Statistical Physics

Jari Kinaret
Department of Applied Physics
Chalmers University of Technology
and Göteborg University
S-41296 Göteborg, SWEDEN

Preface

This is the second draft of these lecture notes, and students' comments would be appreciated as a source for further revisions and improvements.

Much of what follows has been borrowed from various sources. The most important sources for the first part have been *Irreversible Statistical Physics* by Alf Sjölander (Chalmers, Göteborg, 1990), and *Principles of Condensed Matter Physics* by P. M. Chaikin and T. C. Lubensky (Cambridge University Press, Cambridge, 1995). The second part contains material from a graduate course in statistical mechanics taught by Professor Mehran Kardar at the Massachusetts Institute for Technology in the spring of 1988; other important references are the aforementioned book by Chaikin and Lubensky, *Equilibrium Statistical Mechanics* by M. Plischke and B. Bergersen (World Scientific, Singapore, 1994), *Statistical Mechanics* by K. Huang (Wiley, New York, 1987), *Modern Theory of Critical Phenomena* by S.-K. Ma (Benjamin-Cummings, Reading, 1976), and *The Renormalization Group* by N. Goldenfeld (Addison-Wesley, Reading, 1992). A more rigorous and theoretical discussion on many of the topics in the second half of the course can be found in *Statistical Field Theory*, vol. 1, by C. Itzykson and J.-M. Drouffe (Cambridge University Press, Cambridge, 1994).

I would like to thank Dr Göran Wendin and Mr. Peter Sandström for their valuable comments. I would also like to thank Dr Carolyn Berland for her comments on the language of these notes and for producing many of the figures.

Contents

List of Figures

List of Tables

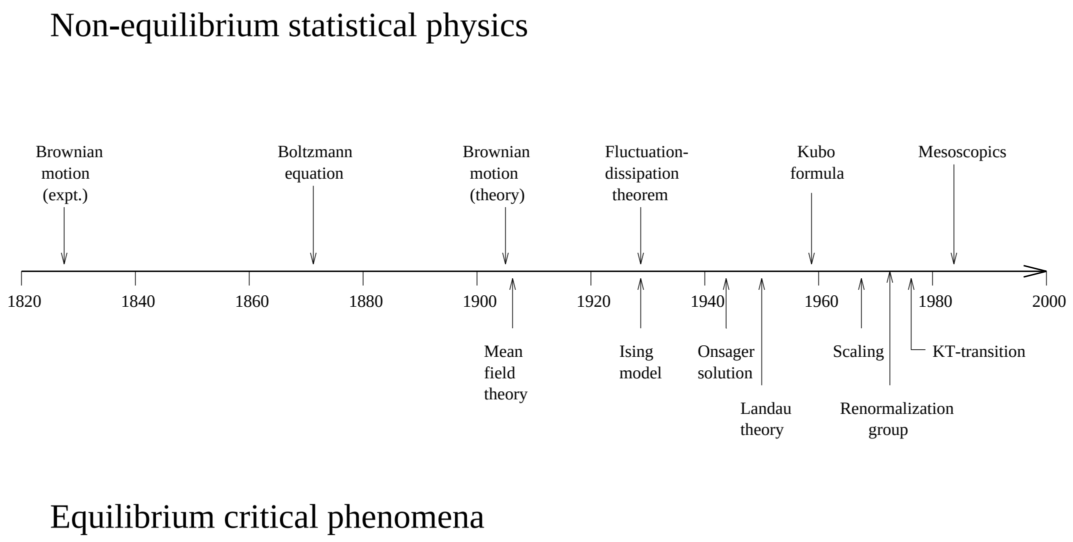


FIG. 1. Approximate times when various topics that are discussed in this course were first introduced or discovered.

I. INTRODUCTION

This course on statistical physics consists of two rather disconnected parts. In the first part of the course we will discuss systems out of equilibrium and focus on issues such as how a system relaxes towards an equilibrium, and the description of time-dependent phenomena in the context of statistical physics. On a practical level we will study various transport equations and their applications to different problems. After completing the first part of the course you should understand the general properties of time-dependent distribution functions and be able to apply your knowledge to transport problems in various contexts.

In the second part of the course we will focus on systems in thermal equilibrium, concentrating on the properties of systems in the vicinity of a phase transition. We will discuss the basic aspects of critical phenomena and familiarize ourselves with a number of exact and approximate methods that are used in the field. After having read the second part you should understand the general characteristics of phase transitions, and be able to apply the approximate techniques (mean field approximation, renormalization group) to specific critical phenomena. The material of the second part is considerably more modern than the material of the first part, which is indicated in Figure 1.

For the most part this course deals with the classical $\hbar = 0$ physics, and quantum mechanical effects are mentioned only occasionally.

I will assume that you have already taken a course in equilibrium statistical mechanics and thermodynamics. In particular, I will assume that you are familiar with concepts like partition function (*"tillståndsumma"*), free energy, specific heat (or heat capacity, *"värmekapacitet"*). I will also assume that you know, either from quantum mechanics or classical mechanics, the concept of a Hamiltonian. Necessary mathematical background includes elements of calculus (what is $\int_{-\infty}^{\infty} dx\, e^{-\alpha x^2}$?) and elements of statistics or probability theory.

II. APPROACH TO EQUILIBRIUM

A. Introduction

A system is in a thermodynamic equilibrium if its state, as described by the macroscopic measurable quantities like temperature, pressure, volume, magnetic field and magnetization, does not change with time. Most, if not all, of the thermodynamics and statistical physics you have learned so far has probably dealt with the properties of systems in thermal equilibria — why is that? Examining the world around us reveals that most systems are not in a time-independent equilibrium state but rather behave quite dynamically; for instance, all living organisms are clearly out of equilibrium. It appears that thermal equilibrium is more of an exception than a rule, and most likely it is an idealized concept which does not occur anywhere in nature. Therefore, understanding the non-equilibrium dynamics is of primary importance if we are to make sense out of the world around us.

Although thermal equilibrium, strictly speaking, is not found anywhere, it is nevertheless an important concept. First of all, there is the pragmatic reason that since from a macroscopic point of view the equilibrium state appears to be static, its description should be rather straightforward, and we have some hope of understanding it ("Let's do something that we can do rather than something that is rigorously correct"). Secondly, most systems have some dissipative mechanisms that convert other forms of energy to heat (second law of thermodynamics). This dissipation drives the system towards an equilibrium, so even if we consider a system that is initially out of equilibrium, it will approach equilibrium if we wait sufficiently long. Thirdly, many systems are near equilibrium, and their state can be described as a small deviation from the equilibrium state. Consequently, their properties are related to the equilibrium properties.

In the first part of this course we will focus on the second and third points: how does a non-equilibrium system approach equilibrium, and how can we relate the properties of a non-equilibrium system to those of a system in equilibrium. To achieve this goal we must start by defining the necessary concepts which we will subsequently use in describing the dynamics. The concepts that we define, as well as many of the general dynamic equations, are applicable quite generally even outside physics.

B. Phase Space and Classical Distribution Functions

1. Definitions

Let us consider a classical system in d dimensions with N particles. The microscopic state of the system is fully defined if we know the position and momentum of each of the N particles. If we denote the position of the i^{th} particle by the d-dimensional vector $\vec{q}_i$ and similarly denote the momentum of the particle by $\vec{p}_i$, the microstate of the system is specified by the vectors $\{\vec{q}_1, \vec{p}_1, \vec{q}_2, \vec{p}_2, \ldots, \vec{q}_N, \vec{p}_N\}$. The $2dN$ dimensional space spanned by these vectors is known as *the phase space* Γ of the system.

For simplicity let us denote the microstate of the system by the $2dN$-dimensional vector $\vec{X}$. As the system evolves in time it traces a trajectory $\vec{X}(t)$ in phase space — the simplest example is the time development of a set of non-interacting particles in the absence of external forces: from Newton's laws we know that the momenta of the particles do not change and the particles' positions change with constant velocities; consequently the state of the system as a function of time is a straight line in phase space. If there are external or internal forces acting upon the system the trajectory in phase space can be considerably more complicated and, under some circumstances, chaotic.

In statistical physics the system usually consists of a huge number of particles ($\sim 10^{23}$), and it is not possible to determine the positions and momenta of all particles. Instead, we define the *macrostate* of the system by specifying some macroscopic quantities like pressure, volume and temperature. Usually particular values of these macroscopic quantities can originate from any one of a large number of different microstates (the exception is if we specify $T = 0$, in which case the system will be in one of a (usually) small number of ground states). Therefore, we define a probability density $\overline{f}_N(\vec{X}, t)$ such that

$$\overline{f}_N(\vec{X}, t)\Delta X = \text{Probability that at time } t \text{ the microstate of the system is in the neighborhood } \Delta X \text{ of the point } \vec{X} \text{ in the phase space.}$$

Since $\overline{f}_N(\vec{X}, t)$ is a probability density it obeys the normalization condition

$$\int d^{2dN}X\overline{f}_N(\vec{X}, t) = 1. \tag{2.1}$$

Let us prepare the system at time $t = t_0$ so that it occupies a particular region $\mathcal{R}$ in phase space (i.e. the distribution function $\overline{f}_N(\vec{X}, t_0)$ is nonzero if $\vec{X} \in \mathcal{R}$). Since each microstate within this region develops in time in a way that is fully determined by the Hamiltonian, we can in principle calculate the precise region $\mathcal{R}(t)$ that the system occupies at some other time t. In practice, however, this task may be well-nigh impossible — interactions within the system (e.g. collision between particles) quickly deform the region $\mathcal{R}$ in a very complicated fashion as I have illustrated in Fig. 2. However, we are able to make some predictions of the system's time-development using *Liouville's theorem*, which states that *the volume of the $2dN$-dimensional region that an arbitrary system occupies in phase space remains constant in time.* Since the number of particles in the system does not change with time we can re-phrase Liouville's theorem and state that *the density of particles in phase space remains constant in time.* Because of this property the time evolution of a system in phase space can be described as a flow of an incompressible fluid.

To prove Liouville's theorem we consider a system that at time t_0 has been prepared so that it occupies the region $\Delta X = \{\vec{X} | \vec{q}_1 \in \Delta q_1, \vec{p}_1 \in \Delta p_1, \vec{q}_2 \in \Delta q_2, \vec{p}_2 \in \Delta p_2, \ldots, \vec{q}_N \in \Delta q_N, \vec{p}_N \in \Delta p_N\}$.[1] Let us now consider the time development of the x-components of the position and momentum of the j^{th} particle. If at time t the x-component of $\vec{q}_j$ is confined to an interval $[x_a(t), x_b(t)]$ (or to a collection of such intervals), then at time $t + \delta t$ it is confined

---

[1] The volume of ΔX is given by $|\Delta X| = \prod_j \prod_\alpha (\Delta q_{j,\alpha} \Delta p_{j,\alpha})$. Here $\Delta q_{j,\alpha}$ and $\Delta p_{j,\alpha}$ denote the lengths of intervals in position and momentum spaces.

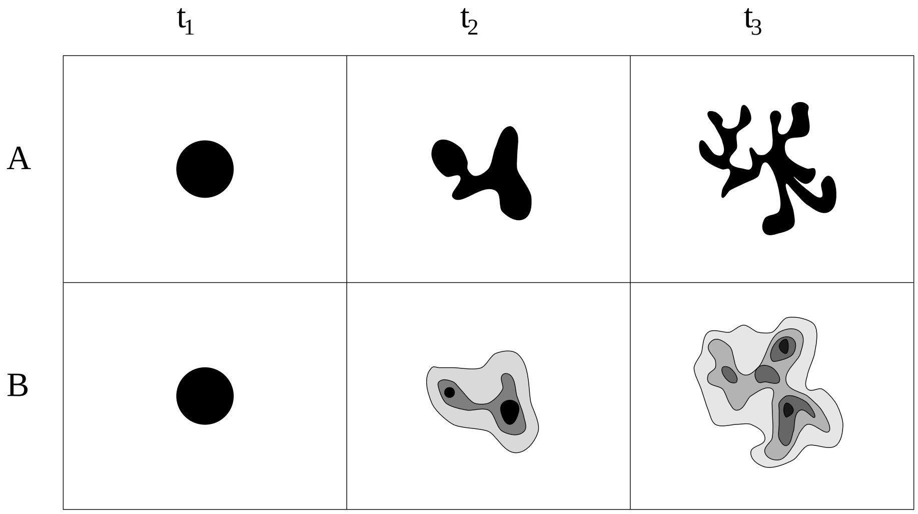


FIG. 2. Illustration of the time evolution of the region $\mathcal{R}$ that a system occupies in phase space. Liouville's theorem states that the volume $|\mathcal{R}(t)|$ of the region remains constant. Here the time series A illustrates the microscopic evolution towards a contorted shape, and B depicts the coarse-grained evolution which is probed by experiments.

to the interval $[x_a(t+\delta t), x_b(t+\delta t)]$ where the time evolution of x_a and x_b is governed by the Hamiltonian. Therefore, the length of the interval at time $t+\delta t$ is given by

$$\begin{aligned} \Delta q_{j,x}(t+\delta t) &= x_b(t+\delta t) - x_a(t+\delta t) \approx x_b(t) + \delta t\, \dot{x}_b(t) - x_a(t) - \delta t\, \dot{x}_a(t) \\ &= \Delta q_{j,x}(t) \left[1 + \delta t \frac{\Delta \dot{q}_{j,x}(t)}{\Delta q_{j,x}(t)} \right] . \end{aligned} \tag{2.2}$$

Similarly we get

$$\Delta p_{j,x}(t+\delta t) = \Delta p_{j,x} \left[1 + \delta t \frac{\Delta \dot{p}_{j,x}(t)}{\Delta p_{j,x}(t)} \right] . \tag{2.3}$$

Repeating this analysis for all the other spatial components α and for all particles j we get, to first order in δt,

$$|\Delta X(t+\delta t)| = |\Delta X(t)| \left[1 + \delta t \sum_{j=1}^{N} \sum_{\alpha=1}^{d} \left(\frac{\Delta \dot{q}_{j,\alpha}(t)}{\Delta q_{j,\alpha}(t)} + \frac{\Delta \dot{p}_{j,\alpha}(t)}{\Delta p_{j,\alpha}(t)} \right) \right] . \tag{2.4}$$

The terms in the sum can be simplified using Hamilton's equations of motion

$$\left\{ \begin{aligned} \dot{q}_{j,\alpha} &= \frac{\partial H}{\partial p_{j,\alpha}} \\ \dot{p}_{j,\alpha} &= -\frac{\partial H}{\partial q_{j,\alpha}} \end{aligned} \right. \tag{2.5}$$

which give for infinitesimally small intervals (i.e. going from finite differences to derivatives)

$$\frac{\Delta\, \dot{q}_{j,\alpha}\,(t)}{\Delta q_{j,\alpha}(t)} + \frac{\Delta\, \dot{p}_{j,\alpha}\,(t)}{\Delta p_{j,\alpha}(t)} \rightarrow \frac{\partial}{\partial q_{j,\alpha}}\frac{\partial H}{\partial p_{j,\alpha}} - \frac{\partial}{\partial p_{j,\alpha}}\frac{\partial H}{\partial q_{j,\alpha}} = 0 \tag{2.6}$$

and therefore $\frac{d}{dt}|\Delta X(t)| = 0$ and we have proven Liouville's theorem.

