

ISSYP 2022 Mentoring

What is magnetism and what does it tell us about phases of
matter?

Session 1 - Probability and Statistics

July 25, 2022

Abstract

In Session 1, we present an introduction to concepts in probability and statistics, for use in statistical physics. We discuss random variables and distributions that these variables follow. The concept of expectation values, and important statistical notions such as the mean and variance of quantities are analysed, and we follow several examples to illuminate motivation for calculations. Given this framework, we are prepared to discuss applications within statistical physics, and relate to examples such as energy and ideal gases in chemical systems. Finally, we discuss analytical and numerical approaches to dealing with difficult distributions and calculations.

Contents

1	Random Variables and Probability Distributions	1
1.1	Random Variables	1
1.2	Probability Distributions	1
1.3	Multiple Random Variables	2
1.4	Examples of Distributions	3
2	Expectation Values	5
2.1	Functions of Random Variables	5
2.2	Mean and Variance	6
3	Statistical Physics	7
3.1	Building up Statistical Models	7
3.2	Use of Fundamental Constants	7
4	Obstacles	9
4.1	Choosing a Distribution	9
4.2	Normalizing Distributions	9
5	Summary	11

1 Random Variables and Probability Distributions

How are statistics and physics related? There is a whole field of physics called *statistical-physics*, or *statistical-mechanics*, which describes materials and systems in terms of random variables and their expectation values. We could think of performing an experiment of measuring a quantity, such as the heat capacity of a material, or how much magnetic field it produces. This quantity will vary depending on the exact configuration of all of the parts of the material, like all of the vibrating atoms and molecules, but will have some expected value, with some variance, or error in the measurement depending on how much the configurations of the material fluctuate.

1.1 Random Variables

The concept of randomness in probability can be described using a so-called random variable X . This random variable can take any possible value x , also known as an *outcome*, within a set $\mathcal{X} = \{x^{(0)}, x^{(1)}, \dots\}$. We say that X follows a distribution

$$X \sim p(X), \tag{1}$$

meaning the random variable takes a value $X = x$ from this set with probability $p(x)$.

What are some examples of random variables in real-life?

How could we estimate the probability distribution of a quantity of interest?

1.2 Probability Distributions

The notion of probability can be understood in terms of an experiment of observing the random variable X , N different times. The probability of a specific value x from the possible set of values \mathcal{X} is equal to the number of times $X = x$ is observed, in the limit of an infinite number N of observations:

$$p(x) = \lim_{N \rightarrow \infty} \frac{\# \text{ times } X = x}{N} . \tag{2}$$

What are some constraints on the values $p(x)$ that all distributions should always obey, no matter the random variable?

Probability distributions have two necessary conditions:

1. *An outcome from the set of possible outcomes must always be observed*: The total probability of observing one of the possible outcomes must be equal to 1 (or 100% probability)

$$\sum_{x \in \mathcal{X}} p(x) = 1 \quad (3)$$

If a distribution satisfies this condition, we say it is *normalized*, otherwise we must *renormalize*: $p(x) \rightarrow p(x) / \sum_{x \in \mathcal{X}} p(x)$.

2. *Each outcome must have bounded, non-negative probability*: For all possible x from the set of possible values \mathcal{X} , the probabilities must be bounded as

$$0 \leq p(x) \leq 1 \quad : \quad x \in \mathcal{X} \quad (4)$$

Intuitively, it does not make sense for the probability of an outcome to be greater than 100%, or for there to be negative probabilities.

1.3 Multiple Random Variables

Instead of just considering one random variable X , we can consider a group of N random variables

$$(X_1, X_2, \dots, X_N) \sim p(X_1, X_2, \dots, X_N) , \quad (5)$$

that follows a *joint distribution* describing the probability of observing a specific value for each variable *component* with subscript index $i = 1, 2, \dots, N$

$$X_i = x_i \quad (6)$$

within its set of allowed possible values

$$x_i \in \mathcal{X}_i . \quad (7)$$

The variables within this group may be *correlated*, meaning the probability of observing a certain $X_i = x_i$, is *dependent* on observing a different variable $X_j = x_j$.

If all variables are *independent*, then their distributions are all separate, $X_i \sim p_i(X_i)$, and the probability of observing a specific outcome of an individual variable is completely independent of observing an outcome of any other random variable. The joint distribution for independent random variables can be written as the separate product of individual distributions

$$p_{\text{independent}}(x_1, x_2, \dots, x_N) = p_1(x_1)p_2(x_2) \dots p_N(x_N) \quad (8)$$

Why is this an appropriate definition of independence of variables?

To simplify notations, we will usually denote groups of variables by a single variable that is assumed to have *components*

$$X = (X_1, X_2, \dots, X_N) \quad (9)$$

where the set of all possible multiple variables is

$$x \in \mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_N) . \quad (10)$$

we will further denote functions of multiple random variables as

$$f(X) = f(X_1, X_2, \dots, X_N) \quad (11)$$

and denote summations over all possible components as

$$\sum_{x \in \mathcal{X}} f(x) = \sum_{x_1 \in \mathcal{X}_1} \sum_{x_2 \in \mathcal{X}_2} \cdots \sum_{x_N \in \mathcal{X}_N} f(x_1, x_2, \dots, x_N) . \quad (12)$$

1.4 Examples of Distributions

There are many examples of random variables and probability distributions:

1. *Uniform distribution*: If the set of possible values has size N , then all values are equally likely, with constant probability

$$p_{\text{uniform}}(x) = 1/N \quad : \quad x \in \mathcal{X} \quad (13)$$

For example, dice have 6, equally likely sides to land on, each side landing has probability $1/6$.

2. *Bernoulli distribution*: If the set of possible values has size 2, $\mathcal{X} = \{0, 1\}$, and the probabilities are

$$\begin{aligned} p_{\text{bernoulli}}(0) &= p \\ p_{\text{bernoulli}}(1) &= 1 - p \end{aligned} \quad (14)$$

For example, a coin that is more likely to be heads 0, than tails 1, could be described with a Bernoulli distribution. The distribution of observing k 0's from n Bernoulli trials is

$$p_{\text{bernoulli}}(\#0's = k) = \binom{n}{k} p^k (1 - p)^{n-k} . \quad (15)$$

How could we derive this?

3. *Gaussian distribution*: If the allowed values are any real value $\mathcal{X} = \mathbb{R}$, and follow the *bell-curve*, or *normal* distribution

$$p_{\text{gaussian}}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \quad (16)$$

For example, many observations in nature and society follow the normal distribution.

4. *Boltzmann distribution*: If the allowed values are any set of values \mathcal{X} , and follow the exponential distribution with a characterizing function $\varepsilon(x)$

$$p_{\text{boltzmann}}(x) = \frac{1}{\sum_{x \in \mathcal{X}} e^{-\varepsilon(x)}} e^{-\varepsilon(x)} \quad (17)$$

For example, many observations in physics, particularly related to the energy and entropy of a system, follow a Boltzmann distribution.

Each of these distributions can be generalized to the *multiple variable* case, with specific modifications depending on how the component variables are *correlated*.

What other example distributions can you think of? Either real-life examples of the distributions above, or new examples?

What do you think the function $\varepsilon(x)$ might represent for Boltzmann distributions?

2 Expectation Values

Random variables and distributions can be fairly abstract, and difficult to get a sense of what the random outcomes really mean. The most intuitive, and descriptive quantities are *expectation values*, which can be thought of as the value of a quantity that is observed on average, if it is distributed according to a certain distribution.

We denote expectation values of a random variable $X \sim p(X)$ as

$$\langle X \rangle = \sum_{x \in \mathcal{X}} p(x) x, \quad (18)$$

which can be seen as a sum of the possible values of X , with each value *weighted* by how likely it is, according to the distribution $p(x)$.

What are some useful properties of expectation values? For example, what is the expectation value of a constant $\langle \alpha \rangle$, or a constant times a random variable $\langle \alpha X \rangle$?

