)$. This description hints that the points in physical space corresponding to a true lattice form regular patterns and grids

Example 2.1. 1D lattices will be especially relevant later; one such is the 'line' of points $\mathbb{Z}$. A second such is the closed 'circle' of N points $\mathbb{Z} \cap [0, N-1]$ (which can form a group with addition defined modulo N, see figure 4).

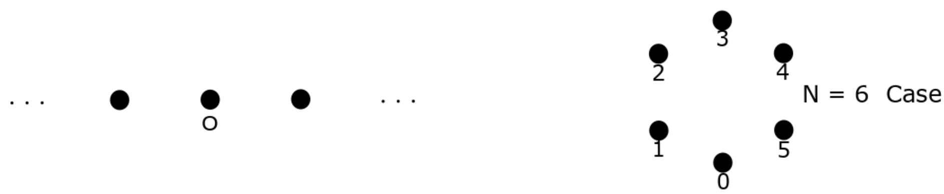

Figure 3: On the left is $\mathfrak{I} = \mathbb{Z}$. On the right is $\mathbb{Z} \cap [0, 5]$.

Though we will take this geometrically regular image henceforth, for the rest of this essay what we call a 'lattice' need not form a group. Critically, *atoms are modelled as having positions indexed by elements in the lattice* upon which the model is defined. Equivalently we can think of the elements as position vectors.

Appealing to the physical interpretation, we call the lattice elements **sites**.

With each site, we associate a quantity representing the moment. The characteristic simplification of an Ising Model of a material is the reduction of the number of degrees of freedom of the external magnetic field and magnetic moments; it assumes constant $H \in \mathbb{R}$, i.e H is a 1D vector, and that moment can only align with or against the field (a choice of two directions given the restriction on H). The quantity present at each site is called **spin**, which takes a value of 1 in one direction and -1 in the other.

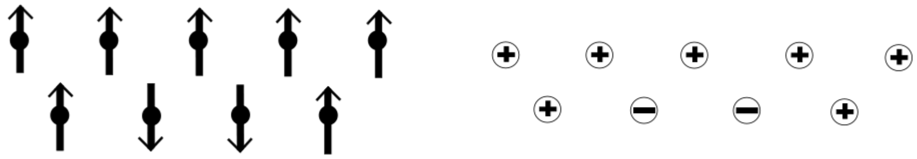

Figure 4: We denote the spins on an example lattice in 2 ways. On the left, up arrows indicate +1 spins, down arrows indicate -1 spins. On the left we simply use + and − markers (in the intuitive way), which is a more suitable representation for larger systems.

We preliminarily define the spin at a lattice site $\sigma_i \in \{-1, 1\}$ for site $i \in \mathfrak{I}$. The magnetisation, as the net moment, will be related to the sum of all spins in the model.

One more requirement is that the model takes into account the unpredictable, thermodynamic nature of atomic states. Magnetic moments will fluctuate rapidly over time, so in our model we can not predict the spin states with certainty. In an Ising model, spin is a random variable.

Though we can't usually list the spins at every site, the possible values of each spin are 1 and -1, so we know the possible states of the model system, called configurations, are ordered combinations of elements of $\{-1, 1\}$:

Definition 2.1. For an Ising Model defined on lattice $\mathfrak{I}$, we say $\omega\colon \mathfrak{I} \to \{-1, 1\}$ is the **configuration** of the model iff $\omega(i)$ is equal to the spin at site i for all $i \in \mathfrak{I}$. The **configuration space** is defined as

$$\Omega_{\mathfrak{I}} : = \{\omega : \omega \text{ is a function from } \mathfrak{I} \text{ into } \{-1, 1\}\}$$

To clarify, a configuration is simply an assignment of sites to spins, and the configuration space is the set of possible ways to do this. The configuration of the model shall be random, so $\Omega_{\mathfrak{I}}$ can be seen as the outcome space, upon which we can properly define spin:

Definition 2.2. For a model defined on $\mathfrak{I}$ in configuration $\omega \in \Omega_{\mathfrak{I}}$, the **spin** at site $i \in \mathfrak{I}$ is the function $\sigma_i\colon \Omega_{\mathfrak{I}} \to \{-1, 1\}$ defined as $\sigma_i(\omega) = \omega(i)$.

The magnetisation of our model needs to reconcile with the observable function of temperature and external field which it exists to mimic. Though configuration could vary for fixed field and temperature, the magnetisation must not.

Definition 2.3. The **magnetisation** of an Ising Model on finite lattice[4] $\mathfrak{I}$ is

$$M\colon = \overline{\sum_{i \in \mathfrak{I}} \sigma_i(\omega)}$$

$\overline{x}$ means the average of x over all configurations $\omega \in \Omega_{\mathfrak{I}}$.

For a particular model, $|M|$ is greater for conditions by which, on average, a greater majority of spins take the same value. We will later show how M is deducible given H, T, via the probability measure on $\Omega_{\mathfrak{I}}$.

We still need to ensure that an Ising Model causes neighbouring spins to couple, with a tendency to align with the field. The probability measure on $\Omega_{\mathfrak{I}}$ will ensure this, by giving a higher probability to configurations which align with H and positively correlate spins in close proximity; such configurations will have a low *energy*.

[4]Models defined on infinite lattices will be discussed later.

2.2 Energy

By now, the reader is aware that we treat spin at a site as a function of configuration; for the remainder of the essay, this will be implicit. We write σ_i as a spin at site $i \in \mathfrak{I}$ for configuration $\omega \in \Omega_{\mathfrak{I}}$.

In defining the interaction between spins, we think of the sites as vertices on a graph. One would draw a generic graph with vertices as points and edges as the line between them. Formally, a graph $G = (V, E)$ is an ordered pair with vertex set V, edge set E, where an edge between vertices $v, w \in V$ is recorded as $\{v, w\} \in E$. For a model defined on $\mathfrak{I}$, the *graph of interacting sites* has

$$V = \mathfrak{I}, \qquad E : = \{\{v, w\} : v, w \in V \text{ and } v, w \text{ are interacting sites}\}$$

We reiterate; edges denote interacting sites [4]. In this essay, we will be considering models in which, interpreting geometrically, only adjacent sites interact. E is called the **interaction edge set**.