2. Averages and Fluctuations

a. Self-Averaging Quantities. Let us consider a system the macrostate of which we have specified with the macroscopic thermodynamic variables, say, temperature and volume. What is the pressure in this system? The answer is not obvious — after all, the system could be in any of a large number of microstates, and there is no guarantee that the pressure is the same in all of them. It sounds reasonable, however, to use the distribution function that we introduced in the previous section and calculate the average pressure weighted by the distribution function. A similar procedure can be used for any macroscopic observable.

This kind of an average where we weight each microstate with the distribution function is called the *ensemble average.* Mathematically, the ensemble average of a quantity A is defined as

$$\langle A(t)\rangle = \frac{\int d^{2dN}X\, A(\vec{X})\overline{f}_N(\vec{X},t)}{\int d^{2dN}X\overline{f}_N(\vec{X},t)}. \tag{2.7}$$

Since the value $A(\vec{X})$ is usually different for different microstates, it is necessary to define also the variance of A, which is given by $\mathrm{var}(A(t)) = \langle [A(t) - \langle A(t)\rangle]^2\rangle$. Occasionally it may even be necessary to consider the full probability distribution $P_A(A,t)$ of the variable A. That is defined by

$$P_A(A_0,t) = \frac{\int d^{2dN}X\delta(A_0 - A(\vec{X}))\overline{f}_N(\vec{X},t)}{\int d^{2dN}X\overline{f}_N(\vec{X},t)} \tag{2.8}$$

where $\delta(x)$ is the Dirac delta function.

Usually, however, we need not know the full probability distribution $P_A(A,t)$. The reason is that most quantities of interest are *self-averaging*, by which we mean that for large systems (i.e. large N) the standard deviation $\delta A(t) = \sqrt{\mathrm{var}(A(t))}$ is much smaller than the average value $\langle A(t)\rangle$. The reason for this behavior is that a large system can be divided into a large number of small subsystems in such a way that the fluctuations in the small systems are statistically independent. It then follows from the central limit theorem of statistics that an extensive quantity (i.e. one whose value for the large system is given as a sum of its values for the small subsystems) has a normal distribution. The mean of the distribution is given by the sum of the average values of the quantity in each subsystem, and the variance is a sum of the variances in the subsystems. Since the number of statistically independent subsystems increases linearly with N, both the average value of the quantity as well as its variance increase linearly with N. The standard deviation is the square root of the variance, so it only increases as $\sqrt{N}$, and the ratio between the standard deviation and the average value decreases as $\sqrt{N}/N \sim N^{-1/2}$. Therefore, fluctuations are relatively unimportant in macroscopic systems with a large number of particles, but quite important in microscopic

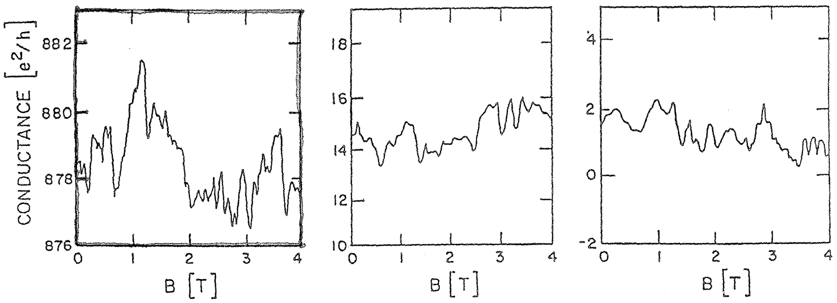


FIG. 3. Conductances of different mesoscopic samples as a function of the magnetic field. Note that the typical fluctuations are independent of the average conductance. Left: gold ring with diameter 0.8μm, center: quasi-1D Si MOSFET, right: numerical simulation. Figure taken from A. D. Stone, *Proc. 2nd Int. Symp. Foundations of Quantum Mechanics*, Tokyo, 1986, pp. 207–217.

systems like atoms where $N \lesssim 100$. We will consider fluctuations more explicitly in the chapter on thermal equilibrium.

Thus, we have concluded that most quantities which can be written as integrals of the distribution function over the entire phase space are self-averaging. This accounts for most quantities of physical interest. Exceptions occur first of all in systems where the energy of a microstate is a very irregular function of the phase space coordinate $\vec{X}$ — this case occurs for instance in spin glasses [1]. Another exception occurs when the quantity that self-averages is not the one that we usually consider: for instance the conductance G of a disordered one-dimensional wire is not self-averaging but its logarithm $\log(hG/e^2)$ is.

A particular example of the lack of self-averaging is seen in small electronic systems [2]. Experimentally it is seen that the conductance G (i.e. the inverse of the resistance) of a set of small, similar conductors varies from one sample to another. Self-averaging would imply that the typical relative fluctuations $\delta G/\langle G\rangle$ should depend on the size L of the conductor and decrease as $L^{-1/2}$. In contrast, measurements show that the fluctuations are *independent* of both L and the average value $\langle G\rangle$, and are given by $\delta G \approx \frac{1}{26k\Omega}$ as is shown in Fig. 3. Thus, the conductance of these systems is not a self-averaging quantity. The lack of self-averaging in this case is an entirely quantum mechanical effect: the quantum mechanical wave function extends over the entire sample and it is impossible to divide the sample into statistically independent subsystems. Since the number of particles is still quite large in these systems, typically $N \sim 10^4 \ldots 10^6$, they are macroscopic from the point of view of the particle number, but from the point of view of the importance of fluctuations these systems are microscopic. This dual character led Yoseph Imry in 1986 [3] to coin the new term *mesoscopic* to describe the intermediate regime between microscopic and macroscopic worlds (from the Greek *mesos*, meaning "between"). For about ten years mesoscopic physics has been a well-defined research area within condensed matter physics.

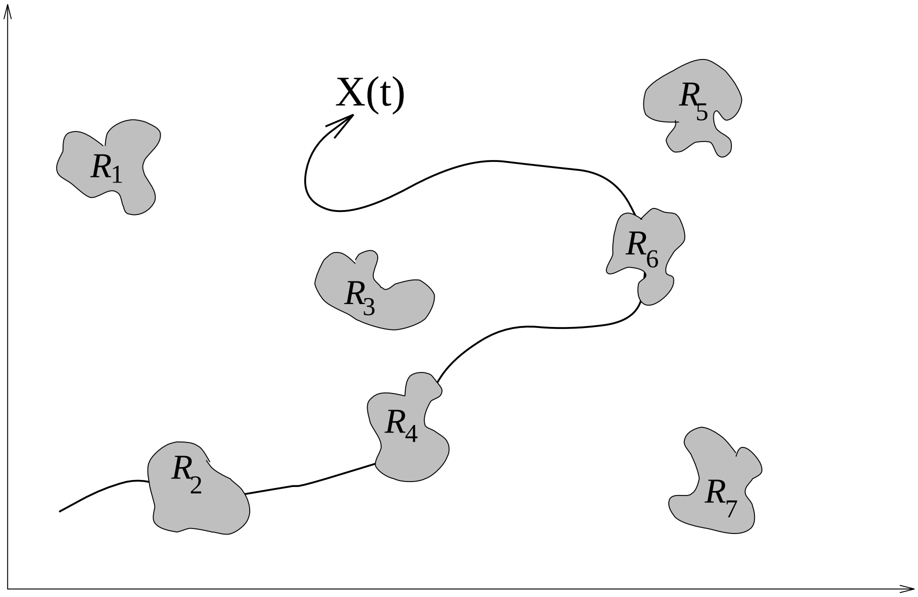


FIG. 4. Schematic trajectory of a particular microstate as a function of time. Note that the trajectory probes several mutually accessible regions of the phase space.

3. Equilibrium Distributions

Let us consider the time evolution of the distribution function $\overline{f}_N(\vec{X},t)$. Since each microstate follows a particular trajectory $\vec{X}(t)$ in phase space, the value of the distribution function at a fixed phase space point $\vec{X}$ will usually vary with time. The set of trajectories $\{\vec{X}(t)\}$ connects regions of phase space to other regions — regions connected by at least one trajectory are called mutually accessible, see Fig. 4. Mutual accessibility, or lack thereof, is usually connected to a conservation law: for instance, in spherically symmetric systems angular momentum is conserved; therefore regions of phase space corresponding to different angular momenta are not mutually accessible. A system that at initial time t_0 is in some neighborhood ΔX_0 of a phase space point X_0 will, as time passes, visit all the phase space regions that are mutually accessible with the initial state. Because different microstates within the initial region ΔX_0 visit the accessible regions in a slightly different order and at a different rate, eventually any memory of the initial state is lost, and the system occupies all accessible regions with equal probability. This final state is known as thermal equilibrium [9], and the corresponding distribution is the equilibrium distribution $f^{\rm eq}(\vec{X})$. Note that the equilibrium distribution is independent of time. The equilibrium distribution is also the distribution that maximizes entropy as will be shown in an exercise session.

Liouville's theorem guarantees that the volume of phase space occupied by a system subject to fixed external constraints does not change. However, the phase space region occupied by the system becomes increasingly distorted as time passes on, and eventually it will reach the neighborhood of every accessible point of the phase space — roughly speaking, the probability density will spread itself more and more evenly over entire accessible phase space. An illuminating example is diffusion of a dye in water: as time passes, dye molecules are more and more evenly spread over the whole container, but the volume occupied by the dye molecules does not change.

An important concept relating to thermal equilibrium is called ergodicity. Ergodicity assumption states that all regions of phase space which correspond to the same energy are mutually accessible. Essentially the ergodicity assumption ignores all conservation laws except for energy conservation. This is frequently a bad assumption in ideal model systems like one planet in a Newtonian gravitational field, but much better an assumption in the real world where the planet interacts with lots of other planets, is bombarded by asteroids etc.. If ergodicity assumption is valid then the equilibrium distribution function $f^{\mathrm{eq}}(\vec{X})$ can only depend on $\vec{X}$ through the energy $E(\vec{X})$. A system that satisfies ergodicity assumption is called *ergodic*, and generally speaking ergodic systems are much easier to deal with than non-ergodic ones.

One important consequence of ergodicity is that if we observe a system that is initially in a particular microstate $\vec{X}_0$, as time passes, it will pass through all other microstates with the same energy, and therefore knowledge of the initial state is lost if we wait sufficiently long. Consequently, the time-average of some quantity A starting from the initial state $\vec{X}_0$,

$$\overline{A}_{\vec{X}_0} = \lim_{t\to\infty} \frac{1}{t} \int_0^t dt' A(\vec{X}(t')), \qquad \vec{X}(0) = \vec{X}_0, \tag{2.9}$$

is independent of $\vec{X}_0$ and is furthermore equal to the equilibrium ensemble average

$$\overline{A}_{\vec{X}_0} = \overline{A} = \langle A \rangle_{\mathrm{eq}} = \int d^{2dN} X \overline{f}_N^{\mathrm{eq}}(\vec{X}) A(\vec{X}). \tag{2.10}$$

The canonical equilibrium distribution function, which is valid for systems with a conserved number of particles, is given by the Gibbs form

$$\overline{f}_N^{\mathrm{eq}}(\vec{X}) = \frac{1}{Z} e^{-E(\vec{X})/k_B T} \tag{2.11}$$

where k_B is the Boltzmann constant and T is temperature. The partition function $Z(T)$ is given by

$$Z(T) = \int d^{2dN} X \, e^{-E(\vec{X})/k_B T} \tag{2.12}$$

and is needed in (2.11) to guarantee proper normalization.

4. Reduced Distribution Functions

The distribution function $\overline{f}_N(\vec{X}, t)$ depends on $(2dN + 1)$ variables, and is therefore not calculable in practice when $N \approx 10^{23}$. On the other hand, it is not particularly interesting, either, since most physical quantities do not depend very sensitively on the individual positions and momenta of all 10^{23} particles. Most of the time we are interested only in knowing how many particles there are in a particular region of space, and how those particles are moving. This information is contained in the reduced one-particle distribution function.

First, let us calculate the probability that particle number 1 is in a particular region of phase space. Since we do not care where all other particles are, we simply integrate over them, and define

$$\overline{f}_1(\vec{r}_1, \vec{p}_1, t) = \int \prod_{j=2}^{n} (d^d r_j d^d p_j) \overline{f}_N(\vec{X}, t) \tag{2.13}$$

(Note that from here on we denote the spatial variable by $\vec{r}$ rather than $\vec{q}$. This is in agreement with the standard notation in classical mechanics where $(\vec{q}, \vec{p})$ are general canonical coordinates, and $(\vec{r}, \vec{p})$ are used in the specific case when $\vec{r}$ is a position vector and $\vec{p}$ the usual momentum.). The probability of finding particle number 1 in the (single particle) phase space element $(\Delta r_1, \Delta p_1)$ is given by $\overline{f}_1(\vec{r}_1, \vec{p}_1, t)\Delta r_1 \Delta p_1$. If we assume that all particles are identical, the more relevant quantity is the number of particles in a phase space volume — we do not care which particles occupy the element. The corresponding distribution is called the one-particle distribution and is given by $f_1(\vec{r}_1, \vec{p}_1, t) = N\overline{f}_1(\vec{r}_1, \vec{p}_1, t)$. In a classical gas or liquid the energy of a particle is given by $\epsilon(\vec{p}) = \frac{p^2}{2m}$ and the equilibrium one-particle distribution function is consequently

$$f_1^{\mathrm{eq}}(\vec{p}) = \frac{N}{V} \frac{1}{(2\pi m k_B T)^{d/2}} e^{-\frac{p^2}{2mk_B T}}. \tag{2.14}$$

Apart from the prefactor N/V this is known as the Maxwell distribution.