Due to probability distribution being normalized, the expectation values of constants are

What is the expectation value of the uniformly distributed random variable X from the set $\mathcal{X} = \{1, 2, \dots, n\}$?

2.1 Functions of Random Variables

Given we have a random variable X , functions of these variables $f(X)$ for $x \in \mathcal{X}$ are like any other function, and the distribution of the outcomes of this function are distributed similarly to the underlying distribution of $X \sim p(X)$.

How could we calculate the expectation value of functions of random variables?

There is the so-called *law of the unconscious statistician*, shown through a *change of variables*, that is a very useful formula for all functions of variables with known distributions:

$$\langle f(X) \rangle = \sum_{x \in \mathcal{X}} p(x) f(x). \quad (19)$$

which can be seen as a sum of the possible values of $f(x)$, with each value *weighted* by how likely the argument $X = x$ is, according to the distribution $p(x)$.

Therefore, we don't need to figure out the, often complicated, distribution for the values $f(X)$, but only require knowing the distribution for X .

What are some useful properties of functions of expectation values? For example, what is the expectation value of weighted sums of functions $\langle \sum_i \alpha_i f_i(X) \rangle$?

What is the expectation value the function of a sum of multiple independent variables $f(X_1, X_2, \dots, X_N) = \sum_i^N f_i(X_i)$?

2.2 Mean and Variance

Two of the most important expectation values that characterize a distribution are the so-called mean and variance, generally denoted respectively as

1. *Mean:*

$$\mu_X = \langle X \rangle = \sum_{x \in \mathfrak{X}} p(x) x \quad (20)$$

which describes the value of a random variable that is expected to be observed on average.

2. *Variance:*

$$\sigma_X^2 = \langle (X - \mu_X)^2 \rangle = \sum_{x \in \mathfrak{X}} p(x) (x - \mu_X)^2 \quad (21)$$

which describes how much a value of a random variable is likely to differ from the expected value.

How can we write the variance in the simplified form $\sigma_X^2 = \langle X^2 \rangle - \mu_X^2$?

For some distributions, if we know the mean and variance, this is sufficient to describe the entire distribution. For example, the *Gaussian* distribution can be written with a mean and variance parameter

$$p_{\text{gaussian}}(x|\mu_X, \sigma_X^2) = \frac{1}{\sqrt{2\pi\sigma_X^2}} e^{-\frac{1}{2\sigma_X^2}(x-\mu_X)^2}. \quad (22)$$

When do mean and variance come up as important quantities in real-life?

3 Statistical Physics

Almost all of the materials and matter around us are behaving randomly at some level, with their constituents being described by random variables that follow some complicated distributions. However, we can build up from simple *models*, descriptions of how these constituents are changing, and interacting with each other, to form simple probability distributions for the behaviours. These distributions can then be used to estimate expectation values, to get a qualitative and somewhat quantitative description for the average behaviours of these complicated materials if we were to measure a material many times in an experiment.

3.1 Building up Statistical Models

These simplified models can be chosen to describe certain aspects of material's behaviour well, and then they can be combined, and knowledge of different results from different models can be accumulated to produce more sophisticated, and accurate models and distributions.

As an example, let us think about ideal gases in chemistry. These gases consist of many (Avogadro's number) particles, all moving somewhat *randomly* around their container. It would be very difficult to observe, or calculate the exact position and configuration of every particle, however we can describe the *average* behaviour of the gas as a whole using statistics. This average behaviour can be described by a few important quantities, such as the volume V of the container that the particles are occupying, the average temperature T of the particles, and the average pressure P the particles are exerting on their container. If the number of particles in the gas N can be estimated, then the ideal gas law $PV = Nk_bT$, where k_b is Boltzmann's constant, can be used to estimate this average pressure or temperature of the gas, without knowing anything about the exact configuration of the gas particles.

This ideal gas law relationship can be derived using solely principles from statistical physics! We will not derive this here, but using calculus, and some concepts from these sessions, it is very accessible. We could then work on developing a more accurate model for the temperature of ideal gases, or the pressure and volume dependence. Combining these models will yield a more precise description. For example, this ideal gas equation can be expanded by taking into account the particles having a non-zero size, resulting in the improved Van der Waals gas equation $(P + aN^2/V^2)(V - bN) = Nk_bT$ for some coefficients a, b that depend on the specific gas particles being considered.

3.2 Use of Fundamental Constants

To simplify expressions, we will use a common convention in physics, to use units of quantities such that particular fundamental constants have value of 1, meaning we do not have to explicitly include them in expressions. If we want to revert back to conventional (i.e) SI units, we just have to perform unit analysis to restore where the constants should appear in expressions.

What are the units and right and left hand sides of the idea gas law $PV = Nk_bT$?

Here, we will set the frequently occurring Boltzmann's constant equal to 1

$$k_b = 1 \tag{23}$$

and so temperatures and energies can now be thought to have the same units!

4 Obstacles

So far, we have seen a formal introduction to some concepts in probability and statistics, and how they can be used in principle to obtain average quantities and behaviours of complicated probabilistic materials and systems. However, we have not discussed the many challenges with taking this statistical approach.

4.1 Choosing a Distribution

How do we even know which distribution a quantity should follow? From theoretical approaches from basic first principles of physics, we can often derive the distributions that certain quantities, or material behaviours should follow. Performing experiments to collect data of observations also gives insight into which distributions are being followed. Fortuitously, by assuming that the materials are in certain environments, for example at a fixed temperature, or within a fixed volume, there are only a few distributions that seem to occur often in nature. The primary distribution for how much *energy* a material possesses, a concept to be discussed in future sessions, is given by some form of a *Boltzmann* distribution. Given some underlying properties of a material are random variables $\sigma \in \Xi$ that can take any allowed value in the set Ξ , and assuming some relationship of the energy as a function of these properties ε_σ and that the material is at a constant temperature T , the Boltzmann probability of a material having a certain energy is

$$p_\sigma = \frac{1}{Z} e^{-\frac{\varepsilon_\sigma}{T}} \quad (24)$$

where the normalization factor $Z = \sum_{\sigma \in \Xi} e^{-\varepsilon_\sigma/T}$ is a sum over all configurations, and is entitled, for historical reasons, the *partition function*.

What does the Boltzmann distribution say about which energetic configuration are most likely at different temperatures?

How does the energy of a particular configuration of particles relate to the concepts of microstates and macrostates

4.2 Normalizing Distributions

We have discussed that probability distributions must be normalized with this partition function factor Z , however have not discussed how to feasibly calculate this summation for complicated distributions p_σ , and a large number of possible values Ξ . Essentially all of statistics, using some calculus, can be described using only this one quantity Z ! There are many analytical approaches to computing these summations exactly, which will be discussed and introduced over future sessions, and much of statistical physics is concerned with finding improved summation methods.

Finally, even if we approach computing these normalization factors using computational methods, there are many technical challenges. If we are trying to sum over all possible

configurations of many variables, attempting to perform these summations can take hours, days, or even years in principle to compute on today's fastest computers. Alternatively to trying to perform the summations exactly, we may attempt to obtain enough samples of these random variables such that our set of samples adequately reflects the distribution they follow, and use these samples to estimate other quantities. To accurately conduct this sampling procedure, entitled *Monte Carlo* methods, also requires significant computational time and resources, and has many other feasibility issues.

How would we conceptually come up with an approach to efficiently perform some summations?

5 Summary

In this session, we have discussed some important concepts in statistics of random variables, and probability distributions. We have described how to calculate expectation values of quantities, and highlighted the mean and variance as central notions. We have gone on to discuss how physics and statistics are intricately related, and how we may start to form these connections. Finally, we have discussed a general approach of starting with as simple a description of a physical system as possible, and slowly accumulating knowledge of the system with more and more complicated models.

In the next session, we will be discussing what even are phases of matter, and how can we start to describe materials and phenomena around us using mathematics and physical models.

Are there any additional questions or comments?

What is everyone taking away from this session?

What would everyone like to learn more about regarding this session?

How was the level of content and delivery of this session?

ISSYP 2022 Mentoring

What is magnetism and what does it tell us about phases of matter?