When magnetic moments interact (with each other or the external field), they acquire an energy, as does the entire system. We define a corresponding energy of the model with the following in mind:

- The energy of the model should be the sum of the energies over all sites
- The strength of the interaction between spins, or between field and spin, should only depend on the spin values involved (not explicitly on the lattice element)
- The energy at site i will have a component caused by interaction with the magnetic field, $-H\sigma_i$ and a component $-\frac{1}{2}J\sigma_i\sigma_j$ (where J is a constant) from each interaction with a site $j \in \mathfrak{I}$ such that $\{i, j\} \in E$; *the interaction between two spins of the same sign should yield a lower energy than that between opposite-valued spins, also a spin aligning with external field (a spin with the same sign as H) should yield a lower energy than a spin aligned against it.*

Definition 2.4. Let E be the interaction edge set of an Ising Model defined on $\mathfrak{I}$.

$$\varepsilon(\omega) : = -J \sum_{\substack{i,j: \\ \{i,j\} \in E}} \sigma_i\sigma_j \quad - H \sum_{i \in \mathfrak{I}} \sigma_i$$

defines the **energy** of the model as a function of its configuration ω.

The first term in this definition considers the energy between all neighbouring spins. The coupling strength J is the same for each interacting pair; for Ferromagnetic materials, $J > 0$. The second term accounts for the energy arising from spins interacting with H.

2.3 Probability

The probability distribution function on $\Omega_{\mathfrak{I}}$ takes a form which can be justified by considering a method of 'statistical ensembles'. A rigorous argument is given in *The Two-dimensional Ising Model*[5].

It is time to link the model to its temperature T. If k_B is the Boltzmann constant, $\beta : = \frac{1}{k_B T}$ is the **inverse temperature**.

Assertion 2.1. *For an Ising Model defined on $\mathfrak{I}$ and at temperature at $T = \frac{1}{k_B \beta}$, let $Z : = \sum_{\omega' \in \Omega_{\mathfrak{I}}} e^{-\beta\varepsilon(\omega')}$. The probability that the model is in configuration ω is given by*[6]

$$\mathbb{P}(\omega) = \frac{e^{-\beta\varepsilon(\omega)}}{Z}$$

[5] Suppose we knew the energy of a finite, *isolated* model. By the fundamental postulate of statistical mechanics[8, pp 7], all configurations with the same energy would be equally likely to occur. Now suppose the model may exchange energy with a larger body, with which it is in thermal equilibrium, forming an 'ensemble'. We can apply the same principle to the energy, which must be constant with respect to time, of the combined system to derive the probability distribution (of the Ising model) which we are about to assert[6]. In this essay we ignore the physical relevance of a model's energy, discussing a model abstract from a restricted physical system.

[6] Having taken ST112, yes, as ω is an outcome we should treat $\mathbb{P}$ as a probability measure properly with probability given by $\mathbb{P}(\{\omega\})$, but we will omit curly braces for simplicity.

Definition 2.5. *Z as above is the* **partition function** *of that Ising Model*

In considering the probability of one configuration relative to the probability of another at some fixed temperature and external field, we can ignore Z. Using ω_{ref} as some fixed configuration, $\frac{\mathbb{P}(\omega)}{\mathbb{P}(\omega_{\text{ref}})} = \frac{e^{-\beta\varepsilon(\omega)}}{e^{-\beta\varepsilon(\omega_{\text{ref}})}}$. The 'relative' probability of each configuration ω is proportional to $e^{-\beta\varepsilon(\omega)}$, so *lower energy configurations are more likely to occur.* Considering the implications of spin-spin coupling and spin-field coupling on the energy, each spin is likely to align with the external field, and neighbouring spins are more likely to point in the same direction (in statistical terms, spins are correlated). *This is how the model couples neighbouring spins and spins with the external field.* The sum of the $e^{-\beta\varepsilon(\omega)}$ weights (associated with each $\omega \in \Omega_{\mathfrak{I}}$) over all possible configurations gives Z; the partition function is effectively the normalisation constant of the probability function in Assertion 2.1 (though many authors would hate to 'belittle' it).

Remark. The energy of the model in configuration $\omega \in \Omega_{\mathfrak{I}}$ is a function of ω, H, and β (we assumed J is constant across the model) just as the probability function must be. Z is dependent on β, H, but not on any specific configuration; Z is a function of all possible $\omega \in \Omega_{\mathfrak{I}}$, a set which doesn't vary for a model. We see that Z is a function of (H, β), or a function of (H, T) equally.

3 Magnetisation and the partition function

By now, we have fully defined the core of the Ising Model (this excludes potential *boundary conditions*, which we will return to). In summary, the Ising model consists of a lattice, each of whose elements is associated with a spin. The probability distribution of configurations of spin values is given in Assertion 2.1, thanks to which we can give a specific form to the magnetisation:

$$M = \sum_{\omega \in \Omega_{\mathfrak{I}}} \left[\mathbb{P}(\omega) \sum_{i \in \mathfrak{I}} \sigma_i \right] = \frac{\sum_{\omega \in \Omega_{\mathfrak{I}}} [e^{-\beta\varepsilon(\omega)} \sum_{i \in \mathfrak{I}} \sigma_i]}{Z} \tag{1}$$

where $\beta :\, = \frac{1}{k_B T}$. In this section, we check the conditions under which properties of M coincide with those in figure 2, revealing the relevance of the partition function Z.
The next three propositions are based on content in *'The Two-dimensional Ising Model'*:

Proposition 3.1. *The magnetisation of an Ising Model defined on a finite lattice is given by*

$$M :\, = \beta^{-1} \frac{\partial}{\partial H} \ln(Z)$$

for inverse temperature β, external field H.