Occasionally we are interested in the relative positions of two particles, and then we need the two-particle distribution function. Again we begin by defining the probability density of finding particle number 1 in phase space element $(\Delta r_1, \Delta p_1)$ and particle number 2 in element $(\Delta r_2, \Delta p_2)$, which is given by

$$\overline{f}_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t) = \int \prod_{j=3}^{n} (d^d r_j d^d p_j) \overline{f}_N(\vec{X}, t). \tag{2.15}$$

Since we in general are not concerned with the labels of the particles, it is customary to define the two-particle distribution function $f_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t) = N(N-1)\overline{f}_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t)$ which measures the number of pairs such that the first particle is at $(\vec{r}_1, \vec{p}_1)$ and the second particle is at $(\vec{r}_2, \vec{p}_2)$.

If we are not interested in the momenta of the particles we can further simplify our notation and introduce the one-particle density $n_1(\vec{r}, t) = \int d^d p f_1(\vec{r}, \vec{p}, t)$ and the two-particle density $n_2(\vec{r}_1, \vec{r}_2, t) = \int d^d p_1 d^d p_2 f_2(\vec{r}_1, \vec{p}_1, \vec{r}_2, \vec{p}_2, t)$. At equilibrium we have (assuming that there are no external forces so that the equilibrium state has a uniform density)

$$\left\{ \begin{array}{lcl} n_1^{\mathrm{eq}}(\vec{r}) & = & \frac{N}{V} \\ n_2^{\mathrm{eq}}(\vec{r}_1, \vec{r}_2) & = & \left(\frac{N}{V}\right)^2 g(|\vec{r}_1 - \vec{r}_2|) \end{array} \right. \tag{2.16}$$

where $g(r)$ is called the pair distribution function. The Fourier transform of the pair distribution function is related to another quantity, the static structure factor $S(\vec{q})$ through

$$S(\vec{q})-1=\frac{N}{V}\int d^dr e^{-i\vec{q}\cdot\vec{r}}\left[g(r)-1\right]. \tag{2.17}$$

The structure factor is of great interest since it is directly measurable in scattering measurements; hence, we can use such measurements to determine relative positions of molecules in a liquid or a solid.

All quantities that depend on the coordinates of single particles or the relative positions of two particles can be written in terms of the one- and two-particle densities. As an example, the average potential energy of a system is given by

$$\left\langle\frac{1}{2}\sum_{i\neq j}v(\vec{r}_i-\vec{r}_j)\right\rangle=\frac{1}{2}\int d^dr_1d^dr_2v(\vec{r}_1-\vec{r}_2)n_2(\vec{r}_1,\vec{r}_2)=\frac{1}{2}VS_{d-1}n^2\int_0^\infty drr^{d-1}g(r)v(r) \tag{2.18}$$

where I assumed that the potential energy only depends on distance, $v(\vec{r})=v(r)$. Here S_{d-1} is the surface area of a d-dimensional unit sphere so that $S_2=4\pi$, $S_1=2\pi$ and $S_0=2$ — note that the surface of a d-dimensional region is a $(d-1)$-dimensional surface, which explains why mathematicians prefer to call the (two-dimensional) surface area of a (three-dimensional) unit sphere S_2 rather than S_3. This means that most measurable quantities can be related to $f_1(\vec{r},\vec{p},t)$ or $f_2(\vec{r}_1,\vec{p}_1,\vec{r}_2,\vec{p}_2,t)$. Since these depend only on $(2d+1)$ and $(4d+1)$ variables, respectively, they are much more manageable quantities than the original N-particle distribution function $\overline{f}_N(\vec{X},t)$. We will usually be interested in f_1 and f_2.

We mentioned earlier that those dynamic variables which are not fixed by external constraints fluctuate around an average value. The fluctuations are easy to calculate in equilibrium. As an example we consider a system at a fixed temperature T and calculate the fluctuations in its total energy. Let the system be described by the Hamiltonian H so that the equilibrium distribution function is given by

$$\overline{f}_N^{\mathrm{eq}}(\vec{X})=\frac{1}{Z}e^{-\beta H(\vec{X})}, \tag{2.19}$$

where $\beta=1/(k_BT)$. The average total energy, $\langle H\rangle$, is also known as the internal energy, and is defined by

$$U(T)=\langle H\rangle=\frac{1}{Z}\int d^{2dN}X\,H(\vec{X})e^{-\beta H(\vec{X})}. \tag{2.20}$$

Consequently, the temperature derivative of $U(T)$ is given by

$$\frac{\partial U}{\partial T}=\frac{1}{k_BT^2}\left\{\frac{1}{Z}\int d^{2dN}X\,H^2(\vec{X})e^{-\beta H(\vec{X})}-\left[\frac{1}{Z}\int d^{2dN}X\,H(\vec{X})e^{-\beta H(\vec{X})}\right]^2\right\}. \tag{2.21}$$

Here the right hand side is $\frac{1}{k_BT^2}\left[\langle H^2\rangle-\langle H\rangle^2\right]=\frac{1}{k_BT^2}\langle[H-\langle H\rangle]^2\rangle=\frac{1}{k_BT^2}(\delta E)^2$, where δE is the standard deviation of the total energy. Taking into account the definition of the heat capacity, $c_V=\frac{1}{V}\frac{\partial U}{\partial T}$, we obtain the final result

$$(\delta E)^2=k_BT^2Vc_V \tag{2.22}$$

which shows explicitly that the variance of the total energy is proportional to the system size V (or the particle number N) as we argued before. Also, we see that the variance is proportional to the specific heat, which is a response coefficient telling how the system responds to a changing temperature. This is our first encounter with the fluctuation-dissipation theorem that connects fluctuations $(\langle(\delta E)^2\rangle)$ to response coefficients (c_V).

5. *Wigner Distribution Function*

If we try to generalize the concept of a distribution function like $f_1(\vec{r},\vec{p},t)$ to quantum mechanical systems, we face a fundamental difficulty. In quantum mechanics we cannot simultaneously determine position and momentum to an arbitrary accuracy, and therefore it makes no sense to talk about the probability of finding a particle in an arbitrarily small phase space element. Consequently, much of the previous discussion must be modified considerably if we wish to extend it to quantum mechanical systems.[2]

The quantum mechanical quantity which most closely represents the one-particle distribution function $f_1(\vec{r},\vec{p},t)$ is the Wigner distribution function $f_W(\vec{r},\vec{p},t)$. Note that since $f_W(\vec{r},\vec{p},t)$ depends on both $\vec{r}$ and $\vec{p}$, it is not a measurable observable. The Wigner distribution function is defined by

$$f_W(\vec{R},\vec{p},t) = \int d^d r e^{-i\vec{p}\cdot\vec{r}} \psi^\dagger(\vec{R}-\frac{1}{2}\vec{r})\psi(\vec{R}+\frac{1}{2}\vec{r}) \tag{2.23}$$

where $\psi^\dagger(\vec{r})$ and $\psi(\vec{r})$ are the operators that create and annihilate a particle at position $\vec{r}$, respectively. Integrating over either $\vec{p}$ or $\vec{R}$ we get

$$\begin{aligned} \int \frac{d^d p}{(2\pi)^d} f_W(\vec{R},\vec{p},t) &= \int d^d r \psi^\dagger(\vec{R}-\frac{1}{2}\vec{r})\psi(\vec{R}+\frac{1}{2}\vec{r}) \int \frac{d^d p}{(2\pi)^d} e^{-i\vec{p}\cdot\vec{r}} &&= \psi^\dagger(\vec{R})\psi(\vec{R}) \\ \int d^d R f_W(\vec{R},\vec{p},t) &= \int d^d r \int d^d r' \psi^\dagger(\vec{r})\psi(\vec{r}')e^{-i\vec{p}\cdot\vec{r}} e^{i\vec{p}\cdot\vec{r}'} &&= \psi^\dagger(\vec{p})\psi(\vec{p}). \end{aligned} \tag{2.24}$$

Thus, $f_W(\vec{r},\vec{p},t)$ integrated over all $\vec{p}$ yields the density $n(\vec{r})$, whereas an integral of $f_W(\vec{r},\vec{p},t)$ over $\vec{r}$ yields the density in momentum space — both these properties are valid also for the classical one-particle density $f_1(\vec{r},\vec{p},t)$. However, we should not push the analogy between $f_W(\vec{r},\vec{p},t)$ and $f_1(\vec{r},\vec{p},t)$ too far; for instance, there are often regions in phase space where $f_W(\vec{r},\vec{p},t) < 0$, which would be unacceptable for a proper density distribution like $f_1(\vec{r},\vec{p},t)$.

---

[2] If you are not familiar with the second quantized form of quantum mechanics you may wish to skip over this section.

C. Transport Equations

1. *The Liouville Equation*

Liouville's theorem gives us an equation of motion for the N-particle distribution function $\overline{f}_N(\vec{X},t)$, also known as a transport equation. The theorem states $\frac{d\overline{f}_N}{dt}=0$, which more explicitly (taking into account the time dependence of the positions and momenta of the particles) gives

$$\boxed{\frac{\partial \overline{f}_N}{\partial t}+\sum_{j=1}^{N}\left\{\frac{1}{m}\vec{p}_j\cdot\frac{\partial \overline{f}_N}{\partial \vec{r}_j}+\vec{F}_j\cdot\frac{\partial \overline{f}_N}{\partial \vec{p}_j}\right\}=0} \tag{2.25}$$

where $\vec{F}_j$ is the force acting on the j^{th} particle. This is known as the classical Liouville equation. We derived this form earlier when we proved Liouville's theorem. The first term in the sum describes free flow of particles with velocities $\frac{1}{m}\vec{p}_j$, and the second term describes momentum changes due to forces acting on the particles. In most cases the forces $\vec{F}_j$ are in part due to external forces, and in part to interactions between the particles. If the interactions can be described through interaction potentials $\Phi(\vec{r})$ between pairs of particles we have

$$\vec{F}_j=\vec{F}^{\text{ext}}(\vec{r}_j,t)-\sum_{i\neq j}\frac{\partial\Phi(\vec{r}_j-\vec{r}_i)}{\partial\vec{r}_j}. \tag{2.26}$$

Note that the equilibrium distribution $\overline{f}_N^{\text{eq}}(\vec{X})=\frac{1}{Z}e^{-\beta H(\vec{X})}$ satisfies the Liouville equation (left as an exercise).

The Liouville equation is an exact equation of motion for the N-particle distribution. As such it is impossible to solve since the solution would require the integration of a differential equation with 10^{23} variables. Furthermore, a solution would be useless since it depends on the initial positions and momenta of all the 10^{23} particles. Therefore, the Liouville equation is usually of little practical importance except as a starting point for the derivation of other approximate, and more tractable, transport equations — the exception is study of small systems. The equation is also reversible: there is a one-to-one correspondence between an initial state and a final state. The reversibility of the Liouville equation is a direct consequence of the reversibility of Newton's equations.

2. *Approximations and Irreversibility*

If we integrate the Liouville equation over all positions and momenta except for $\vec{r}_1$ and $\vec{p}_1$ we obtain an equation of motion for the one-particle distribution. The integrations are straightforward, and using expression (2.26) for the force term we obtain

$$\frac{\partial f_1(\vec{r},\vec{p},t)}{\partial t}+\frac{\vec{p}}{m}\cdot\frac{\partial f_1(\vec{r},\vec{p},t)}{\partial\vec{r}}+\vec{F}^{\text{ext}}(\vec{r},t)\cdot\frac{\partial f_1(\vec{r},\vec{p},t)}{\partial\vec{p}}-\int d^dr'd^dp'\frac{\partial\Phi(\vec{r}-\vec{r}')}{\partial\vec{r}}\cdot\frac{\partial f_2(\vec{r},\vec{p},\vec{r}',\vec{p}',t)}{\partial\vec{p}}=0. \tag{2.27}$$

We see that the equation of motion is not closed: to solve it we must know the two-particle distribution $f_2(\vec{r}, \vec{p}, \vec{r}', \vec{p}', t)$. If we try to derive an equation for f_2, we find that solving it requires the knowledge of f_3 etc.. This hierarchical scheme is known as the BBGKY hierarchy (Bogoliubov-Born-Green-Kirkwood-Yvon); it is exact, and therefore useless. We need an approximation to connect the distribution function f_{M+1} for some M to lower order functions $f_1, f_2, \ldots, f_M$. Such an approximation is equivalent to neglecting $(M+1)$-particle correlations which has profound implications.

This is the proper point to discuss reversibility. We have already stated that since Newton's laws are invariant under time reversal, the Liouville equation must also be time-reversal invariant. If an arbitrary sequence of events is possible, then the time-reversed sequence is equally possible. This is in direct contradiction with the second law of thermodynamics which states that entropy does not decrease in any process. Time reversal invariance would imply that if a process which leads to an increase of entropy can occur, then also the reverse process which leads to a decrease of entropy must be possible. Which of the two theories should we disregard, Newtonian mechanics or thermodynamics?

The commonly accepted answer is that we should not disregard either one of them. Instead, we must examine little more carefully the premises leading to the statement that entropy does not decrease. Thermodynamics, and statistical physics, are probabilistic theories, they are concerned with likely events and not rare phenomena. Newtonian mechanics, in contrast, applies to all phenomena, regardless of how common they are.