Session 2 - Phases of Matter and Lattice Models

July 26, 2022

Abstract

In Session 2, we discuss the concept of phases of matter in materials around us, and discuss similarities and differences between both different phases, and different materials which exhibit different types of phases of matter. We then discuss how we could approach modelling these phases of matter using mathematics, and the language of statistical physics. We finally introduce an important playground for modelling materials, and investigating phases of matter, that of discrete lattices of sites, connected by bonds. These lattices offer an intuitive background for proposing statistical models, and we discuss examples of lattices, and how they can be used to visualize interactions between constituents in a material.

Contents

1	What are Phases of Matter?	1
1.1	Distinguishing Phases of Matter	1
1.2	Phase Transitions	1
2	Models of Materials	3
2.1	Calculations based on Models	3
3	Concept of Energy in a System	4
3.1	Interaction Energies	4
4	Lattices	5
4.1	Performing summations on Lattices	5
4.2	Homogeneity in Lattices	6
5	Lattice Models of Magnetism	7
5.1	Ising Model of Magnetism	7
6	Summary	9

1 What are Phases of Matter?

What do we think of when we talk about phases of matter?

There are many materials around us, all made up of basic building block constituents, depending how far we zoom in, they could be molecules, or atoms, or even sub-atomic particles. Each material has some unique characteristics, but we can further categorize them by what *phase* they are in, and there are usually *universal* characteristics of each phase across a wide range of materials. An important direction in physics is understanding how to classify these phases of matter, and determining how to best characterize these universal phenomena.

1.1 Distinguishing Phases of Matter

What distinguishes phases of matter? What could we measure to determine which phase a material is in?

What are some universal characteristics of solid, liquid, and gas phases across different materials? What distinguishes the materials?

1.2 Phase Transitions

Arguably more important than identifying phases of matter, is understanding how they *transition* from one phase, and what environmental factors *drive* these transitions.

What property of the environment forces water to be in one of its phases?

How could we write mathematically the presence of a phase transition in a particular quantity?

Phases of matter, and phase transitions are somewhat intuitive, however are notoriously difficult to describe mathematically. We will now present several aspects of modelling materials that will help towards obtaining a complete physical and mathematical description of phases of materials.

Why are phase transitions difficult to describe mathematically?

Phase transitions arise when there is an *abrupt* change in some important quantity called an *order-parameter* (i.e) the *density* and configuration of particles in water), when certain independent parameters exceeds certain *critical values* (i.e) the freezing or boiling *temperature* of water). Ideally, these abrupt changes of quantities result in very different macroscopic properties of the material, that we can identify as being in very different phases (i.e) liquid and solid water have very different properties).

We can model this mathematically by first finding a suitable quantity of interest, that from our physical intuition, appears to possibly change as another parameter changes past a

certain point, and then study the *dependence* of these variables. Most behaviours in physics and materials tend to change *smoothly* (i.e) plastics slowly deform when stressed, and don't spontaneously change to some fundamentally different material), and so these abrupt changes can be difficult to conceptualize and describe mathematically.

For complicated systems, it can be difficult and not at all obvious whether there are phases and phase transitions, and which quantities are most appropriate to indicate these phase transitions. Fortunately, as we will discuss throughout these sessions, many materials behave very similarly at a fundamental level, and there are underlying phenomena that cause the existence of similar phase transitions across these different materials.

We have seen that the *Temperature*, and its scale relative to the energy of a system affects the Boltzmann distribution for how likely a material is to have a certain amount of energy. Temperature is in fact often the driving factor that causes phase transitions at a certain *critical-temperature*.

2 Models of Materials

As physicists, we must develop what are known as *models* of materials. These models are mathematical descriptions for how basic constituents of a material, what physicists call *particles*, evolve over time, and interact with each other. In order to start developing these models, and approaches for extracting important properties about the material using calculations, we must first develop some intuition behind how these particles may *interact*, how we can describe the *environment* the particles occupy, and some governing *principles* that dictate how these particles evolve and reach some kind of equilibrium over time.

What are different interactions between particles that you can think of? Hint: Think about electromagnetic and gravitational forces

Imagine we have a particle that can be in one of two states: $\sigma = \pm 1$ and $\Xi = \{\pm 1\}$, a so-called two-level system. How could we represent this state mathematically? How could we write down an interaction between two of these particles?

Imagine we have now a set of N of these particles $\sigma = \{\sigma_i = \pm 1 : i = 1 \dots N\}$, and $\Xi = (\Xi_1, \Xi_2, \dots, \Xi_N)$, where $\Xi_i = \{\pm 1\}$. What would be some intuitive and simple ways of arranging these particles, that intuitively shows and constrains which particles interact most with which (Hint: Nearby) particles?

2.1 Calculations based on Models

When we talk of forming mathematical models of materials, we must also talk about performing *calculations* with these models. Calculating generally refers to extracting useful, intuitive, and *experimentally-relevant* quantities about the materials being described. Some examples could be the heat capacity: how much heat is required to raise a material's temperature, or its bulk modulus: how much its volume changes when pressure is applied, or its magnetic susceptibility: how much magnetic field it gives off when exposed to a magnetic field. One way of calculating these quantities, is to assume that the system has some randomness, and can be in different configurations σ that have an *energy* ε_σ . When it is at a fixed temperature T in equilibrium with its environment, it follows the *Boltzmann* distribution

$$p_\sigma = \frac{1}{Z} e^{-\frac{\varepsilon_\sigma}{T}} \quad (1)$$

where $Z = \sum_{\sigma \in \Xi} e^{-\frac{\varepsilon_\sigma}{T}}$ is the probability normalization factor as the sum over all possible configurations, or (micro)states that the system can be in. To calculate f_σ , that are assumed to be functions of which state the system is in, this amounts to estimating the *expectation value* of these quantities, over this distribution of possible energetic states

$$\langle f \rangle = \sum_{\sigma \in \Xi} p_\sigma f_\sigma \quad (2)$$

How could we write $\langle f \rangle$ as a function of solely the partition function?

3 Concept of Energy in a System

The concept of energy can be abstract, and what do we mean by the energy $\varepsilon = \varepsilon_\sigma$ of a system for a configuration σ ? We have probably seen that materials with mass m can have kinetic energy from being in motion with speed v , $\varepsilon_v = v^2/2m$, or gravitational potential energy from being at a height z , $\varepsilon_z = mgz$, or electromagnetic energy from being attracted or repelled from a particle with charge e a distance r away, $\varepsilon_r = -e/4\pi\epsilon_0 r$. When concerned with the dynamics, or kinematics of a system, we can think of the *forces* acting on particles, causing their motion, or we can equally think about the energy that particles have as they move around, and eventually lose, or gain energy, depending how they collide, or interact with other particles.

How would we describe the total energy of a system of N such particles, each with the energies above?

Energy is essentially a *local* quantity, meaning it can generally be written as a sum over the energy ε_i of each separate particle i , which includes energy that a particle has, plus energy that it possesses from interacting with other particles. Energy can thus be written as a sum over each of N particles

$$\varepsilon_\sigma = \sum_i^N \varepsilon_{\sigma_i} . \quad (3)$$

What governing principle do systems with a defined energy obey? Hint: Do systems want to be in high energy, or low energy states?

Due to the Boltzmann equation having the form of an exponential of the energy, how are the probabilities affected if all energies are shifted by a constant energy value?

3.1 Interaction Energies

Given we can describe the energies, of individual particles, we must now supplement our energy description with the energy representing how particles interact. By the principle of energy minimization, interactions that are favoured, such as gravitational attraction between massive particles, will *decrease* the energy, whereas interactions that are not favoured, such as electromagnetic attraction between equally charged particles, will *increase* the energy.

4 Lattices

Given these notions of the states that particles can be in, the interactions between particles, and this concept of energies of a system of particles, the last important aspect of describing models of materials, is the environment that these particles are occupying. A convenient description when dealing with fixed particles, in nice, uniform systems, are grids of sites where these particles live, called *lattices*.