Proof. Differentiating Z term by term, noting $Z = \sum_{\omega \in \Omega_{\mathfrak{I}}} e^{-\beta\varepsilon(\omega)}$,

$$\begin{aligned} \frac{\partial Z}{\partial H} &= \sum_{\omega \in \Omega_{\mathfrak{I}}} \frac{\partial}{\partial H} e^{-\beta\varepsilon} \\ &= \beta \sum_{\omega \in \Omega_{\mathfrak{I}}} \left[e^{-\beta\varepsilon(\omega)} \sum_{i \in \mathfrak{I}} \sigma_i \right] \end{aligned}$$

Thus, by the chain rule,

$$\begin{aligned} \frac{\partial}{\partial H} \ln(Z) = \frac{1}{Z} \frac{\partial Z}{\partial H} &= \beta \left(\frac{1}{Z} \sum_{\omega \in \Omega_{\mathfrak{I}}} \left[e^{-\beta\varepsilon(\omega)} \sum_{i \in \mathfrak{I}} \sigma_i \right] \right) \\ &= \beta M \qquad \text{by (1).} \end{aligned}$$

□

This will be shown to be a significant result; we set out to describe the magnetisation of a physical system, so we want to find M. Now, if we can calculate Z, we may be able to calculate M after all. As Z is a function of (H, T), M would depend on T (through β) and H. Noting this, we can immediately confirm that M has some of the properties of the observable we are modelling.

Proposition 3.2. *The magnetisation $M(H, T)$ of an Ising Model is such that:*

- $M(0, T) = 0$
- $M(-H, T) = -M(H, T)$ *i.e M is an odd function of H*
- $H > 0 \Rightarrow M(H, T) > 0$, *and* $H < 0 \Rightarrow M(H, T) < 0$

for all $T > 0$

The first two points are derivable by considering symmetry of the system under the transformation $\sigma_i \to -\sigma_i, H \to -H$. Separating the sum over $\omega \in \Omega_{\mathfrak{I}}$ in (1) into those terms giving positive and those terms giving negative values of $\sum_{i \in \mathfrak{I}} \sigma_i$ gives the third point [6, p.19].
Though $M(0, T) = 0$, this does not rule out spontaneous magnetisation due to the specificity of it definition:

Definition 3.1. The **spontaneous magnetisation** M_0 of model defined on a finite lattice is given by

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

The Thermodynamic limit

Proposition 3.3. *For a model defined on a finite lattice $\mathfrak{I}$, M is an infinitely smooth function of (H, T) for all $(H, T) \in \mathbb{R} \times \mathbb{R}^+$*

Proof. The number of configurations in $\mathfrak{I}$ is $2^{|\mathfrak{I}|}$, so Z is the sum of a finite number of terms of the form $e^{-\beta\varepsilon}$. The energy of each configuration ε is linear in H, while $\beta = \frac{1}{k_B T}$, so each term is infinitely smooth for all $(H, T) \in \mathbb{R} \times \mathbb{R}^+$, as is the case for Z.
$\ln(x)$ is infinitely smooth $\forall x \in \mathbb{R}^+$. By the chain rule, as $e^{-\beta\varepsilon} > 0$ for each configuration so $Z > 0$, $\ln(Z)$ is infinitely smooth $\forall (H, T) \in \mathbb{R} \times \mathbb{R}^+$. Thus, $M = k_B T \frac{\partial}{\partial H} \ln(Z)$ is infinitely smooth on $\mathbb{R} \times \mathbb{R}^+$. □

Consider figure 2(c) again. For a model defined on a finite lattice, M is an infinitely smooth function of H for $T > 0$, but we want an Ising Model to predict M with the discontinuity at $H = 0$ corresponding to spontaneous magnetisation when $T_c > T > 0$. For this to be possible, we need to take the model in the limit as $|\mathfrak{I}| \to \infty$. This is called the **thermodynamic limit**.

In hindsight, it doesn't seem too insensible that we would use models in the case of an arbitrarily large number of sites in this study of ferromagnetic materials. Macroscopic systems consist of around 10^{23} or more atoms, each of which we model with a site-spin pair. This number is beyond the threshold for accurately modelling a system as being in the thermodynamic limit[6, 7].

We can expect the magnetisation M to increase in proportion with $|\mathfrak{I}|$ in many cases[6, pp.22], so in the thermodynamic limit it is sensible to choose an alternative representation of how magnetised the material becomes. If we expect most of the spins to align with H most of the time, we expect the 'average' spin to align in the same direction for constant $H \neq 0$ most of the time.

Definition 3.2. For an Ising Model defined on finite lattice $\mathfrak{I}$ with magnetisation M, the **magnetisation per site** is given by

$$m(H, T) := \frac{1}{|\mathfrak{I}|} M(H, T)$$

For an infinite lattice,

$$m_\infty(H, T) := \lim_{|\mathfrak{I}| \to \infty} \frac{1}{|\mathfrak{I}|} M(H, T)$$

The particularities of how $|\mathfrak{I}|$ should go to infinity will be discussed later; the limit's existence and value can be shown to be largely independent of the nature of the sequence of finite lattices defining it. For now, we note that for $|\mathfrak{I}|$ finite, $|M| \leq |\mathfrak{I}|$, so $|m_\infty|, |m| \leq 1$; magnetisation per site is bounded in the thermodynamic limit, as desired. Having been successful here, we complete our translation to the limiting case by the next definition.

Definition 3.3. The **spontaneous magnetisation per site** is given by

$$m_0(H,T) : = \lim_{H \to 0^+} m(H,T)$$

or $m_{\infty_0}(H,T) : = \lim_{H \to 0^+} m_\infty(H,T)$ as appropriate.