Again we can use the illuminating example of a dye diffusing in a container full of water. If we place a drop of dye in water at time $t = t_0$, it will quickly disperse over the container. In the course of this diffusion the dye molecules collide frequently with each other as well as with the water molecules. Since entropy is a measure of disorder, we conclude that entropy indeed increases in this process. At a later time $t = t_1$ all dye molecules and water molecules occupy certain locations in phase space, i.e. are at certain positions and have certain velocities. If we then suddenly reverse all the velocities,[3] the molecules will reverse their trajectories in phase space, and at time $t = t_1 + (t_1 - t_0)$ all dye molecules have returned to form the drop that was introduced at time t_0 — in particular, all collisions between molecules have taken place in a time-reversed manner. It appears that a more ordered state has been recovered, and therefore we conclude that entropy has decreased between times t_1 and $t_1 + (t_1 - t_0)$. However, the entropy has decreased only because the initial condition at time t_1 was chosen in a particular way. In general, if we choose arbitrary initial conditions, entropy increases with probability 1 (i.e. almost certainly) and the entropy-decreasing processes are extremely rare. This is in agreement with our experience and intuition that is based on processes that we have observed in nature, i.e. processes that occur comparatively frequently.

If we want to formally rescue the second law of thermodynamics we must define entropy more carefully. The canonical entropy is related to the N-particle distribution function through

---

[3] This feat requires a *deus ex machina* of some sort, which is not a part of the standard equipment in experimental laboratories, but is a handy tool for theorists.

$$S = -k_B \int d^{2dN}X \overline{f}_N(\vec{X},t) \log \left[\overline{f}_N(\vec{X},t)\right], \qquad (2.28)$$

and it is easy to show that $\frac{dS}{dt} = 0$ or that the canonical entropy is a constant in time. The canonical entropy is not a monotonically increasing function with time, it does not obey the second law of thermodynamics, and therefore it is compatible with the underlying time-reversible equations of motion. The canonical entropy is therefore not the entropy that is usually meant in statistical physics.

The conventional entropy is known as the hydrodynamic entropy, and it is obtained from the one-particle distribution function through

$$S = -k_B \int d^d r d^d p f_1(\vec{r},\vec{p},t) \log \left[f_1(\vec{r},\vec{p},t)\right]. \qquad (2.29)$$

The hydrodynamic entropy is obtained from the one-particle distribution, and since in general $\frac{df_1}{dt} \neq 0$, the hydrodynamic entropy is not necessarily a constant of motion. The famous H-theorem by Ludwig Boltzmann states that $\frac{dS}{dt} \geq 0$ where S is the hydrodynamic entropy (we will prove the H-theorem in the next chapter) and the second law of thermodynamics has been rescued. Physically, the difference between using the N-particle distribution function $\overline{f}_N$ and the one-particle distribution function f_1 is due to neglecting correlations between particles. In the example of dye in water, correlations between the dye molecules, as well between dye molecules and water molecules, at time t_1 were crucial to cause all the necessary collisions which in turn caused a dye droplet to form at time $t_1 + (t_1 - t_0)$. These correlations can be thought of as a special type of order, so that contrary to our previous conclusion, no order was actually created in the time-reversed evolution — in agreement with the statement that the canonical entropy does not change with time. Most experiments, however, are not sufficiently detailed to specify the correlations in the initial state (i.e. to specify $\overline{f}_N(\vec{X},t_0)$), and therefore the hydrodynamic entropy is more relevant than the canonical entropy. Consequently, we frequently make the approximation of ignoring multiparticle correlations. This assumption, or approximation, takes us from time-reversal invariant Newtonian mechanics to thermodynamics or statistical mechanics which is not invariant under time reversal.

3. Boltzmann Equation

a. Justification of the Boltzmann Equation. The Boltzmann equation is one approximation to the transport equation (2.27). It was introduced by Ludwig Boltzmann in 1872 to discuss dynamics of ordinary gases. It is meant to be applicable in cases when the interactions between particles are very short range but strong — the interactions are usually called collisions. Some further approximations are also needed: (i) a particle interacts with at most one other particle at a time (dilute gas approximation: only binary collisions are considered), (ii) collisions are instantaneous, i.e. the duration of a collision is much shorter than any other relevant time scale, (iii) collisions are local in space, i.e. the extent of the interactions is much shorter than any other length scale. Furthermore, we ignore all internal structure of the particles and assume that the collisions are elastic.

In the Boltzmann equation the fundamental dynamic process is a binary collision between two point particles. In a collision event two particles with momenta $\vec{p}$ and $\vec{p}_1$ come together

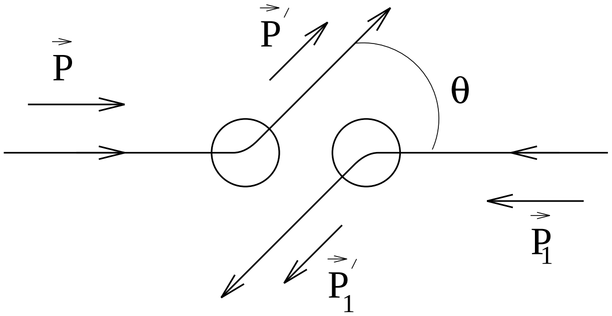


FIG. 5. A binary collision event describing the scattering of an initial state $(\vec{p}, \vec{p}_1)$ to a final state $(\vec{p}\,', \vec{p}_1\!')$.

and exchange momentum so that after the collision their momenta are $\vec{p}\,'$ and $\vec{p}_1\!'$, respectively, see Fig. 5. Between collisions the particles move under the influence of the external fields, usually along straight lines. If the radius of interaction is R and the density of particles is n, the distance ℓ a particle travels between two collisions in the average is given by $\ell = \frac{1}{n\pi R^2}$ in 3D, by $\ell = \frac{1}{nR}$ in 2D, and by $\ell = \frac{1}{n}$ in 1D. This distance is called the mean free path, and the dilute gas approximation requires $\ell \gg R$. We also assume that the mean free path is much shorter than the size of the system, $\ell \ll L$. This is known as the diffusive regime; the dynamics of the gas is quite different in the ballistic regime when $\ell \gtrsim L$ (so-called Knudsen gas).

The last two approximations leading to the Boltzmann equation mean that the equation can only be used to describe phenomena that vary slowly in space and time compared to R and the collision time τ_c. The collision time can be estimated as $\tau_c \approx R/\sqrt{\langle v^2 \rangle} \approx R\sqrt{m/k_BT}$. Therefore, we need to be careful when applying the Boltzmann equation to inherently fast phenomena like light scattering, especially at low temperatures.

The Boltzmann equation is usually written in the form

$$\boxed{\frac{\partial f_1}{\partial t} + \frac{1}{m}\vec{p} \cdot \frac{\partial f_1}{\partial \vec{r}} + \vec{F}^{\text{ext}}(\vec{r}, t) \cdot \frac{\partial f_1}{\partial \vec{p}} = \left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}}} \tag{2.30}$$

where the term on the right hand side is known as the Boltzmann collision term and is a particular approximation to the last term in Eq. (2.27). Terms on the left hand side describe a smooth collisionless evolution of the system and are known as drift terms.

We shall now motivate the commonly used expression for the collision term. More rigorous derivation is possible but beyond the scope of this course. The collision term describes flow of particles to and from a volume element in phase space. Since collisions are local in space, we only need to consider the changes in momenta due to collisions. The distribution function $f_1(\vec{r}, \vec{p}, t)$ changes in part if a particle with momentum $\vec{p}$ collides with another particle, and scatters to a different momentum state, or if two particles with momenta different from $\vec{p}$ collide and after the collision one of them has momentum equal to $\vec{p}$. These scattering-out and scattering-in processes give

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} \Delta\vec{p} = -\int_{\Delta\vec{p}} d^dp \int d^dp_1 \int d^dp' \int d^dp_1' f_1(\vec{r},\vec{p},t) f_1(\vec{r},\vec{p}_1,t) W(\vec{p},\vec{p}_1 \to \vec{p}\,',\vec{p}_1') + \int d^dp' \int d^dp_1' \int_{\Delta\vec{p}} d^dp \int d^dp_1 f_1(\vec{r},\vec{p}\,',t) f_1(\vec{r},\vec{p}_1',t) W(\vec{p}\,',\vec{p}_1' \to \vec{p},\vec{p}_1) \tag{2.31}$$

where $W(\vec{p},\vec{p}_1 \to \vec{p}\,',\vec{p}_1')$ is the probability for transitions from $(\vec{p},\vec{p}_1)$ to $(\vec{p}\,',\vec{p}_1')$ per unit time.

We have assumed that the probability of finding one particle with momentum $\vec{p}$ and another with momentum $\vec{p}_1$ is proportional to $f_1(\vec{r},\vec{p},t)f_1(\vec{r},\vec{p}_1,t)$, an assumption known as the *Stoßzahlansatz*. In terms of the two-particle distribution this corresponds to the approximation $\overline{f}_2(\vec{r},\vec{p},\vec{r},\vec{p}_1,t) \approx \overline{f}_1(\vec{r},\vec{p},t)\overline{f}_1(\vec{r},\vec{p}_1,t)$, which means that we are ignoring any correlations between the two particles. This assumption renders the Boltzmann equation irreversible as we have discussed before.

The time-reversal invariance of Newton's equations implies that the rates $W(\vec{p},\vec{p}_1 \to \vec{p}\,',\vec{p}_1')$ are also invariant under time reversal [5]. Therefore, we must have

$$W(\vec{p},\vec{p}_1 \to \vec{p}\,',\vec{p}_1') = W(\vec{p}\,',\vec{p}_1' \to \vec{p},\vec{p}_1). \tag{2.32}$$

Inserting this result to the expression of the collision term we have (taking the limit $\Delta\vec{p} \to 0$)

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} = -\int d^dp_1 \int d^dp' \int d^dp_1' W(\vec{p},\vec{p}_1 \to \vec{p}\,',\vec{p}_1') \left[f_1(\vec{r},\vec{p},t)f_1(\vec{r},\vec{p}_1,t) - f_1(\vec{r},\vec{p}\,',t)f_1(\vec{r},\vec{p}_1',t)\right]. \tag{2.33}$$

The conservation laws for energy and momentum imply connections between the four momenta which participate in a particular collision. These connections can be used to parameterize the function W differently in terms of the relative incident velocity $\vec{v} = (\vec{p} - \vec{p}_1)/m$ and the angle Ω between the two outgoing momenta $\vec{p}\,'$ and $\vec{p}_1'$. The new parameterization is given in terms of a differential cross section $\frac{d\sigma(\Omega,v)}{d\Omega}$ which is defined as

$$nv\frac{d\sigma(\Omega,v)}{d\Omega}d\Omega = \text{Number of particles scattered into the solid angle } d\Omega \text{ about the direction } \Omega \text{ per unit time}$$

where n is the density of particles. In terms of the differential cross section the collision term is

$$\left(\frac{\partial f_1}{\partial t}\right)_{\text{coll}} = -\int d^dp_1 \int d\Omega v \frac{d\sigma(\Omega,v)}{d\Omega} \left[f_1(\vec{r},\vec{p},t)f_1(\vec{r},\vec{p}_1,t) - f_1(\vec{r},\vec{p}\,',t)f_1(\vec{r},\vec{p}_1',t)\right]. \tag{2.34}$$

Note that in equation (2.33) there are 3 d-dimensional integrals whereas in equation (2.34) there is 1 d-dimensional integral over $\vec{p}_1$ and 1 $(d-1)$-dimensional integral over Ω. The difference is due to the fact that the differential cross section is constructed to obey momentum and energy conservation, which fix $(d+1)$ degrees of freedom. This leaves exactly the $3d - (d+1) = 2d - 1$ degrees of freedom that are integrated over in equation (2.34).

The differential cross section $\frac{d\sigma(\Omega,v)}{d\Omega}$ must be obtained from a specific (usually quantum mechanical) model. Once it is known, the Boltzmann equation (2.30) with the collision term (2.34) forms a non-linear integro-differential equation for the distribution function $f_1(\vec{r},\vec{p},t)$. Even with all the simplifying approximations we have made the task of solving the transport equation is still quite formidable. Solving the Boltzmann equation in all but the simplest cases is only possible numerically, and even then it is quite laborious. In the following we will make further simplifying assumptions regarding the collision term and extract some relevant physical results.

b. Local Thermal Equilibrium. First we will show that the Boltzmann equation implies that the hydrodynamic entropy is an increasing function of time. Since from now on we will only deal with the one-particle distribution function, we will for brevity drop the subscript 1 and from now on denote $f_1(\vec{r},\vec{p},t)$ by $f(\vec{r},\vec{p},t)$. We start by defining a quantity that Boltzmann himself called[4] H,

$$H(\vec{r},t) = \int d^d p f(\vec{r},\vec{p},t)\log[f(\vec{r},\vec{p},t)] \tag{2.35}$$

which is related to the hydrodynamic entropy through

$$S(t) = -k_B \int d^d r H(\vec{r},t) \tag{2.36}$$

(hence, $H(\vec{r},t)$ is entropy density apart from the factor $-k_B$). Boltzmann's H-theorem states $\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} \leq 0$, that is, collisions reduce the function H. To prove this we take the derivative of H, which gives

$$\frac{\partial H}{\partial t} = \int d^d p \left[\frac{\partial f}{\partial t}\log(f) + \frac{\partial f}{\partial t}\right] = \int d^d p \left(\frac{\partial f}{\partial t}\right)_{\text{drift}} [\log(f)+1] + \int d^d p \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \log(f) \tag{2.37}$$

where we used the fact that $\int d^d p \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = \left(\frac{\partial}{\partial t}\int d^d p f(\vec{r},\vec{p},t)\right)_{\text{coll}} = \left(\frac{\partial}{\partial t}n(\vec{r},t)\right)_{\text{coll}} = 0$. Thus, we have

$$\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = \int d^d p \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \log(f). \tag{2.38}$$

Inserting the expression for the collision term and suppressing the space and time variables we have

$$\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = -\int d^d p \int d^d p_1 \int d\Omega v \frac{d\sigma(\Omega,v)}{d\Omega}[f(\vec{p})f(\vec{p_1}) - f(\vec{p}\,')f(\vec{p}_1')]\log(f(p)). \tag{2.39}$$

Interchanging $\vec{p}$ and $\vec{p_1}$, adding up the two equations, and dividing by 2 gives (using $\log(f(\vec{p})) + \log(f(\vec{p_1})) = \log[f(\vec{p})f(\vec{p_1})]$)

---

[4] Note that this H is not the Hamiltonian!

$$\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = -\frac{1}{2}\int d^d p \int d^d p_1 \int d\Omega v \frac{d\sigma(\Omega, v)}{d\Omega}\left[f(\vec{p})f(\vec{p}_1) - f(\vec{p}\,')f(\vec{p}_1')\right]\log[f(\vec{p})f(\vec{p}_1)]. \tag{2.40}$$

Now interchanging $(\vec{p}, \vec{p}_1) \leftrightarrow (\vec{p}\,', \vec{p}_1')$, adding up the two equations and dividing by two yields

$$\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = -\frac{1}{4}\int d^d p \int d^d p_1 \int d\Omega v \frac{d\sigma(\Omega, v)}{d\Omega}\left[f(\vec{p})f(\vec{p}_1) - f(\vec{p}\,')f(\vec{p}_1')\right]\log\left[\frac{f(\vec{p})f(\vec{p}_1)}{f(\vec{p}\,')f(\vec{p}_1')}\right]. \tag{2.41}$$

Using the result $(x - y)\ln(x/y) \geq 0$ for all positive x, y (where the equality holds only for $x = y$), we have the desired result $\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} \leq 0$.