What are lattices? How can we picture them, and describe them mathematically?

Lattices can be described by a set of sites $\Lambda = \{i : i = 1 \cdots N\}$, connected by bonds $\Delta = \{\delta = (i, j) : i, j \in \Lambda\}$. For example, we could have a lattice that is *symmetrical*, or *translationally invariant*, with all sites having the same *coordination number*, or number of nearest-neighbour bonds z . Examples of these lattice, are the *square*, *triangular*, or *hexagonal* lattices.

How many nearest-neighbours does each of the square, triangular, or hexagonal lattices have?

4.1 Performing summations on Lattices

We often come across summations over the sites and bonds of a lattice in models of interactions between particles that are arranged on a lattice. Performing the correct combinatorics of counting these numbers of sites and bonds, and keeping track of the symmetries present in the lattice is an important skill.

For example, if we are summing over all sites in a lattice, we often denote this for functions of a site f_i as

$$\sum_i f_i , \quad (4)$$

whereas if we are summing over all nearest-neighbour bonds in a lattice, we often denote this for functions of a bond f_{ij} as

$$\sum_{\langle ij \rangle} f_{ij} , \quad (5)$$

and we can also denote this as a summation over all lattice sites, and each sites' nearest-neighbours

$$\sum_{\langle ij \rangle} f_{ij} = \frac{1}{2} \sum_i \sum_{\langle ji \rangle} f_{ij} , \quad (6)$$

and denote all nearest neighbours of site i as $\langle ji \rangle$, or just $\langle i \rangle$.

Why is there the factor of 1/2 when initially summing over all lattice sites, then nearest-

neighbours?

How might we perform this summation where functions f_i are of sites,

$$\sum_{\langle ij \rangle} f_i + f_j \tag{7}$$

for a lattice with N sites, and constant coordination number z ?

4.2 Homogeneity in Lattices

Depending on the system that is being modelling on the lattice, and the quantities being calculated, the system may be *homogenous*, meaning all sites are essentially identical, and a particle at one site is *indistinguishable* from a particle at another site. We can also discuss taking what is known as the *thermodynamic limit* of the lattice, where the number of particles

$$N \rightarrow \infty \tag{8}$$

to represent realistic systems with at least an Avogadro's number of particles.

Lattices can also be in spaces of different numbers of *spatial dimension* d . We live in $d = 3$ dimensions, but could also look at a $d = 1$ dimensional *chain* of sites, or a $d = 2$ dimensional *plane* of sites, or even a more abstract $d > 3$ dimensional *hyper-lattice* of sites.

How does the dimensionality of the lattice affect the coordination number? Enumerate the nearest-neighbours in d dimensions for a hyper-cubic lattice.

5 Lattice Models of Magnetism

Now that we have a sense of some interactions between some 2-level particles, and the lattices that these particles may live on, how may we define a model that represents this system? This concept of energy ε , and the principle of energy minimization, is very general, and offers an intuitive framework for modelling the present state of a system by how much energy it possesses. Given a system will tend towards its lowest energy configuration, this energetic description indicates how the system will evolve towards this state. One important method of modelling a system therefore consists of determining an accurate, but as simple as possible, mathematical description of the energy of a system of particles. We also want to develop models that are as *general* as possible, and serve to describe as many materials and systems as possible, so as to determine the most fundamental underlying phenomena that govern many materials, and to avoid having to develop a new model each time we encounter a new material.

5.1 Ising Model of Magnetism

We will start, perhaps not intuitively, with the most fundamental model of magnetism, the *Ising* model. It will seem incredibly simple at first, but throughout our analysis, we will slowly begin to see its subtleties, and how all of its features, and intuition that we gain from it, will show up constantly in any more sophisticated models that we subsequently analyse. Furthermore, even though it is initially posed as a model for magnetism, the fact it is a 2-level system, means other systems where particles can be in one of two states, can be translated, or *mapped*, into a form that resembles the Ising model. This concept of *mapping* one description of a system, usually one that appears very complicated, and difficult to perform calculations with, to a related system that we already have intuition, or calculated results about, is an essential part of physical modelling as well.

Let us now define the Ising particles $\sigma_i = \pm 1$ as random variables on a lattice of N sites. We can think of each site as being a mini bar-magnet, either pointing up or down.

Each particle only interacts with its z *nearest-neighbours*, with a *ferromagnetic* interaction, meaning the system wants to align all of its nearest-neighbours to point along the same direction, this being the lowest energy state. There also may be an external magnetic field, that the particles will want to *align* themselves with.

How may we write this model?

We can thus write the energy of the Ising model for a given configuration $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ of N particles on a lattice as

$$\varepsilon_\sigma = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (9)$$

where J represents how strongly the particles will tend to align with each other, and h represents the strength of the external magnetic field that forces the particles to align along

its direction.

What is the local energy of each particle in the Ising model?

Notice, that we have not indicated what kind of lattice the particles of this model live on, or which number of *spatial dimensions* we are in. In fact this model is very general, and only requires us to know for each particle and site i , its z nearest-neighbours $\langle ij \rangle$, and so could potentially model magnetism of particles on any lattice we choose.

Does this model make sense? Why are there minus-signs? How would we write this model for a $d = 1$ dimensional chain of Ising particles?

What is the ground-state of the ferromagnetic Ising model on a triangular lattice?

What if there were no minus signs? What would be the ground-state of the anti-ferromagnetic Ising model on a triangular lattice?

We can think of each Ising particle interacting with each of its nearest-neighbours, and this external field, and please note the *minus-signs*, to represent that it is *energetically-favourable* for the particles to align along these interactions.

Given this model of energy in an Ising model, how does this relate to our previous discussion of the Boltzmann distribution, and most energetically favourable states?

Is there any reason to think there could be different phases for particles following this Ising model?

What could be possible quantities that drive any phase transitions, and what quantities could indicate which phase the particles are in?

For example, we could calculate the average magnetization of each particle

$$\langle \sigma \rangle = \frac{1}{N} \sum_i \left(\sum_{\sigma \in \Xi} p_{\sigma} \sigma_i \right), \quad (10)$$

or the average energy of a lattice per particle

$$\langle \varepsilon \rangle = \frac{1}{N} \sum_i \left(\sum_{\sigma \in \Xi} p_{\sigma} \varepsilon_{\sigma_i} \right), \quad (11)$$

We will see how these quantities arise when determining whether there are phases and phase transitions in our model of magnetism, whether they ever abruptly change, and how temperature affects these quantities.

6 Summary

In this session, we have introduced the concepts of phases of matter, and transitions between these phases, and discussed the difficulties and many concepts required to develop a mathematical description of these phases. We then discussed the process that physicists use to develop models for materials, and kinds of interactions between fundamental constituents, or particles that are being described by these models. We then introduced lattices, and how a particle can occupy each site, and interact through bonds between particles on this lattices, and gained some familiarity with summations over these particles and bonds on the lattice. Finally, we used these concepts to introduce our first model of magnetism, the ferromagnetic Ising model, with nearest-neighbour interactions.

In the next session, we will discuss what it means for a model to be easy or difficult to study. We will study the Ising model further, and introduce an important approach to making predictions based on a model called Mean-Field Theory. We will then start some calculations with this Ising model, using our new mean field theoretic approaches, and start to see how particles following the Ising model could exhibit distinct phases.

Are there any additional questions or comments?

What is everyone taking away from this session?

What would everyone like to learn more about regarding this session?

How was the level of content and delivery of this session?

ISSYP 2022 Mentoring

What is magnetism and what does it tell us about phases of
matter?

Session 3 - Solving Lattice Models and Mean-Field Theory

July 27, 2022

Abstract

In Session 3, we discuss the concept of solving a model for physical systems, and the underlying reasons why it may be easier or harder to extract important quantities of interest from a model and an understanding of how the materials behave. Within the context of the Ising model of magnetism, we discuss different approaches that could be taken to better understand systems that follow this model, and what types of quantities we aspire to extract from this description. We introduce the important preliminary technique to calculating quantities of interest of Mean-Field Theory, and discuss what assumptions this approach makes, and how that affects its predictions. We finally calculate the mean-field effective energy of the Ising model and begin to introduce the mean-field calculation of the partition function.