4 Exact Solutions

4.1 Boundary conditions, interaction edge sets and lattice types

Suppose that we are producing an Ising model for which not all functions of the form $\omega\colon \mathfrak{I} \to \{-1,1\}$ are possible configurations. To express this, we might restrict the configuration space to a subset of $\Omega_{\mathfrak{I}}$. Restricting the configuration space like so is equivalent to imposing **boundary conditions**[7] (BCs).

To make precise models, we need to consider model specifications like BCs; another specification is the interaction edge set E, which we haven't restricted other than stating that only neighbouring sites (from the perspective of sites as points in space) should be able to interact. Different ferromagnetic materials have their atoms arranged in different patterns; the lattice's 'geometry' is significant as this determines which sites we would call neighbours. In summary, BCs, edge sets, particularly though lattice geometry, could affect the behaviour of the model, possibly in the thermodynamic limit.

Many of the lattice shapes we see in nature have very regular structures, so much that they allow the following simplification; for the remainder of the essay we will treat lattices as if they were of the form $\mathfrak{I} \subseteq \mathbb{Z}^d$. Some structures would be harder to visualise like this, but this is not an issue in our computation.

Example 4.1. The line of points from example 2.1 can form a model with *free boundary condition*; in this case, $E = \{\{i,j\} : i,j \in \mathbb{Z} = \mathfrak{I}, |i-j| = 1\}$. Only neighbouring sites interact, but the model's configurations are unrestricted on $\mathfrak{I}$.

Example 4.2. We will focus on the *periodic boundary condition* in particular. Let $\mathfrak{I} \subseteq \mathbb{Z}^d$ be a finite lattice subset for some $d \in \mathbb{N}_+$. We write $i \in \mathfrak{I}$ as $i = (i_1, \dots, i_d)$ with $i_j \in \mathbb{N} \cap [1,n]$ for $j \in \mathbb{N} \cap [1,d]$, period $n \in \mathbb{N}$. For simplicity, we assume the same period in each component of i; the BC requires

$$\begin{aligned} E =& \{\{i,j\} : i,j \in \mathfrak{I}, i_m = j_m \,\forall m \text{ except at some } m' \in \mathbb{N} \cap [1,d], \\ & |i_{m'} - j_{m'}| = 1 \text{ mod } n\} \end{aligned}$$

For $d = 1$,

$$E = \{\{i,j\} : i,j \in \mathfrak{I}, |i-j| = 1 \text{ mod } n\}$$

In such a case, setting $\sigma_{n+1} = \sigma_1$ as per the BC,

$$\varepsilon(\omega) = -J \sum_{i=1}^{n} \sigma_i \sigma_{i+1} \quad - H \sum_{i=1}^{n} \sigma_i$$

With the position vector interpretation of $\mathfrak{I}$, the $d = 1$ periodic lattice is best visualised as a ring of points like that mentioned in example 2.1.

[7] The classes of BCs which are physically relevant are those such that, for the restriction $\Omega_{\mathfrak{I},R}$ of $\Omega_{\mathfrak{I}}$, applying this class of restriction to a sequence of increasingly large subsets, $\frac{|\Omega_{\mathfrak{I}}| - |\Omega_{\mathfrak{I},R}|}{|\Omega_{\mathfrak{I}}|} \to 0$ as $|\mathfrak{I}| \to \infty$. We think of only such cases when we think of BCs [4]

We are going to make use of specific BCs, but we need not dwell on their variety. It is theoretically verified that the value of $m_\infty(H,T)$ is not dependent on BCs, nor the way in which $|\mathfrak{I}| \to \infty$. For an infinite lattice $\mathfrak{I}_\infty$, $(\mathfrak{I}_k)_{k\in\mathbb{N}}$ a sequence of finite lattices, we say $\lim_{k\to\infty} \mathfrak{I}_k = \mathfrak{I}_\infty$ iff $\forall \mathfrak{I} \subset \mathfrak{I}_\infty$ with $\mathfrak{I}$ finite, $\exists N \in \mathbb{N}$ such that $k > N \Rightarrow \mathfrak{I}_k \subseteq \mathfrak{I}$. Think of $\mathfrak{I}_k \to \mathfrak{I}_\infty$ as shorthand for this.

The following theorem is adapted from '*Statistical mechanics of lattice systems*'.

Theorem 4.1. *The magnetisation per site m_∞ on infinite lattice $\mathfrak{I}_\infty \subseteq \mathbb{Z}^d$ for some $d \in \mathbb{N}_+$ is well-defined, independent of the sequence $(\mathfrak{I}_k)_{k\in\mathbb{N}}$ of finite subsets of $\mathfrak{I}_\infty :=\lim_{k\to\infty} \mathfrak{I}_k$, and independent of BCs on $\mathfrak{I}_k$ for $H \in \mathbb{R} \setminus S_\beta$, where $S_\beta \subseteq \mathbb{R}$ is dependent on β and is countable. Furthermore, $\forall H \in \mathbb{R} \setminus S_\beta$,*

$$m_\infty(H,T) = \frac{\partial}{\partial H} \lim_{k\to\infty} \frac{1}{|\mathfrak{I}_k|} \ln(Z_{\mathfrak{I}_k})$$

($Z_{\mathfrak{I}_k} = Z$ on $\mathfrak{I}_k$)

Note that this theorem also gives us some freedom to swap limiting processes, which we can use later to compute $m_\infty(H,T)$.

Even excluding S_β from the possible values of H, we can still consider a function on $\mathbb{R}$ excluding a countable number of points, which is dense at all points in $\mathbb{R}$. The practical ramifications are minimal. For $d = 1$, we can assume convenient BCs, *giving the same magnetisation per site for each $H \in \mathbb{R} \setminus S_\beta$, $T > 0$ as for any model defined on an infinite 1D lattice.* In the $d = 1$ case, $S_\beta = \emptyset$,[4] leading to an unexpected smoothness of m_∞.