The equality holds only for $f(\vec{p})f(\vec{p}_1) = f(\vec{p}\,')f(\vec{p}_1')$ in which case $\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = 0$ and the distribution $f(\vec{r}, \vec{p}, t)$ is unchanged by collisions. Expanding the distribution function as

$$f(\vec{r}, \vec{p}, t) = \exp\left(\sum_{j=0}^{\infty} a_j(\vec{r}, t)(\vec{p})^j\right), \tag{2.42}$$

where $a_j(\vec{r}, t)$ are arbitrary functions[5] we find that the condition $\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = 0$ implies

$$\exp\left(\sum_{j=0}^{\infty} a_j(\vec{r}, t)\left[(\vec{p})^j + (\vec{p}_1)^j\right]\right) = \exp\left(\sum_{j=0}^{\infty} a_j(\vec{r}, t)\left[(\vec{p}\,')^j + (\vec{p}_1')^j\right]\right). \tag{2.43}$$

Conservation of momentum and energy imply that $\vec{p} + \vec{p}_1 = \vec{p}\,' + \vec{p}_1'$ and $p^2 + (p_1)^2 = (p')^2 + (p_1')^2$ so that a general solution for f is given by $a_j(\vec{r}, t) = 0$ for $j > 2$. The function $a_0(\vec{r}, t)$ can be determined by the normalization condition, so we have the result that $\left(\frac{\partial H}{\partial t}\right)_{\text{coll}} = 0$ or $\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = 0$ implies

$$f^{\text{l.eq.}}(\vec{r}, \vec{p}, t) = \frac{n(\vec{r}, t)}{[2\pi m k_B T(\vec{r}, t)]^{d/2}} \exp\left\{-\frac{[\vec{p} - m\vec{u}(\vec{r}, t)]^2}{2 m k_B T(\vec{r}, t)}\right\}. \tag{2.44}$$

This is known as the *local thermal equilibrium* solution. It is characterized by the local density $n(\vec{r}, t)$, local temperature $T(\vec{r}, t)$, and local flow velocity $\vec{u}(\vec{r}, t)$, which are known as hydrodynamic variables. They are related to the distribution function through

$$\begin{aligned} n(\vec{r}, t) &= \int d^d p f(\vec{r}, \vec{p}, t) \\ n(\vec{r}, t)\vec{u}(\vec{r}, t) &= \int d^d p \frac{\vec{p}}{m} f(\vec{r}, \vec{p}, t) \\ \frac{d}{2} k_B n(\vec{r}, t) T(\vec{r}, t) &= \int d^d p \frac{[\vec{p} - m\vec{u}(\vec{r}, t)]^2}{2m} f(\vec{r}, \vec{p}, t). \end{aligned} \tag{2.45}$$

---

[5] Note that for an odd j the function $a_j(\vec{r}, t)$ must be vector valued. Products between two vectors in this expansion are taken to be scalar products.

The local thermal equilibrium distribution is not changed by the collision term, but the drift terms on the left hand side of the Boltzmann equation do change $f^{\text{l.eq.}}(\vec{r},\vec{p},t)$, and drive the system towards a full equilibrium.

c. Applications of the Boltzmann Equation. In its general integro-differential form the Boltzmann equation is quite difficult to solve. A common, drastic assumption is to replace the collision term by a simple relaxation mechanism like

$$\left(\frac{\partial f}{\partial t}\right)_{\text{coll}} = -\frac{1}{\tau}\left[f(\vec{r},\vec{p},t) - f^{\text{l.eq.}}(\vec{r},\vec{p},t)\right] \tag{2.46}$$

where τ is a relaxation time that drives the system towards a local thermal equilibrium. Occasionally we go one step further and replace the local thermal equilibrium distribution function here by the full equilibrium distribution (2.14). Note that in the relaxation time approximation the Boltzmann equation is a partial differential equation, whereas with the full collision term it is an integro-differential equation — thus, the simplification is quite substantial! We will now apply the Boltzmann equation to calculate the conductivity tensor of a 3D electron gas in constant electric and magnetic fields.

Our starting point is the Boltzmann equation for a charged particle in the relaxation time approximation

$$\left[\frac{\partial}{\partial t} + \frac{\vec{p}}{m}\cdot\frac{\partial}{\partial \vec{r}} + e\left(\vec{E} + \frac{\vec{p}}{m}\times\vec{B}\right)\cdot\frac{\partial}{\partial \vec{p}}\right] f(\vec{r},\vec{p},t) = -\frac{1}{\tau}\left[f - f^{\text{eq}}\right]. \tag{2.47}$$

Since we are interested in the static linear conductivity, we want to find the distribution function to linear order in magnetic and electric fields. To keep track of the order of various quantities we use the same trick that is often used in quantum mechanics in deriving perturbation theory: we replace $\vec{E} \to \lambda_E \vec{E}$ and $\vec{B} \to \lambda_B \vec{B}$, write $f = f^{\text{eq}} + \lambda_E f^{(E)} + \lambda_B f^{(B)} + \lambda_E\lambda_B f^{(EB)}$, and keep track of only terms up to the order $\lambda_E\lambda_B$. At the end of the calculation we set $\lambda_E = \lambda_B = 1$ to obtain the desired result.

The equilibrium distribution is given by

$$f^{\text{eq}}(p) = \frac{n}{(2\pi m k_B T)^{3/2}} e^{-\beta\frac{p^2}{2m}} \tag{2.48}$$

where n is the density of the electron gas. The momentum derivative (gradient) of $f^{\text{eq}}(p)$ is given by

$$\frac{\partial f^{\text{eq}}}{\partial \vec{p}} = -\beta\frac{\vec{p}}{m} f^{\text{eq}}(p). \tag{2.49}$$

Since all external fields and $f^{\text{eq}}(p)$ are static (no time dependence), the first term $\frac{\partial f}{\partial t}$ drops out from the Boltzmann equation and we have

$$\begin{aligned}
&\frac{\vec{p}}{m}\cdot\frac{\partial}{\partial \vec{r}}\left(\lambda_E f^{(E)} + \lambda_B f^{(B)} + \lambda_B\lambda_E f^{(EB)}\right) \\
+\ & e\left[\lambda_E \vec{E}\cdot\left(-\beta\frac{\vec{p}}{m}\right) f^{\text{eq}}(p) + \lambda_E\lambda_B \vec{E}\cdot\frac{\partial f^{(B)}}{\partial \vec{p}}\right. \\
&\left. + \lambda_B \frac{\vec{p}}{m}\times\vec{B}\cdot\left(-\beta\frac{\vec{p}}{m}\right) f^{\text{eq}}(p) + \lambda_E\lambda_B\frac{\vec{p}}{m}\times\vec{B}\cdot\frac{\partial f^{(E)}}{\partial \vec{p}}\right] \\
=\ & -\frac{1}{\tau}\left[\lambda_E f^{(E)} + \lambda_B f^{(B)} + \lambda_B\lambda_E f^{(EB)}\right].
\end{aligned} \tag{2.50}$$

Since all fields are constant in space, also the spatial derivatives drop out. Setting both sides equal to each other order by order we obtain

$$-\beta\frac{\vec{p}}{m}\cdot E f^{\mathrm{eq}}(p) = -\frac{1}{\tau e} f^{(E)}(\vec{p}) \tag{2.51a}$$

$$0 = -\frac{1}{\tau e} f^{(B)}(\vec{p}) \tag{2.51b}$$

$$\vec{E}\cdot\frac{\partial f^{(B)}}{\partial\vec{p}} + \frac{\vec{p}}{m}\times\vec{B}\cdot\frac{\partial f^{(E)}}{\partial\vec{p}} = -\frac{1}{\tau e} f^{(EB)}(\vec{p}). \tag{2.51c}$$

Solving this set of equations gives

$$f^{(E)}(\vec{p}) = \beta\tau e\frac{\vec{p}}{m}\cdot\vec{E} f^{\mathrm{eq}}(p) \tag{2.52a}$$

$$f^{(B)}(\vec{p}) = 0 \tag{2.52b}$$

$$f^{(EB)}(\vec{p}) = -\beta\frac{\tau^2 e^2}{m^2}\vec{p}\times\vec{B}\cdot\vec{E} f^{\mathrm{eq}}(p). \tag{2.52c}$$

The electric current density is given by $\vec{j} = \int d^3p\frac{e\vec{p}}{m}f(\vec{r},\vec{p},t)$, so now we have (setting $\lambda_E = \lambda_B = 1$)

$$\vec{j} = \beta\frac{\tau e^2}{m^2}\int d^3p\,\vec{p}(\vec{p}\cdot\vec{E})f^{\mathrm{eq}}(p) - \beta\frac{\tau^2 e^3}{m^3}\int d^3p\,\vec{p}(\vec{p}\times\vec{B}\cdot\vec{E})f^{\mathrm{eq}}(p). \tag{2.53}$$

The integrals are easiest to do by components, and using

$$\int_{-\infty}^{\infty} dx e^{-\frac{1}{2}\frac{x^2}{\sigma^2}} = \sqrt{2\pi\sigma^2} \tag{2.54a}$$

$$\int_{-\infty}^{\infty} dx x^2 e^{-\frac{1}{2}\frac{x^2}{\sigma^2}} = \sqrt{2\pi\sigma^2}\,\sigma^2 \tag{2.54b}$$

we obtain

$$\vec{j} = \frac{ne^2\tau}{m}\left(\vec{E} + \frac{e\tau}{m}\vec{E}\times\vec{B}\right). \tag{2.55}$$

Here we recognize the prefactor as the Drude conductivity $\sigma_0 = \frac{ne^2\tau}{m}$, so defining a new vector $\vec{\omega}_c = \frac{eB}{m}\hat{B} = \omega_c\hat{B}$, where $\omega_c = \frac{eB}{m}$ is the cyclotron frequency, we find that the current is given by

$$\vec{j} = \sigma_0\left(\vec{E} + \tau\vec{E}\times\vec{\omega}_c\right). \tag{2.56}$$

Writing the right hand side in component form we obtain the conductivity tensor

$$\overset{\leftrightarrow}{\sigma} = \sigma_0\begin{pmatrix} 1 & \tau\omega_{cz} & -\tau\omega_{cy} \\ -\tau\omega_{cz} & 1 & \tau\omega_{cx} \\ \tau\omega_{cy} & -\tau\omega_{cx} & 1 \end{pmatrix} \tag{2.57}$$

We can treat other transport properties in a similar fashion. For instance, the heat current in three dimensions is given by

$$\vec{J}_Q(\vec{r},t) = \int d^3p\vec{v}\left(\frac{p^2}{2m} - \frac{5}{2}k_BT\right) f(\vec{r},\vec{p},t) \tag{2.58}$$

and the thermal conductivity λ is defined as

$$\vec{J}_Q(\vec{r},t) = -\lambda\nabla T(\vec{r},t). \tag{2.59}$$

Using the relaxation time approximation you can derive the result $\lambda = \frac{5}{2}\frac{nk_B^2T\tau}{m}$. Note that in this case you must consider relaxation towards a local equilibrium since the temperature is position dependent.

4. Langevin equation

a. Brownian Motion. In 1828 the botanist Robert Brown reported observing the erratic motion of pollen particles on the surface of water. He noticed that particles remained at rest for long periods of time, but would occasionally jump an arbitrary distance in a random direction. In modern language we say that the pollen particles perform a random walk. Random walks are important not only in physics but in many other fields as well — one application that has recently become quite popular is the study of share prices in the stock market which jump seemingly randomly from one day to another.

The explanation of Brownian motion is due to Einstein in 1905. We will follow the formulation by Paul Langevin (1908). We assume that the particle that performs a random walk is subject to a *stochastic force* $\vec{\zeta}(t)$ which varies randomly with time, see Fig. 6. We also assume that there is some damping due to friction or viscosity so that the equation of motion for the particle is

$$\boxed{m\dot{v}(t) + m\eta v(t) = \zeta(t).} \tag{2.60}$$

For simplicity we will consider only the one-dimensional case. If there are any conventional (non-stochastic) forces, we include them on the left hand side of the equation of motion as usual. The equation (2.60) is known as the Langevin equation. It is an equation of motion for one particle, and as such fundamentally different from the Boltzmann equation which gives the time evolution of the distribution $f(r,p,t)$.

Note that the Langevin equation is irreversible due to the presence of the dissipation term $\eta v(t)$ which, in the absence of the random force, drives the velocity towards zero. Hence, there is a competition between the stochastic force, which gives the system random "kicks", and the dissipation term, which tries to bring the velocity to zero.