Contents

1	Solving Models of Physical Systems	1
1.1	Possible Approximations of Models	2
1.2	Uncorrelated Model Assumptions	3
2	Mean-Field Theory	4
2.1	Assumption of State Perturbations	4
2.2	Mean-Field Theory of Ising Model	4
3	Summary	7

1 Solving Models of Physical Systems

In previous sessions, we have introduced the concept of describing fundamental constituent particles in physical systems, and their interactions between each other, using a mathematical description of the energy ε_σ that these N particles possess. *Any* configuration of particles, described by the random variables $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ has a probability distribution

$$p_\sigma = \frac{1}{Z} q_\sigma \quad (1)$$

where the unnormalized distribution is

$$q_\sigma = e^{-\frac{\varepsilon_\sigma}{T}} \quad (2)$$

for the Boltzmann distribution at a given temperature T with energy ε_σ . This allows us to understand the most energetically favourable configurations.

To normalize this distribution requires computing the *partition function* as the sum over all possible configurations of the particles

$$Z = \sum_{\sigma \in \Xi} q_\sigma . \quad (3)$$

We have shown that any expectation value of any quantity f_σ distributed according to p_σ

$$\langle f \rangle = \frac{1}{Z} \sum_{\sigma \in \Xi} q_\sigma f_\sigma \quad (4)$$

can be calculated once we know the partition function of a perturbed energy partition:

$$\varepsilon_\sigma \rightarrow \varepsilon_\sigma - \alpha f_\sigma \quad (5)$$

$$Z_\alpha \rightarrow \sum_{\sigma \in \Xi} q_\sigma e^{-\alpha f_\sigma} , \quad (6)$$

leading to the form of expectation values as *derivatives of a so-called free energy*

$$\langle f \rangle = \left. \frac{\partial F_\alpha}{\partial \alpha} \right|_{\alpha=0} , \quad (7)$$

where the free energy is

$$F_\alpha = -T \ln Z_\alpha . \quad (8)$$

This physically interpretable quantity of the free energy, with units of energy given temperature T also has units of energy, can be thought of as a different, total energy description of the system, and is related to the ε_σ definition of energy through the above definition.

This method of perturbations that are set to zero at the end of the calculation, is known as a *Legendre transformation*, or a *Lagrange multiplier*, and indicates the *response* of the system when it is suddenly influenced by an *external field* α , acting on the quantity f_σ . Expectation values can then be thought of as derivatives of this perturbed free energy, in the limit of zero perturbation.

In the case of the quantity of interest already being a term in the energy function of the form αf_σ , then the perturbation doesn't have to be applied to the system, and the original partition function and its free energy can be derived to obtain the expectation value

$$\langle f \rangle = \frac{\partial F}{\partial \alpha}, \quad (9)$$

where the free energy is

$$F = -T \ln Z. \quad (10)$$

Therefore, we can reduce practically all of statistical physics to calculating the summation in the partition function and the free energy of the system!

Our objective is to calculate expectation values of quantities of interest to predict whether there are any distinct phases of materials, that follow a particular model of energy, and to determine whether we can show mathematically that there is a phase transition when we tune a particular parameter of the model.

However, as discussed in previous sessions, depending on the form of p_σ , or ε_σ for the Boltzmann distribution, this deceptively simple summation can be *impossible* to perform analytically, and *approximations*, or *numerical methods* must be used.

1.1 Possible Approximations of Models

If the partition function for a given physical system that follows the Boltzmann distribution with an energy ε_σ , is not able to be computed *exactly*, then the model must be *simplified* by modifying the form of the energy to an *effective* energy

$$\tilde{\varepsilon}_\sigma \approx \varepsilon_\sigma, \quad (11)$$

where any quantities associated with these effective assumptions will be denoted with a tilde $\tilde{f} \approx f$.

We also note that the energy, and its effective approximation are *local* quantities, meaning they can be described by a summation of the energy of each particle

$$\varepsilon_\sigma = \sum_i \varepsilon_{\sigma_i}, \quad \tilde{\varepsilon}_\sigma = \sum_i \tilde{\varepsilon}_{\sigma_i}. \quad (12)$$

This approximate form of the energy ideally still describes many of the phenomenon that are occurring within the physical system, but also has a form that lends itself to ensuring the partition function can be calculated.

1.2 Uncorrelated Model Assumptions

A system's particles may have many complicated interactions, making the particles *correlated*. These correlations make the summations $\sum_{\sigma \in \Xi}$ particularly difficult, and one guiding principle in making simplifying assumptions about the models, is to approximate the model by an effective *uncorrelated* model.

Here, the amount of correlation between two particles σ_i, σ_j is defined as the expectation value of how much they both co-vary from their expected value, also known as the *covariance*

$$\langle (\sigma_i - \langle \sigma_i \rangle) (\sigma_j - \langle \sigma_j \rangle) \rangle = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle . \quad (13)$$

What is the correlation/covariance of two independent particles?

In general, assuming particles are independent is far too restrictive, and we are unable to adequately represent our physical system by an effective model with independent particles because most physical phenomena arise from the interactions between particles. However using an effective model with uncorrelated particles can significantly simplify our calculations, while retaining some interactions between particles in our model.

We recall the energy of the Ising model for a given configuration $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N)$ of N particles on a lattice is

$$\varepsilon_\sigma = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (14)$$

where J represents how strongly the particles will tend to align with each other, and h represents the strength of the external magnetic field that forces the particles to align along its direction.

What is the local energy of each particle in the Ising model?

How do correlations arise in the Ising model?

What would be form of the energy, and resulting Boltzmann distribution if each particle is independent from each other, and the particles do not interact?

How would the resulting calculation of the partition be simplified for this effective model?

When may particles behave independently and this would be a good assumption to make?

How would we change the Ising model of energy to make its particles uncorrelated?

2 Mean-Field Theory

Given this notion of making approximations, and deriving an effective model for the energy of a system of interest, we will now introduce an important *mean-field* approach to approximations. Mean-field theory helps us decrease the correlations between particles in model, but still retains the types of interactions of the original model to be able to get a qualitative description.

Note that we will also still delay specifying the states that each particle can occupy, which lattice we are studying, or how many number of spatial dimensions we are in. This mean-field theory approach we will see does not rely on these specifications, and again only relies on knowing how each particle i 's z nearest-neighbours $\langle ji \rangle$.

2.1 Assumption of State Perturbations

Our main assumption, or what is known as an *ansatz*, is to represent each of the particles magnetic states as being equal to an average, *mean* value, plus an additional *perturbation*

$$\sigma_i = \langle \sigma_i \rangle + \delta_i , \quad (15)$$

and if we also assume the lattice is *homogeneous*, then each particle's average magnetization value are identical, and are equal to the average magnetization per particle across the lattice

$$\langle \sigma_i \rangle = \langle \sigma \rangle = \frac{1}{N} \sum_i \langle \sigma_i \rangle . \quad (16)$$

It can be useful to think of the perturbation as the difference between the magnetization value and the mean value

$$\delta_i = \sigma_i - \langle \sigma_i \rangle . \quad (17)$$

What is the expectation value of the perturbation?

How can we use this ansatz of writing the particle states with small perturbations to remove correlations from the model?

2.2 Mean-Field Theory of Ising Model

Now that we understand this notion of making our particles uncorrelated by assuming $\delta_i \ll \langle \sigma_i \rangle$, for states $\sigma_i = \langle \sigma_i \rangle + \delta_i$, we can insert this assumption into the Ising model, and see how this simplifies the effective energy.