4.2 Exact solution in 1D via the transfer matrix method

Consider the $d = 1$ periodic boundary condition model as described in example 4.2. Our aim is to calculate Z for a finite lattice from which we can deduce m_∞. The method we employ is that of the transfer matrix[8]; we can split the energy of a configuration into the sum of terms involving a pair of neighbouring spins. Let $|\mathfrak{I}| = n \geq 3$:

$$\varepsilon(\omega) = -J \sum_{i=1}^{n} \sigma_i\sigma_{i+1} \quad - H\sum_{i=1}^{n} \sigma_i = -\sum_{i=1}^{n} \left[J\sigma_i\sigma_{i+1} + \frac{1}{2}H(\sigma_i + \sigma_{i+1})\right] \tag{2}$$

Therefore

$$\begin{aligned} Z &= \sum_{\omega\in\Omega_{\mathfrak{I}}} \exp[-\beta\varepsilon(\omega)] \qquad = \sum_{\omega\in\Omega_{\mathfrak{I}}} \exp \sum_{i=1}^{n} \left[\beta J\sigma_i\sigma_{i+1} + \frac{1}{2}\beta H(\sigma_i + \sigma_{i+1})\right] \\ &= \sum_{\omega\in\Omega_{\mathfrak{I}}} \prod_{i=1}^{n} \exp\left[\beta J\sigma_i\sigma_{i+1} + \frac{1}{2}\beta H(\sigma_i + \sigma_{i+1})\right] \\ &= \sum_{\omega\in\Omega_{\mathfrak{I}}} \prod_{i=1}^{n} V(\sigma_i, \sigma_{i+1}) \end{aligned}$$

where $V : \{-1,1\}^2 \to \mathbb{R}$ is defined by $V(\sigma_i,\sigma_j) := e^{[\beta J\sigma_i\sigma_j + \frac{1}{2}\beta H(\sigma_i+\sigma_j)]}$. The domain of V naturally inspires a representation of V by a 2×2 matrix; the **transfer matrix** of this model is

$$\mathbf{V} := \begin{pmatrix} V(1,1) & V(1,-1) \\ V(-1,1) & V(-1,-1) \end{pmatrix}$$

The clever part of this standard argument is writing $V(\sigma_i,\sigma_{i+1})$ in terms of $\mathbf{V}$, a representation which will clarify how we can decompose the sum over $\omega \in \Omega_{\mathfrak{I}}$ usefully. For generality of case to be

[8]The strategy we follow in this chapter for determining Z is not attributed to any one source as it is so widely known [1, 6, 4, 7]

represented, we introduce the notation

$$|\sigma_i\rangle = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \sigma_i = 1 \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \sigma_i = -1 \end{cases}$$

as well as[9] $\langle \sigma_i| = |\sigma_i\rangle^T$. In this way, we can write

$$V(\sigma_i, \sigma_j) = \langle \sigma_i|\mathbf{V}|\sigma_j\rangle$$

For example, if $\sigma_i = 1, \sigma_j = -1$,

$$\langle \sigma_i|\mathbf{V}|\sigma_j\rangle = \langle\, 1\,|\mathbf{V}| - 1\rangle = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} V(1,1) & V(1,-1) \\ V(-1,1) & V(-1,-1) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = V(1,-1)$$

Now we write

$$Z = \sum_{\omega\in\Omega_{\mathfrak{J}}} \prod_{i=1}^{n} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle$$

Let $\mathrm{Tr}(\mathbf{A})$ denote the trace of matrix $\mathbf{A}$, which is the sum of its diagonal elements.

Lemma 4.2. *$Z = \mathrm{Tr}(\boldsymbol{V}^n)$ for all $n \in \mathbb{N}_{\geq 3}$*

Proof. We will consider expanding, in the previous representation of Z, the sum over all configuration into cases of $\sigma_n = 1$ and cases of $\sigma_n = -1$:

$$Z = \sum_{\substack{\omega\in\Omega_{\mathfrak{J}}:\\ \omega(n)=1}} \prod_{i=1}^{n} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle \quad + \quad \sum_{\substack{\omega\in\Omega_{\mathfrak{J}}:\\ \omega(n)=-1}} \prod_{i=1}^{n} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle$$

$$Z = \sum_{\substack{\omega\in\Omega_{\mathfrak{J}}:\\ \omega(n)=1}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}|\sigma_n\rangle\langle \sigma_n|\mathbf{V}|\sigma_1\rangle$$

$$+ \sum_{\substack{\omega\in\Omega_{\mathfrak{J}}:\\ \omega(n)=-1}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}|\sigma_n\rangle\langle \sigma_n|\mathbf{V}|\sigma_1\rangle$$

In each sum, we are summing over distinct ordered combinations of $\sigma_1, \ldots, \sigma_{n-1}$ only, so in each we can assert the value of σ_n and restrict the sum to being equivalently over $\omega \in \Omega_{\mathfrak{J}\setminus\{n\}}$.

$$Z = \sum_{\omega\in\Omega_{\mathfrak{J}\setminus\{n\}}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}|1\rangle\langle 1|\mathbf{V}|\sigma_1\rangle$$

$$+ \sum_{\omega\in\Omega_{\mathfrak{J}\setminus\{n\}}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}| - 1\rangle\langle -1|\mathbf{V}|\sigma_1\rangle$$

$$= \sum_{\omega\in\Omega_{\mathfrak{J}\setminus\{n\}}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}\Big(|1\rangle\langle 1| + |-1\rangle\langle -1|\Big)\mathbf{V}|\sigma_1\rangle$$

$$= \sum_{\omega\in\Omega_{\mathfrak{J}\setminus\{n\}}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)\mathbf{V}|\sigma_1\rangle$$

$$= \sum_{\omega\in\Omega_{\mathfrak{J}\setminus\{n\}}} \left[\prod_{i=1}^{n-2} \langle \sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle \sigma_{n-1}|\mathbf{V}^2|\sigma_1\rangle$$