Before we attempt to solve the Langevin equation we will consider the stochastic force in some detail and introduce some useful concepts for describing time-dependent stochastic variables in general. We assume that the stochastic force has no preferred direction, so that its average over time vanishes, $\overline{\zeta(t)} = 0$. Time averages are somewhat cumbersome to deal with, and it is more convenient to consider an ensemble of systems which are otherwise

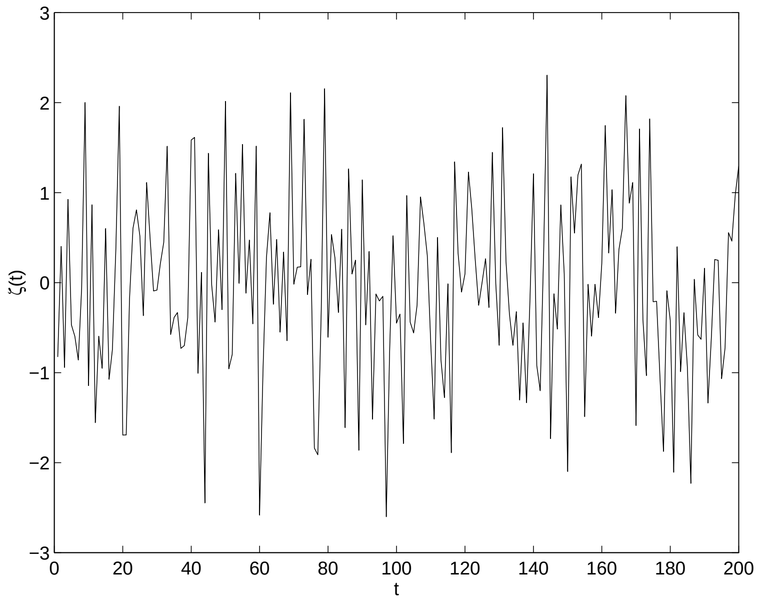


FIG. 6. Time-dependent stochastic force $\zeta(t)$ that results in the noisy dynamics describe the Langevin equation.

identical but have different stochastic forces. We can then view $\zeta(t)$ as a random variable which has some probability distribution. Let us denote the average with respect to this probability distribution by $\langle \ldots \rangle$. Our statement that $\zeta(t)$ has no preferred direction implies

$$\langle \zeta(t) \rangle = 0. \tag{2.61}$$

We will also assume that the random force is uncorrelated, that is, the $\zeta(t)$ is independent of $\zeta(t')$ for $t \neq t'$. Mathematically this is expressed as

$$\langle \zeta(t)\zeta(t') \rangle = A\delta(t-t') \tag{2.62}$$

where A is some constant that gives the strength of the random force. Often the random force $\zeta(t)$ is due to a sequence of many independent events which are not observable (e.g. collisions between a large pollen grain and a small water molecule), in which case the central limit theorem implies that the probability distribution of $\zeta(t)$ is Gaussian,

$$P[\zeta(t)] = \frac{1}{\sqrt{2\pi A}} e^{-\frac{1}{2A}\int dt\, \zeta^2(t)} \tag{2.63}$$

Random forces like $\zeta(t)$ give rise to noisy behavior in the observables and are therefore known as noise sources. A noise source is characterized by its power spectrum. The power spectrum — also called the spectral function — of a stochastic variable $\xi(t)$ is defined as

$$f_\xi(\omega) = \lim_{T_0 \to \infty} \frac{1}{2T_0} \left| \int_{-T_0}^{T_0} dt\, e^{i\omega t} \xi(t) \right|^2 \tag{2.64}$$

which can in equilibrium be simplified to give

$$f_\xi(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t} \langle \xi(t)\xi(0)\rangle_{\mathrm{eq}}. \tag{2.65}$$

This result is known as the Wiener-Khintchine theorem. We can prove it as follows:

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \left| \int_{-T_0}^{T_0} dt'\, e^{i\omega t'} \xi(t') \right|^2 \\ &= \int_{-T_0}^{T_0} dt'\, \xi(t') \int_{-T_0}^{T_0} dt''\, \xi(t'') \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{i\omega(t'-t''-t)} \\ &= \int_{\max(-T_0,-T_0+t)}^{\min(T_0,T_0+t)} dt'\, \xi(t')\xi(t'-t) \end{aligned} \tag{2.66}$$

Dividing both sides by $2T_0$ and taking the limit $T_0 \to \infty$ gives

$$\int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} f_\xi(\omega) = \lim_{T_0\to\infty} \left[\frac{1}{2T_0} \int_{-T_0}^{T_0} dt'\, \xi(t')\xi(t'-t) + \mathcal{O}\left(t \frac{\xi(T_0) + \xi(-T_0)}{2T_0} \right) \right] \tag{2.67}$$

Now, if $\lim_{T\to\infty}(\xi(T)/T) = 0$, the right hand side of this expression is the time average of $\xi(t'+t)\xi(t')$ which is in equilibrium equal to the thermal average. Thus, we have the result that the equilibrium power spectrum of a fluctuating variable is given by the Fourier transform of the thermal average of the autocorrelation function.

An alternative, often useful, form of the Wiener-Khintchine theorem can be derived by considering the correlation function $C_\xi(t,t') = \langle \xi(t)\xi(t')\rangle$. If we write $\xi(t)$ and $\xi(t')$ in terms of the Fourier transform $\xi(\omega)$, we have

$$C_\xi(t,t') = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} e^{-i\omega t} e^{-i\omega' t'} \langle \xi(\omega)\xi(\omega')\rangle. \tag{2.68}$$

Introducing $T = \frac{1}{2}(t+t')$ and $\tau = t - t'$ we get

$$C_\xi(T,\tau) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} e^{-i(\omega+\omega')T} e^{-i\frac{1}{2}(\omega-\omega')\tau} \langle \xi(\omega)\xi(\omega')\rangle. \tag{2.69}$$

If the correlation function $C_\xi(T,\tau)$ does not depend on the time T, as is the case e.g. in thermal equilibrium, the right hand side of this equation cannot depend on T, either. That is only possible if the average $\langle \xi(\omega)\xi(\omega')\rangle$ is proportional to $\delta(\omega+\omega')$. Changing variables in the frequency integrations to $\frac{1}{2}(\omega+\omega')$ and $\omega - \omega'$ yields the desired result

$$\langle \xi(\omega)\xi(\omega')\rangle_{\mathrm{eq}} = C_\xi^{\mathrm{eq}}(\omega) 2\pi\delta(\omega+\omega'). \tag{2.70}$$

Note that the first form of the Wiener-Khintchine theorem states that $C_\xi^{\mathrm{eq}}(\omega) = f_\xi(\omega)$, i.e. the power spectrum $f_\xi(\omega)$ is equal to the correlation function in Fourier space. The equation (2.70) then relates the power spectrum to the Fourier transform of the stochastic variable itself.

In the specific case of a fluctuating force the autocorrelation function is $\langle \zeta(t)\zeta(t')\rangle = A\delta(t-t')$ and its Fourier transform is independent of frequency, $f_\zeta(\omega) = A$. This kind of a noise which is frequency independent is called *white noise.*

b. Correlation Functions. We are now ready to solve Langevin's equation. We will do that in two different ways, first by solving it directly, and then by employing Wiener-Khintchine theorem.

The Langevin equation (2.60) is a first order ordinary differential equation. Its solution is given by

$$v(t) = v(0)e^{-\eta t} + \frac{1}{m}\int_0^t dt'\, e^{-\eta(t-t')}\zeta(t') \tag{2.71}$$

as you can prove by direct substitution. Notice that the memory of the initial velocity $v(0)$ is lost after a time of the order η^{-1}. The velocity-velocity correlation function $C_v(t,0) = \langle v(t)v(0)\rangle$ is now given by

$$C_v(t,0) = v^2(0)e^{-\eta t} + \frac{1}{m}\int_0^t dt'\, e^{-\eta(t-t')}v(0)\langle\zeta(t')\rangle = v^2(0)e^{-\eta t} \tag{2.72}$$

which confirms the statement that knowledge of initial state is lost after time η^{-1}. The average squared velocity at time t is given by

$$\begin{aligned}\langle v^2(t)\rangle &= v^2(0)e^{-2\eta t} + \frac{2}{m}\int_0^t dt'\, e^{-\eta(2t-t')}v(0)\langle\zeta(t')\rangle \\ &\quad + \frac{1}{m^2}\int_0^t dt' \int_0^t dt''\, e^{-\eta(2t-t'-t'')}\langle\zeta(t')\zeta(t'')\rangle \\ &= v^2(0)e^{-2\eta t} + \frac{A}{m^2}\int_0^t dt'\, e^{-2\eta(t-t')} \\ &= v^2(0)e^{-2\eta t} + \frac{A}{2m^2\eta}\left[1 - e^{-2\eta t}\right]\end{aligned} \tag{2.73}$$

Thus, we get

$$\lim_{t\to\infty}\langle v^2(t)\rangle = \frac{A}{2m^2\eta}. \tag{2.74}$$

If the system approaches equilibrium as $t \to \infty$, the average kinetic energy $\frac{1}{2}m\langle v^2\rangle$ is by equipartition theorem given by $\frac{1}{2}k_BT$. This implies that $\langle v^2\rangle = k_BT/m$ and therefore we must have

$$\frac{A}{2m^2\eta} = \frac{k_BT}{m} \tag{2.75}$$

or

$$A = 2\eta k_BTm \tag{2.76}$$

if the Langevin equation describes a system that relaxes to equilibrium as $t \to \infty$. This is one form of the fluctuation-dissipation theorem: the magnitude of fluctuations, A, is proportional to temperature multiplied by the dissipation coefficient η.

An alternative way of solving Langevin's equation is to Fourier transform it first. That gives

$$-i\omega m v(\omega) + m\eta v(\omega) = \zeta(\omega) \tag{2.77}$$

or

$$v(\omega) = \frac{\zeta(\omega)}{-i\omega m + m\eta}. \tag{2.78}$$

Fourier transforming yields a solution that is independent of boundary conditions at $t = 0$, so that correlation functions like $C_v(t, t')$ only depend on the time difference $t - t'$. We can therefore apply the Wiener-Khintchine theorem both to the velocity $v(t)$ and to the noise $\zeta(t)$. From the second form of the Wiener-Khintchine theorem we know that $C_v(\omega) = \int_{-\infty}^{\infty} \frac{d\omega'}{2\pi} \langle v(\omega)v(\omega')\rangle_{\text{eq}}$, so we form the product $v(\omega)v(\omega')$ and take the ensemble average. The only term affected by the averaging is $\zeta(\omega)\zeta(\omega')$, which, according to the Wiener-Khintchine theorem is $2\pi\delta(\omega + \omega')f_\zeta(\omega)$. Substituting this into the integral shows that the velocity correlation function $C_v(\omega)$ is given by $\langle v(\omega)v(-\omega)\rangle_{\text{eq}}$, i.e.

$$C_v(\omega) = \frac{f_\zeta(\omega)}{m^2(\omega^2 + \eta^2)} \tag{2.79}$$

(this is only valid if the system is in equilibrium). Since $\langle v^2\rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} C_v(\omega)$ we have immediately

$$\langle v^2\rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{A}{m^2(\omega^2 + \eta^2)} = \frac{A}{2\pi m^2 \eta} \int_{-\infty}^{\infty} \frac{dx}{1 + x^2} = \frac{A}{2m^2\eta} \tag{2.80}$$

which gives the fluctuation-dissipation theorem as we saw previously. The combination of the Wiener-Khintchine theorem and a Fourier transform is often a powerful technique in the analysis of stochastic differential equations like the Langevin equation.

c. Johnson noise. One consequence of the fluctuation-dissipation theorem is that if the system has some dissipation, i.e. $\eta \neq 0$, it must also show fluctuations. This result is quite general and has several applications. A particular example is an electric circuit which has non-zero resistance: due to the fluctuation-dissipation theorem the current through the circuit will fluctuate (assuming voltage bias). This results in a temperature-dependent noise called Johnson noise which was discovered experimentally in 1928. The origin of the noise was analyzed by Harry Nyquist in the same year.

Let us consider the RLC circuit depicted in Fig. 7 and described by the equation

$$L\frac{dI}{dt} + RI + \frac{1}{C}\int^{t} dt'\, I(t') = V_N(t). \tag{2.81}$$

Here $V_N(t)$ is a noise voltage acting across the circuit. From the fluctuation-dissipation theorem we expect that the noise power spectrum is proportional to the temperature and the dissipative resistance R. Fourier transforming we get

$$I(\omega) = \frac{V_N(\omega)}{iL\omega + R - i\frac{1}{C\omega}} \tag{2.82}$$

and therefore

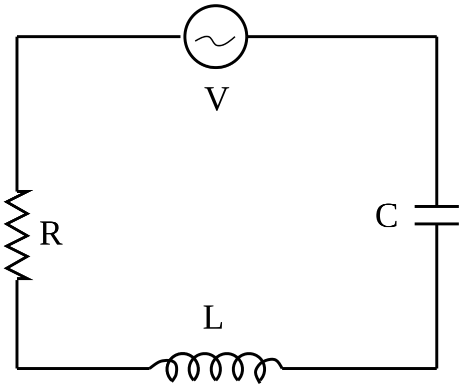


FIG. 7. Equivalent circuit for an RLC circuit including a thermal noise source $V_N(t)$.

$$\langle I^2\rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{f_V(\omega)}{R^2 + \left(L\omega - \frac{1}{C\omega}\right)^2}. \tag{2.83}$$

The integrand on the right hand side is small except near the resonant frequency $\omega_0 = 1/\sqrt{LC}$. Therefore, we approximate it as

$$\langle I^2\rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\omega^2 f_V(\omega)}{L^2(\omega^2-\omega_0^2)^2 + R^2\omega^2} \approx f_V(\omega_0) \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{\omega^2}{L^2(\omega^2-\omega_0^2)^2 + R^2\omega^2}. \tag{2.84}$$

The integrand has four poles $\omega = \pm i\frac{R}{2L} \pm \sqrt{\omega_0^2 - \left(\frac{R}{2L}\right)^2}$ and can be evaluated using the residue theorem, and yields

$$\langle I^2\rangle = \frac{f_V(\omega_0)}{2LR} \tag{2.85}$$

The energy stored in the coil is given by $\frac{1}{2}LI^2$, which according to the equipartition theorem is on the average equal to $\frac{1}{2}k_BT$, implying $\langle I^2\rangle = k_BT/L$. This gives

$$f_V(\omega) = 2k_BTR. \tag{2.86}$$

Note that the noise level is independent of the non-dissipative impedances L and C^{-1} which determine the resonant frequency ω_0. Consequently, it is the same at all frequencies and we have dropped the subscript 0 from the frequency argument. Note also that the integration over ω is carried out over both positive and negative frequencies, so the effective noise level is $4k_BTR$ per interval of positive frequencies. Equation (2.86) is known as the Nyquist theorem and gives the thermal noise level (Johnson noise) in any resistor.