The exact Ising model energy for a homogeneous lattice can be approximated as an effective

mean-field energy

$$\varepsilon_\sigma = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i \quad (18.a)$$

$$= -J \sum_{\langle ij \rangle} (\langle \sigma_i \rangle + \delta_i)(\langle \sigma_j \rangle + \delta_j) - h \sum_i (\langle \sigma_i \rangle + \delta_i) \quad (18.b)$$

$$= -J \sum_{\langle ij \rangle} (\langle \sigma_i \rangle \langle \sigma_j \rangle + \langle \sigma_j \rangle \delta_i + \langle \sigma_i \rangle \delta_j + \delta_i \delta_j) - h \sum_i (\langle \sigma_i \rangle + \delta_i) \quad (18.c)$$

$$\approx -J \sum_{\langle ij \rangle} (\langle \sigma_i \rangle \langle \sigma_j \rangle + \langle \sigma_j \rangle \delta_i + \langle \sigma_i \rangle \delta_j) - h \sum_i (\langle \sigma_i \rangle + \delta_i) \quad (18.d)$$

$$= -J \sum_{\langle ij \rangle} (\langle \sigma_i \rangle \langle \sigma_j \rangle + \langle \sigma_j \rangle (\sigma_i - \langle \sigma_i \rangle) + \langle \sigma_i \rangle (\sigma_j - \langle \sigma_j \rangle)) - h \sum_i \sigma_i \quad (18.e)$$

$$= -J \sum_{\langle ij \rangle} (\langle \sigma \rangle^2 + \langle \sigma \rangle (\sigma_i + \sigma_j - 2 \langle \sigma \rangle)) - h \sum_i \sigma_i \quad (18.f)$$

$$= J \sum_{\langle ij \rangle} (\langle \sigma \rangle^2 - \langle \sigma \rangle (\sigma_i + \sigma_j)) - h \sum_i \sigma_i \quad (18.g)$$

$$= \frac{1}{2} J \sum_i \sum_{\langle ji \rangle} (\langle \sigma \rangle^2 - \langle \sigma \rangle (\sigma_i + \sigma_j)) - h \sum_i \sigma_i \quad (18.h)$$

$$= \frac{1}{2} J \sum_i \sum_{\langle ji \rangle} \langle \sigma \rangle^2 - 2 \frac{1}{2} J \langle \sigma \rangle \sum_i \sum_{\langle ji \rangle} \sigma_i - h \sum_i \sigma_i \quad (18.i)$$

$$= \frac{zN}{2} J \langle \sigma \rangle^2 - zJ \langle \sigma \rangle \sum_i \sigma_i - h \sum_i \sigma_i \quad (18.j)$$

$$= \frac{zN}{2} J \langle \sigma \rangle^2 - (h + zJ \langle \sigma \rangle) \sum_i \sigma_i \quad (18.k)$$

We can therefore define the homogeneous effective mean-field Ising model energy as

$$\tilde{\varepsilon}_\sigma = \tilde{\Omega} - \tilde{h} \sum_i \sigma_i \quad (19)$$

where the effective mean-field depends on the average magnetization from a particle's nearest neighbours

$$\tilde{h} = h + zJ \langle \sigma \rangle \quad (20)$$

and there is a constant energy

$$\tilde{\Omega} = \frac{zN}{2} J \langle \sigma \rangle^2 . \quad (21)$$

As discussed, the constant energy can be neglected as it has no effect on the probabilities,

and

$$\tilde{\mathcal{E}}_{\sigma} = -\tilde{h} \sum_i \sigma_i . \quad (22)$$

Using mean-field theory of assuming the particles are uncorrelated, we have developed an effective model for the energy where the particles magnetization only interacts with a *mean-field* of the average of all the particles nearest to it!

How does this mean-field approximation limit our understanding of the full Ising model?

How will this mean-field effective energy be useful when computing expectation values and the partition function?

What if our system has disorder and is not homogeneous, and so each particle's average magnetization $\langle \sigma_i \rangle$ are not identical? How does the form of our effective energy change?

3 Summary

In this session, we have introduced the difficulties associated with calculating expectation values for models with complicated forms of their energy. More sophisticated correlations between particles lead to difficult analytical summations, particularly for the partition function. We then discussed approximations that can be made to alleviate some of these difficulties by developing an effective energy description of the model, particularly when the particles in this description are less correlated. We then introduced the important approximation method of Mean-Field theory, and show how by assuming particles' states only deviate by a small amount from their mean values, we can assume they are uncorrelated. Finally we return to our Ising model of magnetism, and calculate the effective energy using this mean-field approach, and find the effective energy is simply all the particle's magnetizations aligning with a mean-external-field.

In the next session, we will use this effective energy form to calculate the mean-field partition function, and other expectation values for the Ising model. Given these effective expectation values, we will be able to understand how certain quantities depend on different parameters of the model such as temperature, or field strength. We will finally get a qualitative sense of why this model of magnetization has phases, and a phase transition. We will finally discuss how our mean-field approximations affect our predictions, and compare to known exact results about phase transitions for the Ising model.

Are there any additional questions or comments?

What is everyone taking away from this session?

What would everyone like to learn more about regarding this session?

How was the level of content and delivery of this session?

ISSYP 2022 Mentoring

What is magnetism and what does it tell us about phases of
matter?

Session 4 - Phase Transitions in the Ising Model

July 28, 2022

Abstract

In Session 4, we use the effective mean-field energy of the Ising model to calculate the partition function and the effective expectation values of the energy and magnetization. Using these effective expectation values, we will observe there are some self-consistency equations that must be satisfied, and performing these checks, we will observe there are abrupt changes in certain quantities, indicating the presence of a phase transition at a specific critical temperature. We will discuss the phases of the Ising model, and what our mean-field approach has illuminated about phase transitions beyond just in the specific case of magnetic systems. We will finally discuss specific lattices and numbers of spatial dimensions for the exact Ising model, and compare some results that can be calculated with analytical and numerical methods to determine the successes and limitations of mean-field theory.

Contents

1	Effective Mean Field of the Ising Model	1
1.1	Effective Mean-Field Partition Function of the Ising Model	1
1.2	Effective Mean-Field Expectation Values of the Ising Model	3
1.3	Effective Expected Energy of the Ising Model	4
1.4	Effective Expected Magnetization of the Ising Model	4
2	Consistency within Mean-Field Theory	5
3	Accuracy of Mean-Field Theory	7
4	Monte Carlo Methods	8
5	Summary	9

1 Effective Mean Field of the Ising Model

Given the Ising model of energy for particles with state $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_N) \in \Xi$ on a lattice with N particles, each with z nearest neighbours

$$\varepsilon_\sigma = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i , \quad (1)$$

in previous sessions we have derived the effective mean-field energy

$$\tilde{\varepsilon}_\sigma = -\tilde{h} \sum_i \sigma_i , \quad (2)$$

as solely the particles' magnetizations interacting with an external effective magnetic field

$$\tilde{h} = h + zJ \langle \sigma \rangle . \quad (3)$$

This system at temperature T follows the Boltzmann distribution for being in state σ

$$p_\sigma = \frac{1}{Z} q_\sigma \quad (4)$$

where the unnormalized distribution is

$$q_\sigma = e^{-\frac{\varepsilon_\sigma}{T}} \quad (5)$$

the exact normalizing partition function is

$$Z = \sum_{\sigma \in \Xi} q_\sigma , \quad (6)$$

and exact expectation values are

$$\langle f \rangle = \frac{1}{Z} \sum_{\sigma \in \Xi} q_\sigma f_\sigma . \quad (7)$$

1.1 Effective Mean-Field Partition Function of the Ising Model

We are now able to calculate the effective partition function $\tilde{Z} \approx Z$ with the effective energy of these essentially independent particles

$$\tilde{Z} = \sum_{\sigma \in \Xi} \tilde{q}_\sigma , \quad (8)$$

where the effective unnormalized distribution is

$$\tilde{q}_\sigma = e^{-\frac{\tilde{\varepsilon}_\sigma}{T}} , \quad (9)$$

and in the case of independent particles $\tilde{\varepsilon}_\sigma = \sum_i \tilde{\varepsilon}_{\sigma_i}$

$$\tilde{q}_\sigma = \prod_i \tilde{q}_{\sigma_i} = \prod_i e^{-\frac{\tilde{\varepsilon}_{\sigma_i}}{T}} . \quad (10)$$