[9]Wherever T appears as a matrix exponent, it means transpose, not temperature

We can repeat this process, splitting the sum into cases of $\omega(j) = 1$ and of $\omega(j) = -1$, restricting each new sum to being over all $\omega \in \Omega_{\mathfrak{I}\setminus\{n,\ldots,j+1,j\}}$, then recombining for all $j \in \mathbb{N} \cap [3, n]$. The result at each stage is

$$Z = \sum_{\omega\in\Omega_{\mathfrak{I}\setminus\{n,\ldots,j+1,j\}}} \left[\prod_{i=1}^{j-2} \langle\sigma_i|\mathbf{V}|\sigma_{i+1}\rangle\right] \langle\sigma_{j-1}|\mathbf{V}^{n-j+2}|\sigma_1\rangle$$

$j = 3$ gives

$$\begin{aligned} Z = &\sum_{\omega\in\Omega_{\{1,2\}}} \langle\sigma_1|\mathbf{V}|\sigma_2\rangle\langle\sigma_2|\mathbf{V}^{n-1}|\sigma_1\rangle \quad = \sum_{\omega\in\Omega_{\{1\}}} \langle\sigma_1|\mathbf{V}^n|\sigma_1\rangle \\ =&\langle 1|\mathbf{V}^n|1\rangle + \langle -1|\mathbf{V}^n|-1\rangle \\ =&\mathrm{Tr}(\mathbf{V}^n) \end{aligned}$$

□

Because $\mathbf{V}$ is a symmetric matrix, we can evaluate $\mathrm{Tr}(\mathbf{V}^n)$ cleanly by diagonalising it; we know $\exists \mathbf{P} \in \mathbb{R}^{2,2}$ which is orthogonal such that $\mathbf{P}^T\mathbf{V}\mathbf{P}$ is diagonal. We can write $\mathbf{P}^T\mathbf{V}\mathbf{P} = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}$ for some $\lambda_+, \lambda_- \in \mathbb{R}$ with $\lambda_+ \geq \lambda_-$. It is easy to show that, for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{2,2}$, $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$, given which[5]

$$\begin{aligned} \mathrm{Tr}(\mathbf{V}^n) =&\mathrm{Tr}(\mathbf{V}^n\mathbf{P}\mathbf{P}^T) && \text{as } \mathbf{P}\mathbf{P}^T = \mathbf{I}_{2,2} \\ =&\mathrm{Tr}(\mathbf{P}^T\mathbf{V}^n\mathbf{P}) \\ =&\mathrm{Tr}(\left[\mathbf{P}^T\mathbf{V}\mathbf{P}\right]^n) = \mathrm{Tr}\begin{pmatrix} \lambda_+^n & 0 \\ 0 & \lambda_-^n \end{pmatrix} \end{aligned}$$

Therefore $Z = \mathrm{Tr}(\mathbf{V}^n) = \lambda_+^n + \lambda_-^n$.

λ_+, λ_- are the eigenvalues of $\mathbf{P}^T\mathbf{V}\mathbf{P}$, which are also the eigenvalues of the similar matrix $\mathbf{V}$, so to compute Z for this model all we need are the eigenvalues of

$$\mathbf{V} = \begin{pmatrix} V(1,1) & V(1,-1) \\ V(-1,1) & V(-1,-1) \end{pmatrix} = \begin{pmatrix} e^{\beta(J+H)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-H)} \end{pmatrix}$$

The characteristic equation can be deduced as

$$C_{\mathbf{V}}(\lambda) = \lambda^2 - 2e^{\beta J}\cosh(\beta H)\lambda + 2\sinh(2\beta J)$$

giving $\lambda_\pm = e^{\beta J}\cosh(\beta H) \pm e^{\beta J}\sqrt{\sinh^2(\beta H) + e^{-4\beta J}}$.

Let us summarise what we have achieved:

Theorem 4.3. *Let $\mathfrak{I} = \mathbb{N} \cap [1, n]$. For a model defined on $\mathfrak{I}$ with periodic boundary condition, energy given by equation (2), the partition function is*

$$Z = \lambda_+^n + \lambda_-^n$$

where $\lambda_\pm = e^{\beta J}\cosh(\beta H) \pm e^{\beta J}\sqrt{\sinh^2(\beta H) + e^{-4\beta J}}$.

Determining the exact form of the magnetisation per site in the limit as $n \to \infty$ is straightforward, other than being computationally taxing, using proposition 3.1, theorem 4.1 and theorem 4.3. A typical technique[6, 7] begins by recognising that, as λ_+ is strictly greater than λ_- for fixed $J > 0$, $\left(\frac{\lambda_-}{\lambda_+}\right)^n \to 0$ as $n \to \infty$ for fixed $\beta > 0$. Writing $Z = \lambda_+^n\left[1 + \left(\frac{\lambda_-}{\lambda_+}\right)^n\right]$ makes use of this to calculate $\frac{\ln Z}{n}$ in the limit as $n \to \infty$.

I quote one simplification of m_∞ here[6, pp.38]:

$$m_\infty = \frac{\sinh(\beta H)}{\sqrt{\sinh^2(\beta H) + e^{-4\beta J}}} \tag{3}$$

We have come 'full circle'; it is time to consider the physical interpretation of what the model has predicted. In doing so, we compare the behaviour of m_∞ with what we expected of the magnetisation by the end of chapter 1. m_∞ is an odd function of H, as should be. We can also show that m_∞ is an increasing function of H.

Let $c \in \mathbb{R}^+$ be a constant. With variable $x \in \mathbb{R}_{\geq 0}$,

$$\begin{aligned} \sqrt{x^2 + c} &= \sqrt{(x + a(x))^2} && \text{for some } a\colon \mathbb{R}_{\geq 0} \to \mathbb{R}^+ \\ &= x + a(x) \end{aligned}$$

We choose a such that $c = 2xa(x) + [a(x)]^2$. If the value of x increases, to ensure this equality holds, $a(x)$ decreases; a is a strictly decreasing function with respect to x. Furthermore, as $x \to \infty$, $a(x) \to 0$.

$$\begin{aligned} \sqrt{x^2 + c} - x &= a(x) > 0 \\ \frac{x}{\sqrt{x^2 + c}} &= 1 - \frac{a(x)}{\sqrt{x^2 + c}} > 0 \end{aligned}$$

Thus, $f(x) = \frac{x}{\sqrt{x^2+c}}$ is an increasing function for $x > 0$, with $\sup_{x \in \mathbb{R}^+} f(x) = 1$. As f is odd with respect to x, f is an increasing function for all $x \leq 0$ too.