As an example, let us consider a 1kΩ resistor at room temperature (300K): we have $f_V = 4 \times 1.38 \times 10^{-23}\mathrm{JK}^{-1} \times 300\mathrm{K} \times 1\mathrm{k}\Omega = 1.7 \times 10^{-17}\mathrm{V}^2\mathrm{s}$; if we measure the voltage with a device that has a frequency cutoff of 100MHz, the noise voltage is $\langle V_N^2\rangle = 100\mathrm{MHz} \times 1.7 \times 10^{-17}\mathrm{V}^2\mathrm{s} = 1.7 \times 10^{-9}\mathrm{V}^2$ or the rms noise is $\sqrt{\langle V_N^2\rangle} = 40\mu\mathrm{V}$. Note that the measured noise level increases with the cutoff frequency and our result actually diverges if we can measure arbitrarily large frequencies or arbitrarily fast fluctuations. The reason is that our classical approach fails at high frequencies when the energy of a single quantum is comparable to k_BT.

a. Connection to the Langevin Equation. We have now studied two approximate dynamic equations, the Boltzmann equation and the Langevin equation. They were quite different — Boltzmann equation is a deterministic[6] equation governing the time-evolution of the distribution function $f(\vec{r},\vec{p},t)$ whereas Langevin equation is a stochastic equation of motion for one particle. It is intuitively clear that if we have a system of particles that move according to the Langevin equation, we should be able to describe the time evolution of the system also in terms of the distribution function, and obtain something like the Boltzmann equation. We will now use the Langevin equation to derive the corresponding deterministic dynamic equation for the distribution function. The resulting transport equation is known as the Fokker-Planck equation.

We start from the Langevin equation which we now write as

$$\frac{dv}{dt} = -\eta v(t) + \frac{1}{m}\zeta(t). \tag{2.87}$$

For simplicity we assume that the system is spatially homogeneous so that r does not appear in the distribution function. We will therefore consider the probability density $P(v_1,t_1|v_0,t_0)$ that the particle has velocity v_1 at time t_1 under the condition that it is known to have velocity v_0 at time t_0. Mathematically this probability is given by the ensemble average (over different realizations of the stochastic force)

$$P(v_1,t_1|v_0,t_0) = \langle \delta(v_1 - v(t_1))\rangle_{v_0,t_0}. \tag{2.88}$$

The probability distribution at time $t_0+\Delta t$ is given by

$$P(v_1,t_0+\Delta t|v_0,t_0) = \langle \delta(v_1 - v(t_0+\Delta t))\rangle_{v_0,t_0}. \tag{2.89}$$

From the equation of motion we have

$$v(t_0+\Delta t) = v(t_0) - \eta\int_{t_0}^{t_0+\Delta t} dt\, v(t) + \frac{1}{m}\int_{t_0}^{t_0+\Delta t} dt\, \zeta(t) \tag{2.90}$$

and substituting this into the above expression we have

$$\begin{aligned}
&P(v_1,t_0+\Delta t|v_0,t_0)\\
&= \Big\langle \delta(v_1 - v(t_0)) - \left[-\eta\int_{t_0}^{t_0+\Delta t} dt\, v(t) + \frac{1}{m}\int_{t_0}^{t_0+\Delta t} dt\,\zeta(t)\right]\delta'(v_1 - v(t_0))\\
&+ \frac{1}{2}\left[\eta^2\left(\int_{t_0}^{t_0+\Delta t} dt\, v(t)\right)^2 + \frac{1}{m^2}\left(\int_{t_0}^{t_0+\Delta t} dt\,\zeta(t)\right)^2\right.\\
&\quad\left. -2\frac{\eta}{m}\left(\int_{t_0}^{t_0+\Delta t} dt\, v(t)\right)\left(\int_{t_0}^{t_0+\Delta t} dt\,\zeta(t)\right)\right]\delta''(v_1 - v(t_0))\Big\rangle_{v_0,t_0}
\end{aligned} \tag{2.91}$$

---

[6] By deterministic we mean that it does not explicitly contain the random noise term.

where we kept terms that appear at first sight to be second order in Δt. Now we calculate the averages over the different realizations of the stochastic force $\zeta(t)$. Only the terms that depend on $\zeta(t)$ are affected, and using

$$\begin{aligned} \langle \zeta(t) \rangle &= 0 \\ \langle \zeta(t)\zeta(t') \rangle &= A\delta(t-t') \end{aligned} \tag{2.92}$$

we obtain

$$\begin{aligned} P(v_1, t_0+\Delta t | v_0, t_0) &= \delta(v_1 - v_0) \\ &+ \left[\eta v(t_0)\Delta t + \frac{1}{2}\eta\, \dot{v}\, (t_0)(\Delta t)^2 \right] \delta'(v_1 - v_0) \\ &+ \left[\frac{1}{2}\eta^2 v^2(t_0)(\Delta t)^2 + \frac{1}{2m^2} A\Delta t \right] \delta''(v_1 - v_0) \end{aligned} \tag{2.93}$$

Note that the last term, $\int_{t_0}^{t_0+\Delta t} dt \int_{t_0}^{t_0+\Delta t} dt' \langle \zeta(t)\zeta(t') \rangle = \int_{t_0}^{t_0+\Delta t} dt \int_{t_0}^{t_0+\Delta t} dt' A\delta(t-t') = \int_{t_0}^{t_0+\Delta t} dt A = A\Delta t$ is only first order in Δt since the correlation function $\langle \zeta(t)\zeta(t') \rangle$ is proportional to a delta function. Now we drop all terms that are proportional to $(\Delta t)^2$, and use $P(v_1, t_0 | v_0, t_0) = \delta(v_1 - v_0)$ to obtain

$$P(v_1, t_0+\Delta t | v_0, t_0) = P(v_0, t_0 | v_0, t_0) + \Delta t \eta v(t_0) \frac{d}{dv} P(v, t_0 | v_0, t_0)_{v=v_0} + \Delta t \frac{1}{2m^2} A \frac{d^2}{dv^2} P(v, t_0 | v_0, t_0)_{v=v_0} \tag{2.94}$$

or

$$\frac{\partial f}{\partial t} = \eta p \frac{\partial f}{\partial p} + \frac{1}{2} A \frac{\partial^2 f}{\partial p^2}, \tag{2.95}$$

where we renamed the probability distribution function $f(p,t)$ and used $p = mv$ to bring the notation into agreement with our previous definitions. Finally, we use the fluctuation-dissipation theorem to relate A to η and get

$$\boxed{\frac{\partial f}{\partial t} = \eta p \frac{\partial f}{\partial p} + \eta m k_B T \frac{\partial^2 f}{\partial p^2}} \tag{2.96}$$

which is known as the Fokker-Planck equation. Since we derived the Fokker-Planck equation from the Langevin equation, it is subject to the same assumptions and limitations. In particular, we have ignored interactions (including collisions) between the particles of interest and only taken into account frequent, statistically independent random kicks by the environment.

The Fokker-Planck equation can be generalized to more complicated Hamiltonians when the damping term in the Langevin equation is written as $-\Gamma \frac{\partial H}{\partial p}$ (thus, $\Gamma = m\eta$ in the simple Langevin equation that we have studied), in which case the Fokker-Planck equation reads

$$\frac{\partial f}{\partial t} = \Gamma k_B T \frac{\partial}{\partial p} \left[\left(\frac{1}{k_B T} \frac{\partial H}{\partial p} + \frac{\partial}{\partial p} \right) f \right] \tag{2.97}$$

The right hand side of this equation vanishes if $f(p) \sim e^{-\beta H(p)}$, which is exactly the equilibrium distribution function — thus, equilibrium is a stationary solution of the Fokker-Planck equation as it should be. If the initial distribution $f(p, t_0)$ deviates from the equilibrium distribution $f_{\mathrm{eq}}(p)$, it will evolve according to the Fokker-Planck equation and approach the equilibrium distribution as $t \to \infty$.

b. Connection to Quantum Mechanics. Mathematically the Fokker-Planck equation is a diffusion equation in momentum space: the first time derivative of f is proportional to the second momentum derivative. Another diffusion equation which we are familiar with is the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi}{\partial r^2} + V(r)\psi(r,t). \tag{2.98}$$

Therefore, we can expect that many of the methods which are used to solve the Schrödinger equation can be applied to the Fokker-Planck equation, too. While this is the case there are a few complications: (i) the Schrödinger equation has an extra i on the left hand side, and (ii) in quantum mechanics expectation values are calculated as $\int dr\, \psi^*(r,t) O \psi(r,t)$, i.e. the distribution function is $|\psi(r,t)|^2$, whereas in non-equilibrium statistical mechanics the distribution function is $f(p,t)$. Thus, roughly speaking, the Schrödinger equation applies to the square root of the distribution function.

6. Master Equation

The last dynamic equation that we will discuss is the master equation. We assume that the system of interest can be in a number of discrete states which we shall denote by j. These states depend on the particular problem that we are studying: they could represent excited states of an atom (if we are considering lasers), or charge states of a small semiconductor island (in mesoscopics). We further assume that there is some probability Γ_{ij} that if the system is in state j at time t it will be in state i at time $t + \Delta t$. We assume that the transitions are instantaneous so that the state of the system is well defined at any given time. The system is described by the distribution function $P(j,t)$ which gives the probability that at time t the system is in state j. The distribution function changes in time because the system can leave the state j and move to another state j', or leave a state j' and enter the state j. Thus, we have the master equation

$$\boxed{\frac{dP(j,t)}{dt} = \sum_{j' \neq j} \left[\Gamma_{jj'} P(j',t) - \Gamma_{j'j} P(j,t)\right], \quad \forall j} \tag{2.99}$$

We require that the equilibrium state $P_{\mathrm{eq}}(j) \propto e^{-\beta\epsilon(j)}$ is a steady-state solution to the master equation, which implies

$$0 = \sum_{j'} \left[\Gamma_{jj'} e^{-\beta\epsilon(j')} - \Gamma_{j'j} e^{-\beta\epsilon(j)}\right], \quad \forall j \tag{2.100}$$

which sets a constraint on the acceptable transition rates $\Gamma_{jj'}$. The easiest way to satisfy this constraint is to assume *detailed balance*, i.e. to assume that each term in the sum vanishes independently. That implies

$$\Gamma_{jj'}e^{-\beta\epsilon(j')} = \Gamma_{j'j}e^{-\beta\epsilon(j)} \quad \forall j, j'. \tag{2.101}$$

The most straightforward applications of the master equation are usually systems where the number of relevant states is small. As an example, let us consider a three-state gas laser [6], the energy diagram of which is shown in Fig. 8. The numbers of gas atoms in states 1,2, and 3 are given by N_1, N_2 and N_3, respectively, and their time evolutions are controlled (before laser action starts) by the master equation

$$\begin{aligned}
\frac{dN_1}{dt} &= -(\Gamma + \gamma_{12} + \gamma_{13})N_1 + \gamma_{21}N_2 + \gamma_{31}N_3 \\
\frac{dN_2}{dt} &= \gamma_{12}N_1 - (\gamma_{21} + \gamma_{23})N_2 + \gamma_{32}N_3 \\
\frac{dN_3}{dt} &= (\Gamma + \gamma_{13})N_1 + \gamma_{23}N_2 - (\gamma_{31} + \gamma_{32})N_3,
\end{aligned} \tag{2.102}$$

where γ_{ij} denotes the probability rate per unit time for transitions from state i to state j (note that the convention regarding the order of subscripts is in laser physics opposite to what is common in other fields!), and Γ denotes pumping, i.e. the probability per unit time of exciting an atom in the ground state to the excited state 3. The total number of atoms is $N = N_1 + N_2 + N_3$ which allows us to eliminate one of the three equations. A further simplification is possible if we consider temperatures that are sufficiently low so that $\gamma_{12} \ll \gamma_{21}$, $\gamma_{13} \ll \gamma_{31}$, and $\gamma_{23} \ll \gamma_{32}$ which allows us to neglect the slow rates. The steady state solution is then given by

$$\begin{aligned}
N_1 &= \frac{\gamma_{21}(\gamma_{31} + \gamma_{32})}{\gamma_{21}(\gamma_{31} + \gamma_{32}) + (\gamma_{21} + \gamma_{32})\Gamma}N \\
N_2 &= \frac{\gamma_{32}\Gamma}{\gamma_{21}(\gamma_{31} + \gamma_{32}) + (\gamma_{21} + \gamma_{32})\Gamma}N.
\end{aligned} \tag{2.103}$$

If the pumping is sufficiently strong,

$$\Gamma > \gamma_{21}\left(1 + \frac{\gamma_{31}}{\gamma_{32}}\right), \tag{2.104}$$

the excited state is more populated than the ground state, and the system may begin to lase. Thus, we see that to achieve lasing at a moderate pumping level, it is favorable to have small γ_{21}, i.e., slow transition between the two states participating in lasing, and at the same time have $\gamma_{32} \gg \gamma_{31}$, i.e. fast decay from the uppermost level to the laser level.