The effective partition function for the Ising model can therefore be split into summations over the individual distributions for each particle

$$\tilde{Z} = \left[\prod_i \sum_{\sigma_i \in \Xi_i} \right] \left[\prod_i \tilde{q}_{\sigma_i} \right] \quad (11)$$

$$= \prod_i \left[\sum_{\sigma_i \in \Xi_i} \tilde{q}_{\sigma_i} \right] \quad (12)$$

$$= \prod_i \left[\sum_{\sigma_i \in \Xi_i} e^{\frac{\tilde{h}}{T} \sigma_i} \right] \quad (13)$$

$$= \prod_i \tilde{Z}_i \quad (14)$$

and our calculation of the effective partition function has reduced to calculating the effective partition function for each particle

$$\tilde{Z}_i = \sum_{\sigma_i \in \Xi_i} e^{\frac{\tilde{h}}{T} \sigma_i} , \quad (15)$$

and given $\Xi_i = \{\pm 1\}$ each individual effective partition function is identical

$$\tilde{Z}_i = \underline{\tilde{Z}} = 2 \cosh(\tilde{h}/T) . \quad (16)$$

Therefore our effective partition function for the Ising model is

$$\tilde{Z} = \underline{\tilde{Z}}^N \quad (17)$$

$$= 2^N \cosh^N(\tilde{h}/T) \quad (18)$$

$$= 2^N \cosh^N((h + zJ \langle \sigma \rangle)/T) . \quad (19)$$

We can see the effective partition function from our mean-field approach has a simple form that depends on the field values in the model h, J , the coordination number of the lattice z , the expected homogeneous magnetization $\langle \sigma \rangle$, and the temperature T .

1.2 Effective Mean-Field Expectation Values of the Ising Model

Given our calculation of the effective partition function, we are now able to calculate effective expectation values $\langle \widetilde{f} \rangle \approx \langle f \rangle$ of quantities of interest

$$\langle \widetilde{f} \rangle = \frac{1}{\widetilde{Z}} \sum_{\sigma \in \Xi} \widetilde{q}_{\sigma} \widetilde{f}_{\sigma} , \quad (20)$$

where notice that depending on the quantity f , we are making approximations in both the effective distribution \widetilde{p}_{σ} , as well as possibly the quantity itself f if it depends on our approximations of uncorrelated particles.

In the case of independent particles, *local* quantities that only depend on each particle individually, such as the average of a quantity per particle

$$f_{\sigma} = \frac{1}{N} \sum_i f_{\sigma_i} , \quad (21)$$

and does not depend on any interactions (products) of particles' states, have expectation values that can be split into summations over the individual distributions for each particle

$$\langle \widetilde{f} \rangle = \frac{1}{N} \sum_i \langle \widetilde{f_{\sigma_i}} \rangle = \frac{1}{N} \sum_i \left[\frac{1}{\widetilde{Z}_i} \sum_{\sigma_i \in \Xi_i} \widetilde{q}_{\sigma_i} \widetilde{f_{\sigma_i}} \right] . \quad (22)$$

Note that there are two different averages occurring, both the expected value of each particle's state, which averages over the Boltzmann distribution of probabilities of that particle being in that energy state (the $\langle \rangle$ average), and the average over each of these particles' expectation values over the whole lattice (the $1/N \sum_i$ average).

For our effective Ising model, we will look at different quantities, and see if there is an abrupt change in any of them as we change any parameters of the model, such as the temperature, or field strengths h, J past some *critical value*.

We also should note that our effective model depends on an effective field, which depends on the *macroscopic* property of the average expected magnetization per particle

$$\widetilde{h} = h + zJ \langle \sigma \rangle \quad (23)$$

and any expectation values, such as the average magnetization per particle, that are also macroscopic, averaged properties across the lattice, may require *consistency* checks to ensure the left and right hand side of these equations are consistent with each other.

1.3 Effective Expected Energy of the Ising Model

We will first calculate the effective average energy per particle (a local quantity) for this (homogenous) effective model

$$\langle \widetilde{\varepsilon} \rangle = \frac{1}{N} \sum_i \langle \widetilde{\varepsilon}_i \rangle \quad (24)$$

$$\langle \widetilde{\varepsilon}_i \rangle = -\tilde{h} \langle \widetilde{\sigma}_i \rangle \quad (25)$$

$$= -\tilde{h} \langle \widetilde{\sigma} \rangle . \quad (26)$$

Therefore the effective average energy per particle

$$\langle \widetilde{\varepsilon} \rangle = -\tilde{h} \langle \widetilde{\sigma} \rangle \quad (27)$$

is just the effective field times the effective average magnetization per particle.

1.4 Effective Expected Magnetization of the Ising Model

We will now calculate the effective average magnetization per particle (a local quantity) for this (homogenous) effective model

$$\langle \widetilde{\sigma} \rangle = \frac{1}{N} \sum_i \langle \widetilde{\sigma}_i \rangle \quad (28)$$

$$\langle \widetilde{\sigma}_i \rangle = \frac{1}{\tilde{Z}_i} \sum_{\sigma_i \in \Xi_i} \tilde{q}_{\sigma_i} \sigma_i \quad (29)$$

$$= \frac{1}{2 \cosh(\tilde{h}/T)} \sum_{\sigma_i \in \{\pm 1\}} e^{-\tilde{h} \sigma_i / T} \sigma_i \quad (30)$$

$$= \frac{1}{2 \cosh(\tilde{h}/T)} 2 \sinh(\tilde{h}/T) \quad (31)$$

$$= \tanh(\tilde{h}/T) . \quad (32)$$

Therefore the effective average magnetization per particle

$$\langle \widetilde{\sigma} \rangle = \tanh(\tilde{h}/T) \quad (33)$$

and we have finally calculated approximate expectation values of the Ising model using mean-field theory!

2 Consistency within Mean-Field Theory

If we assume that our exact average magnetization

$$\langle \sigma \rangle \approx \widetilde{\langle \sigma \rangle} \quad (34)$$

we used in our mean-field ansatz is approximately equal to the effective average magnetization, then what do we notice about this equation for the magnetization above? Hint: How is the effective magnetic field defined?

It turns out in using our effective magnetic field in our calculation for the effective average magnetization, we have caused there to be possible *inconsistencies* in this value of the effective magnetization unless we impose that the following is true:

$$\widetilde{\langle \sigma \rangle} = \tanh((h + zJ\widetilde{\langle \sigma \rangle})/T) . \quad (35)$$

This *self-consistency* equation must be true if our model is to be valid and to predict consistent results for expectation values. However, we can not analytically find the values of $\widetilde{\langle \sigma \rangle}$ that satisfy this equation, and must *plot* both sides of the equation to see where they are equal. We will focus on the simplified case where the magnetic field $h = 0$.

What do both sides of the self-consistency equation look like when plotted as a function of $\widetilde{\langle \sigma \rangle}$?

We see that this self-consistency equation has different solutions for the average magnetization, depending on the value of the temperature!

We can see from our graphs that our mean-field theory approach predicts that the magnetization *must* be 0 for temperatures greater than a *critical-temperature* T^* , however can be a *non-zero* value σ for temperatures less than a critical-temperature:

$$\widetilde{\langle \sigma \rangle} = \begin{cases} 0 & T > zJ \\ \pm \tilde{\phi} & T \leq zJ \end{cases} , \quad (36)$$

and our mean-field theory approach predicts the Ising model to have a critical temperature of

$$\tilde{T}^* = zJ . \quad (37)$$

Why can the magnetization take two possible non-zero values $\pm \tilde{\phi}$ at temperatures below the critical temperature?

How does this notion of there being distinctly different magnetization values at different temperatures above and below a critical temperature T^* , relate to our previous discussions of

phases and phase transitions?