Let $x = \sinh(\beta H)$; if we fix $\beta > 0$, then x increases iff H increases. Setting $c = e^{-4\beta J}$, one can show that $m_\infty = \frac{\sinh(\beta H)}{\sqrt{\sinh^2(\beta H) + c}}$ is an increasing function of H using what we have shown for f, with $\sup_{H \in \mathbb{R}} m_\infty = 1$, $\inf_{H \in \mathbb{R}} m_\infty = -1$.

One critical behaviour is missing. For finite β, m_∞ is still an infinitely smooth function! In the last paragraph's argument, we could have fixed $H \neq 0$ and stated x increases iff β increases. This gives

$$\lim_{\beta \to \infty} m_\infty(H, T) = \begin{cases} 1 & H > 0 \\ 0 & H = 0 \\ -1 & H < 0 \end{cases}$$

Only when $T = 0\,(\beta = \infty)$ does m_∞ behave discontinuously in H, i.e the model predicts $T_c = 0$ for 1D models, in which case spontaneous magnetisation does not occur at an attainable temperature.

5 Conclusion - Peierls' argument

The Ising model bears the name of Ernst Ising. In 1925, while a PhD student of Lenz, Ising's article calculating the partition function in 1 dimension was published; this was the first exact computation of a thermodynamically significant quantity held by the model. Contrary to his expectation, the result he achieved was in line with the conclusion we reached in the previous chapter; no spontaneous magnetisation was predicted at non-zero temperature.

By theorem 4.1, magnetisation per site given by equation (3) holds for all relevant boundary conditions, so all 1 dimensional ($d = 1$) Ising models fail to produce spontaneous magnetisation. Ising's assumption was that, in higher dimensions, the same failure would persist. This was shown not to be true. In 1936, Rudolf Peierls made the case for the $H = 0$ scenario of a $d = 2$ model, via what has come to be known as Peierls' argument. The following (up to the next 'Remark') is based on the rendition of Peierls' argument from '*Statistical mechanics of lattice systems*'.

Let $H = 0$ for a model on $\mathfrak{I} = (\mathbb{Z} \cap [-n, n])^2, n \geq 1$, with a *positive boundary condition*; we restrict $\Omega_{\mathfrak{I}}$ to

$$\Omega_{\mathfrak{I}}^+ := \{\omega \in \Omega_{\mathfrak{I}} : \omega(\pm n, a), \omega(a, \pm n) = 1 \text{ for all } a \in \mathbb{Z} \cap [-n, n]\}$$

The interaction edge set is $E = \{\{i, j\} : i, j \in \mathfrak{I}, |i - j| = 1\}$, identically to that of the free boundary condition for this $\mathfrak{I}$ (see fig 5).

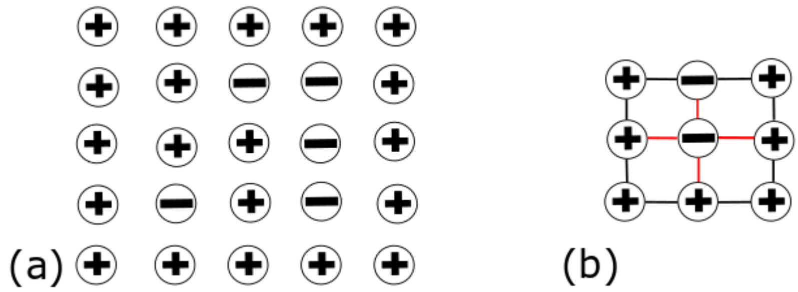

Figure 5: The $n = 2$ case of the 'square' model we're considering is shown in (a). A subset of the vertices and the edges between them are shown in (b), with the edges involving the central site highlighted in red. Boundary sites interact with fewer than 4 other sites.

The tactic is showing that the central site $(0, 0) \in \mathfrak{I}$ achieves an average value > 0 for some positive temperature; in the thermodynamic limit this average is equal to the spontaneous magnetisation per site[4] m_{∞_0}, so if we can prove this we can prove the presence of spontaneous magnetisation for $d = 2$. Let σ_0 denote the value of the spin at $(0, 0)$. As the cases $\sigma_0 = 1, \sigma_0 = -1$ are mutually exclusive and together exhaustive,

$$\begin{aligned} \overline{\sigma_0} = \sum_{\substack{\omega \in \Omega_{\mathfrak{I}}^+ : \\ \omega(0,0)=1}} \mathbb{P}(\omega) - \sum_{\substack{\omega \in \Omega_{\mathfrak{I}}^+ : \\ \omega(0,0)=-1}} \mathbb{P}(\omega) &= \mathbb{P}(\omega : \sigma_0 = 1) - \mathbb{P}(\omega : \sigma_0 = -1) \\ &= 1 - 2\,\mathbb{P}(\omega : \sigma_0 = -1) \end{aligned}$$

(By $\mathbb{P}(\omega : \sigma_0 = -1)$, we mean the probability that $\sigma_0 = -1$)

Suppose we could bound $\mathbb{P}(\omega : \sigma_0 = -1)$ by a function of $\beta = \frac{1}{k_B T}$, namely δ with $\delta(\beta) \to 0^+$ as $\beta \to \infty$. Then we have $\mathbb{P}(\omega : \sigma_0 = -1) \leq \delta(\beta)$, so $\overline{\sigma_0} \geq 1 - 2\delta(\beta)$. Then $\exists \alpha \in \mathbb{R}^+ : \beta > \alpha \Rightarrow \delta(\beta) < \frac{1}{2}$, i.e $T < \frac{1}{k_B \alpha} \Rightarrow \overline{\sigma_0} > 0$. If we prove such $\delta(\beta)$ exists, we prove that the model produces spontaneous magnetisation at a positive temperature.

The characteristic aspect of Peierls' argument is the way we arrive at $\delta(\beta)$. A lattice can be divided into regions of -1 spins and regions of $+1$ spins under a fixed configuration. We denote the set of closed curves which are boundaries of these regions Γ. Each $\gamma \in \Gamma$ is the union of lines of length 1 (of some unit), and each encloses only -1 spins in its interior (each γ lies on a regular rectangular grid with 90 degree angles between grid lines, as in figure 6). Note Γ is dependent on configuration $\omega \in \Omega_{\mathfrak{I}}^+$, so is random.

First, $\mathbb{P}(\omega : \sigma_0 = -1) \leq K$, where K is the probability of some $\gamma \in \Gamma$ enclosing $(0, 0)$. We can further bound K by functions dependent on possible $\gamma \in \Gamma$ enclosing $(0, 0)$, the length $|\gamma|$ of each, and β; Γ specifies the configuration enough to relate each $|\gamma|$ to its energy contribution and determine the configuration energy, so we can relate the probability of the boundary set Γ to a function of β and of its members' lengths. Combinatorial reasoning[10] from here produces a potential $\delta(\beta)$ which remains well defined in the limit as $n \to \infty$, with the properties initially desired. For $d = 2$ there is a critical temperature $T_c > 0$.

Remark. We could have employed the same argument but with a negative boundary condition and found $\overline{\sigma_0} < 0$ for some $T > 0$. This means m_{∞_0} takes different values depending on the BC, so in the notation of theorem 4.1, $S_\beta \ni 0$ for some $\beta \in \mathbb{R}^+$.

[10]Though one of the essences of Peierls' argument, for the sake of essay concision the details are omitted here.

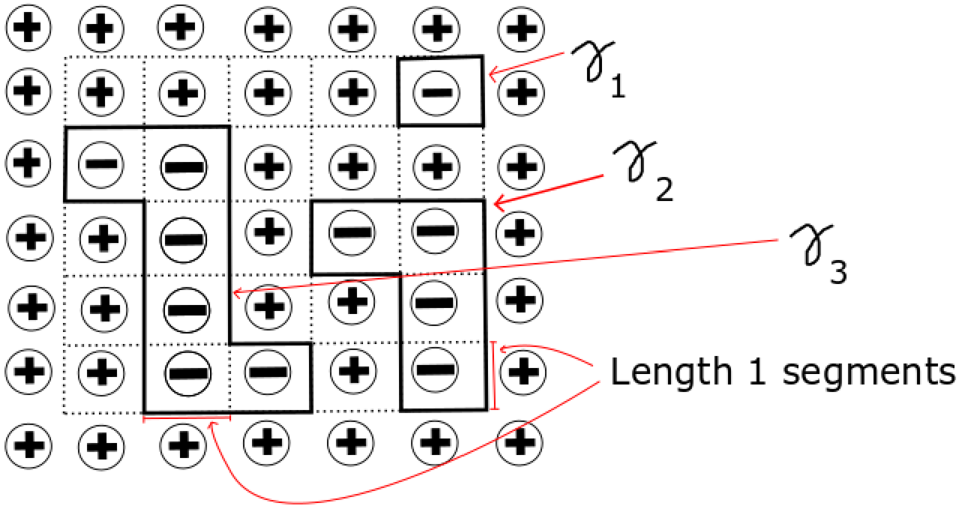

Figure 6: An example of a configuration of $\mathfrak{I}$ with $n = 3$ and the resultant boundary curve set Γ illustrated; $\Gamma = \{\gamma_1, \gamma_2, \gamma_3\}$ with $|\gamma_1| = 4$, $|\gamma_2| = 10$, $|\gamma_3| = 14$.

Since Peierls' contribution, the success of the Ising model has been steadily realised, for example by Crammers and Wannier in 1941 who introduced the transfer matrix technique to calculate T_c for $d = 2$. Other novel behaviours of the model for $d = 2$ have been discovered and studied, such as magnetic hysteresis.[6]

An exact analytic solution for $d = 3$ remains an open problem in mathematical physics.

References

[1] Rodney J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press limited, 1982.

[2] Stephen G Brush. History of the lenz-ising model. *Reviews of modern physics*, 39(4):883, 1967.

[3] Barry A Cipra. An introduction to the ising model. *The American Mathematical Monthly*, 94 (10):937–959, 1987.

[4] Sacha Friedli and Yvan Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction.* Cambridge University Press, 2017.

[5] Israel N Herstein. *Topics in algebra.* John Wiley & Sons, 1975.

[6] Barry M McCoy and Tai Tsun Wu. *The two-dimensional Ising model.* Harvard Univesity Press, 1973.

[7] Giuseppe Mussardo. *Statistical field theory: an introduction to exactly solved models in statistical physics.* Oxford University Press, 2010.

[8] Ramamurti Shankar. *Quantum Field Theory and Condensed Matter: An Introduction.* Cambridge University Press, 2017.

6 Notation

H	External field strength
T	(Absolute) Temperature
$\mathfrak{I}$	Lattice subset of an Ising Model of $\lvert\mathfrak{I}\rvert$ sites
$\overline{x}$	The average of x over all configurations $\omega \in \Omega_{\mathfrak{I}}$
σ_i	The value of the spin at the site i, implicitly for configuration ω
k_B	The Boltzmann constant
β	Inverse temperature; $\beta := \frac{1}{k_B T}$
Z	The partition function (a function of H, β)
$\mathrm{Tr}(\mathbf{A})$	The Trace of matrix $\mathbf{A}$, which is the sum of the diagonal elements