From our mean-field analysis we can predict that the Ising model has *two* distinct phases, and can identify the *magnetization* as the *order-parameter* that distinguishes the phases. The magnetization order parameter indicates at *low temperatures* below a critical temperature, the material is *ordered*, all of its particle's magnets tend to *align* along one direction, and cause the material to be magnetized. At temperatures above a critical temperature, the material is *disordered*, all of its particle's magnets tend to align *randomly*, and causing cancellation all of their magnetic fields, and the material to be not magnetized. These predictions are summarized in Table 1.

Table 1: Labels, magnetization values, and temperature ranges of phases of mean-field Ising model, with a mean-field critical temperature of $T^* = zJ$.

Phase	$\langle\sigma\rangle$	T
Ordered	$\pm\phi$	$0 < T < T^*$
Disordered	0	$T^* \leq T < \infty$

Have we made any assumptions about what type of material we are studying, such as its shape, or number of spatial dimensions? What does this say about the generality of mean-field theory ansatz?

Are these mean-field predictions reasonable?

Other than assuming we have a two-level system, and calling the order parameter a magnetization, have we assumed anything about this system's particles being magnets? Could we apply our approach and knowledge of phase transitions to study systems other than magnets? What if we have more than 2 possible values for our particle's states?

3 Accuracy of Mean-Field Theory

It turns out, that mean-field theory is *incorrect* for lattices with spatial dimension $d \leq 4$! The exact reasons for this are beyond these sessions, but for $d = 2, 3, 4$ essentially it is because in these lower dimensions, each particle has *less* nearest neighbours than in higher dimensions, for example hyper-cubic lattices have $z = 2d$, and mean-field theory *overestimates* a material's tendency to *order* at low temperatures and become magnetized. This leads to mean-field predicting a critical temperature that is too high, and in reality only at very low temperatures, are the states that cause ordering adequately probable to ensure magnetization. For example, the exact critical temperature for a $d = 2$ *square* lattice with $z = 4$ nearest neighbours is

$$T^* = \frac{2}{\ln(1 + \sqrt{2})}J \approx 2.27J < \tilde{T}^* = 4J. \quad (38)$$

In $d = 1$ dimensions for a *chain* of particles, the situation is even more subtle, and there is in fact *no phase transition*! The reason for this is essentially at all temperatures, the system is more likely to disorder, then order, and so no magnetization ever forms, and there is no ordered phase.

Fortunately, this mean-field theory approach has been very useful to get a sense of how temperature affects the phases of a material, and drives these phase transitions. We have learned that the Ising model of magnetism may exhibit distinct phases, depending on the temperature, and number of spatial dimensions you are in. Even more important, we have kept track of our assumptions, and learned about the *universality* of these approaches and predictions to many systems that follow a Boltzmann distribution, and have a discrete number of possible states.

How could we improve our effective energy model to account for these discrepancies with exact results?

We have dealt with strictly classical materials, what if our material behaved quantum-mechanically?

The $d = 1$ chain and $d = 2$ square lattice Ising models can be solved exactly, how could we go about this? Hint: Wait until university! And if you can solve for the exact critical temperature for $d > 2$ dimensions, a Nobel Prize could be in your future!

4 Monte Carlo Methods

We have seen how difficult analytical calculations of expectation values can be, and numerical methods, including *Monte Carlo sampling* procedures to estimate integrals are an important and effective alternative. Many algorithms obtain *samples* of configurations from the 2^N possible total configurations, by performing a series of *updates*, also known as *spin-flips*, to the lattice, with the objective of obtaining updated configurations that contribute significantly to the expectation values.

The updates to the lattice consist of a *proposal* and an *acceptance* step, where the updated configuration may be *rejected* if it is not energetically favourable to evolve to this new configuration. The algorithm has only a few steps, which are repeated many, for example $M = O(N^2)$, times

1. Randomly initialize N lattice values σ at a temperature T
2. Perform updates M times
 - (a) Propose a randomly chosen site i to update $\sigma_i \rightarrow -\sigma_i$
 - (b) Calculate the energy difference Δ_σ between the current and proposed configuration
 - (c) Accept and update to the proposed configuration with probability $e^{-\Delta_\sigma/T}$
 - (d) Calculate any quantities for the current configuration
3. Average quantities over the M samples

Here, we have described what is known as a *Metropolis-Hastings* update, of flipping a *single-site*. Analysis for why this algorithm, and its variants are so successful, are outside the scope of these sessions, however they are very effective at estimating expectation values by finding important configurations to sample. A sample estimate of the average energy and magnetization per particle for the Ising model as a function of temperature, from [*here*](#) is shown in Figure 1.

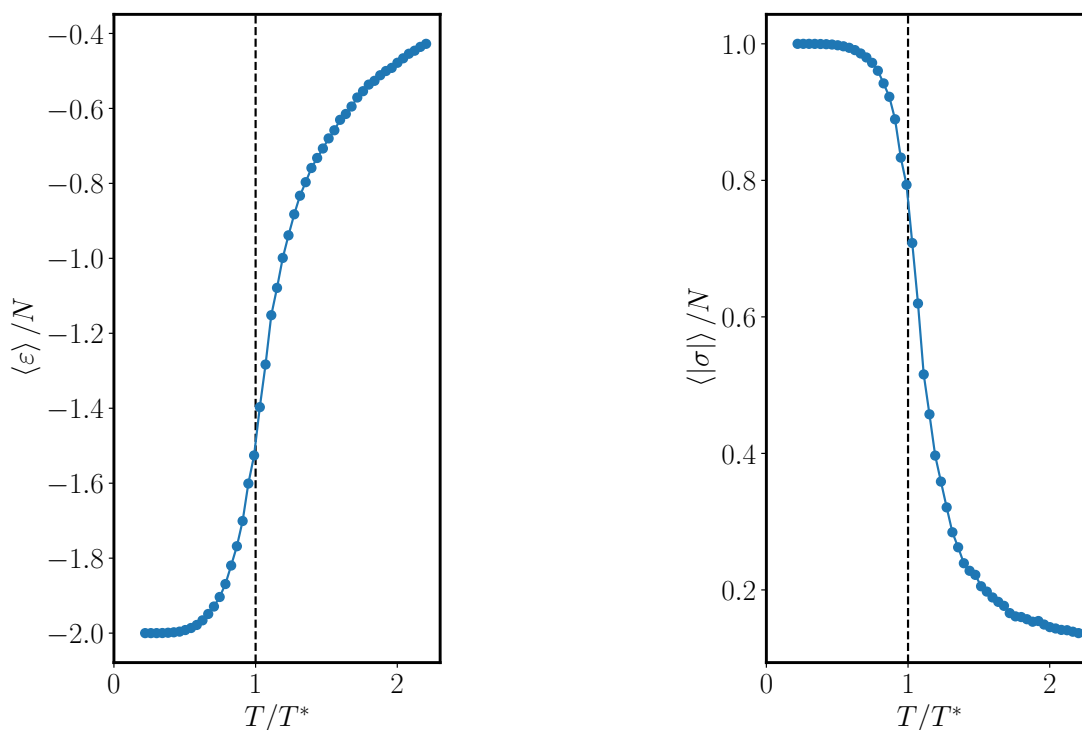


Figure 1: Ising model average energy and magnetization per particle as a function of temperature T , for a square lattice, with $J = 1$, $h = 0$ fields, $N = 10^2$ particles, and $M = N^2$ samples.

5 Summary

In this session, we have applied a mean-field ansatz of an effective energy model for the Ising model to explicitly calculate the partition function, as well as the effective average energy per particle, and the effective average magnetization per particle. From the magnetization calculations, we have seen mean-field theory can lead to there being self-consistency equations that must be satisfied. The solutions of these equations for the average magnetization lead to predictions a critical temperature, below which, the material's magnetization is non-zero and the material's particles are aligned, producing magnetization. Above the critical temperature, the material's magnetization is zero, and the material's particles are not aligned, producing no magnetization. We finally discuss the universality of the Ising model, and the mean-field approach, and compare its predictions to exact results.

Are there any additional questions or comments?

What is everyone taking away from this session?

What would everyone like to learn more about regarding this session?

How was the level of content and delivery of this session?