Note that since we have neglected the stimulated transitions which depend on the amplitude of the electromagnetic field in the laser cavity, our discussion is only valid if there is no laser activity. Once the laser is operating, we must include an additional equation that describes the time evolution of the intensity of the radiation field W, as well as a coupling between the radiation field and the populations N_1 and N_2 (stimulated emission and absorption). The last term takes the form $B(N_2 - N_1)W$, where B is the so-called Einstein's B coefficient. This term renders the time evolution equations nonlinear.

We can generalize the master equation to systems with a continuum of states by regarding $P(x,t)$ as a probability density. In this case we can write the equation as

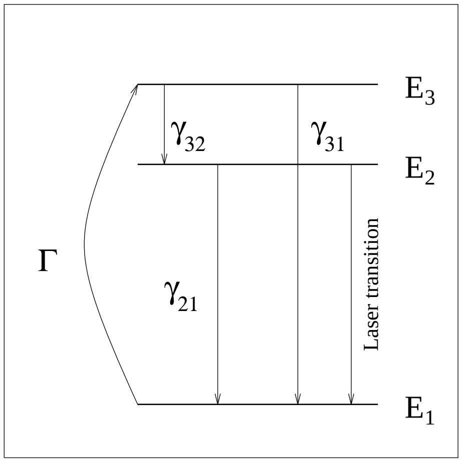


FIG. 8. Energy diagram of a three-state laser. The system is pumped from its ground state 1 to the excited state 3, subsequently relaxing back to the ground state through the intermediate state 2. Lasing takes place between states 2 and 1.

$$\frac{\partial P}{\partial t} = \int dx \, \left[\Gamma(x,x')P(x',t) - \Gamma(x',x)P(x,t)\right] . \tag{2.105}$$

Before we can use the master equation in any particular application we must determine the transition rates $\Gamma_{jj'}$. This is usually done by means of a quantum mechanical calculation, and can be rather a complicated undertaking. The master equation itself is not an exact quantum mechanical equation since it only keeps track of the probability density $P(j,t) = |\psi(j,t)|^2$ but not the phase of the wave function $\psi(j,t)$. It has nevertheless proven quite useful in describing a large variety of phenomena.

D. Linear Response

1. Classical Damped Oscillator

We have now discussed several transport equations (Liouville, Boltzmann, Langevin, Fokker-Planck, and master), and found that they are quite useful in describing systems slightly out of equilibrium. Usually the system is out of equilibrium because of some external force (e.g. the electric field) that results in a response (e.g. electric current) in the system. The transport equations can be used to calculate the response, and we have found that the response can often be expressed in terms of equilibrium quantities. We shall now approach more generally and derive a few general properties of the response functions.

The general theory of linear response is usually covered in courses on many-particle physics, and is discussed for instance in the books of Mahan [7] or Fetter and Walecka [8]. In this course we will content ourselves with a brief exposition of the relevant ideas and results.

Let us start by considering a classical one-dimensional damped harmonic oscillator which is subject to a time-dependent force $F(t)$, and is described by the equation of motion

$$m\frac{d^2x}{dt^2} + 2m\eta\frac{dx}{dt} + m\omega_0^2 x = F(t) \tag{2.106}$$

where m is the mass of the oscillating particle, ω_0 is the frequency of the undamped oscillator, and η is the damping. Note that for later convenience we have included an extra factor of 2 in the dissipation term as compared to the Langevin equation (2.60). The solution of (2.106) is given by

$$x(t) = x_0(t) + \int_{-\infty}^{t} dt' \left\{ e^{-\eta(t-t')}\frac{\sin[\omega_1(t-t')]}{m\omega_1} \right\} F(t') \tag{2.107}$$

where $\omega_1 = \sqrt{\omega_0^2 - \eta^2}$ is frequency of the damped oscillator ($\eta < \omega_0$) and $x_0(t)$ is a solution of the homogeneous equation. The reason for including the factor of 2 in the dissipation term is now apparent — we have the same exponential decay $e^{-\eta(t-t')}$ as previously.

Let us now assume that the force $F(t)$ consists of a stochastic part $\zeta(t)$ and of a deterministic part $F^{\text{ext}}(t)$ — the former describes thermal variation, and the latter describes an external perturbation like gravity or electric field. We are interested in describing the behavior of the system to linear order in the perturbation $F^{\text{ext}}(t)$. Following the analysis of equation (2.60) we consider an ensemble of systems described by (2.106), differing only in the realization of the stochastic force $\zeta(t)$. Since the equation of motion (2.106) is linear, the effects of the stochastic and deterministic forces are simply additive, and we find that the ensemble-averaged position is given by

$$\langle x(t)\rangle = \langle x(t)\rangle_0 + \int_{-\infty}^{\infty} dt'\, \chi(t-t')F^{\text{ext}}(t') \tag{2.108}$$

where $\langle x(t)\rangle_0$ is the average position in the absence of the perturbation, and $\chi(t)$ is given by

$$\chi(t) = \begin{cases} e^{-\eta t}\frac{\sin(\omega_1 t)}{m\omega_1}, & t > 0 \\ 0, & t < 0 \end{cases} \tag{2.109}$$

The function $\chi(t)$ is called the *retarded response function:* it is a response function since it describes how the system responds to the perturbation $F^{\text{ext}}(t)$, and it is retarded since the response to the perturbation $F^{\text{ext}}(t')$ occurs only at times t that are later than t', $t > t'$ (this is guaranteed by setting $\chi(t)$ to zero for $t < 0$). The response function satisfies the differential equation

$$m\frac{d^2\chi}{dt^2} + 2m\eta\frac{d\chi}{dt} + m\omega_0^2\chi = \delta(t) \tag{2.110}$$

with the boundary condition $\chi(t) = 0$ for $t < 0$ — thus, $\chi(t)$ is a response to a delta function impulse at time zero.

It is often more useful to consider the response in Fourier space. Taking the Fourier transform of $\chi(t)$ we have

$$\chi(\omega) = \int_{-\infty}^{\infty} dt\, e^{i\omega t}\chi(t) = \int_0^{\infty} dt\, \cos(\omega t)\chi(t) + i\int_0^{\infty} dt\, \sin(\omega t)\chi(t) = \chi'(\omega) + i\chi''(\omega) \tag{2.111}$$

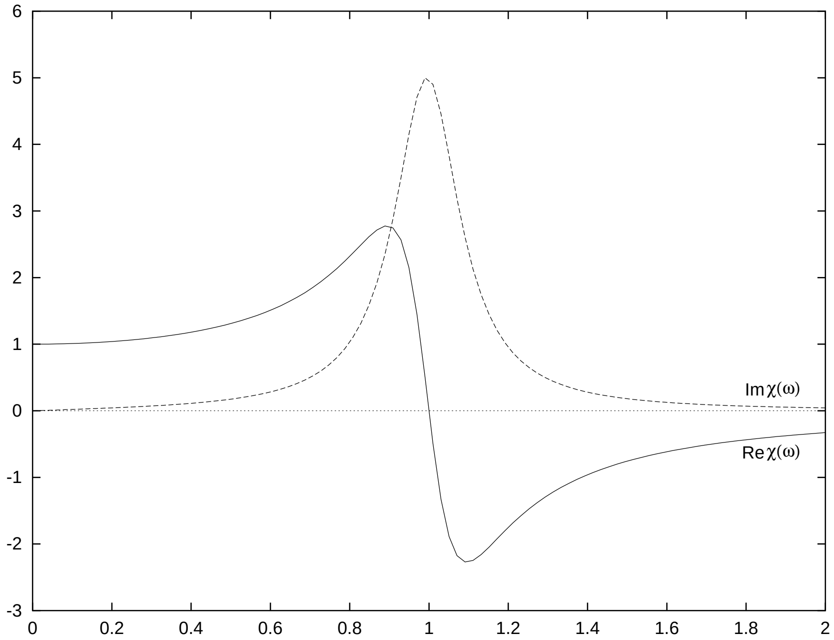


FIG. 9. The real and imaginary parts of the response function $\chi(\omega)$

where $\chi'(\omega)$ and $\chi''(\omega)$ are the real and imaginary parts of $\chi(\omega)$. Using the explicit form (2.109) we obtain

$$\begin{aligned} \chi'(\omega) &= \tfrac{1}{m}\tfrac{\omega_0^2-\omega^2}{(\omega_0^2-\omega^2)^2+4\eta^2\omega^2} \\ \chi''(\omega)) &= \tfrac{1}{m}\tfrac{2\eta\omega}{(\omega_0^2-\omega^2)^2+4\eta^2\omega^2} \end{aligned} \tag{2.112}$$

The functions $\chi'(\omega)$ and $\chi''(\omega)$ have been plotted in Fig. 9.

It is actually often more straightforward to first derive $\chi(\omega)$, and then obtain $\chi(t)$ by a Fourier transform. Using this approach, we start by Fourier transforming the equation of motion and solving for $x(\omega)$ which yields

$$x(\omega) = \frac{1}{m}\frac{F(\omega)}{\omega_0^2 - \omega^2 - 2i\eta\omega}. \tag{2.113}$$

Noticing that effects of ζ and F^{ext} decouple, we can immediately read off the response function $\chi(\omega)$. We can also determine the correlation function $C_x(\omega)$ in the absence of the external force, which yields

$$C_x(\omega) = 2k_BT\frac{1}{m}\frac{2\eta}{(\omega^2-\omega_0^2)^2+4\eta^2\omega^2} \tag{2.114}$$

where we used the connection $A = 4k_BTm\eta$ (note the extra factor of two due to the factor of two in the dissipation term!). Comparing this with the expression of $\chi''(\omega)$ allows us to write down the classical form of the fluctuation-dissipation theorem

$$\chi''(\omega) = \frac{\omega}{2k_BT} C_x(\omega). \tag{2.115}$$

If we had considered a case with $\langle x \rangle \neq 0$, we should use the correlation function $\langle (x(t) - \langle x \rangle)(x(0) - \langle x \rangle) \rangle$ in the place of $C_x(t)$.

One final step to justify the name fluctuation-dissipation theorem is to show that $\chi''(\omega)$ measures the dissipated power. We do that by considering the energy absorbed by the particle from the external field, which is given by $dW = F^{\text{ext}} dx$. Consequently, the absorbed power is $P(t) = \langle \frac{dW}{dt} \rangle = F^{\text{ext}}(t) \partial_t \langle x(t) \rangle$. Expressing $\langle x(t) \rangle$ in terms of the response function $\chi(t)$ and Fourier transforming we find, for $F^{\text{ext}}(t) = F_0 \cos(\omega_{\text{ext}} t)$, that the average power absorption during one cycle of $F^{\text{ext}}(t)$ is given by $\overline{P(t)} = \frac{1}{2\pi} F_0^2 \omega_{\text{ext}} \chi''(\omega_{\text{ext}})$. Thus, the rate at which the system absorbs energy from and external source is indeed given by $\chi''(\omega)$.

2. Linear Response in Quantum Mechanics

The quantum mechanical treatment of linear response at finite temperature is somewhat subtle, and we do not have the time to address all the interesting questions.[7] We start by assuming that some time in the distant past, which we call $t = -\infty$, the system is in thermal equilibrium and described by the Hamiltonian $\hat{H}$. A time-dependent perturbation $\hat{H}_1(t)$ is slowly turned on, and we wish to determine how the observables are affected by the perturbation. The expectation value of an observable is given by

$$\langle \hat{A}(t) \rangle = \text{Tr}(\hat{\rho}(t) \hat{A}) \tag{2.116}$$

where $\hat{\rho}(t)$ is the density matrix describing the system. In equilibrium the density matrix is given by $\hat{\rho}_{\text{eq}} = \frac{1}{Z} e^{-\beta \hat{H}}$, but when the system is out of equilibrium it can be more complicated. The density matrix, like any other quantum mechanical observable, obeys the Heisenberg equation of motion[8]

$$\frac{d}{dt} \hat{\rho}(t) = \frac{i}{\hbar} [\hat{\rho}(t), \hat{H}_0 + \hat{H}_1(t)]. \tag{2.117}$$

Let the density matrix in the absence of the perturbation be $\hat{\rho}_0$ and write $\hat{\rho}(t) = \hat{\rho}_0 + \hat{\rho}_1(t)$. If the perturbation is weak, $\hat{\rho}_1(t)$ is small, and to first order in small quantities the Heisenberg equation of motion gives

$$i\hbar \frac{d}{dt} \hat{\rho}_1(t) - [\hat{H}_0, \hat{\rho}_1(t)] = [\hat{H}_1(t), \hat{\rho}_0]. \tag{2.118}$$

The left hand side can be rewritten as

$$e^{-i\hat{H}_0 t/\hbar} \left[i\hbar \frac{d}{dt} \left(e^{i\hat{H}_0 t/\hbar} \hat{\rho}_1(t) e^{-i\hat{H}_0 t/\hbar} \right) \right] e^{i\hat{H}_0 t/\hbar}$$

---

[7] This section follows very closely the treatment of Mahan [7].

[8] Here all operators are in the Heisenberg picture, see e.g. Sakurai: Modern Quantum Mechanics.

which gives a first order differential equation that can be solved to yield

$$\hat{\rho}_1(t) = -\frac{i}{\hbar} e^{-i\hat{H}_0 t/\hbar} \left\{ \int_{-\infty}^{t} dt' \, [\hat{\tilde{H}}_1(t'), \hat{\rho}_0] \right\} e^{i\hat{H}_0 t/\hbar} \tag{2.119}$$

where $\hat{\tilde{H}}_1(t) = e^{i\hat{H}_0 t/\hbar} \hat{H}_1(t) e^{-i\hat{H}_0 t/\hbar}$ is the perturbation in the interaction picture.

Now we know how the density matrix changes due to the perturbation $\hat{H}_1(t)$, and we can therefore determine how the expectation values of observables change. We obtain

$$\langle \hat{A}(t) \rangle = -\frac{i}{\hbar} \mathrm{Tr} \left\{ e^{-i\hat{H}_0 t/\hbar} \left[\int_{-\infty}^{t} dt' \, [\hat{\tilde{H}}_1(t'), \hat{\rho}_0] \right] e^{i\hat{H}_0 t/\hbar} \hat{A} \right\}, \tag{2.120}$$

where we assumed that $\langle \hat{A}(t) \rangle$ vanishes in the absence of the perturbation. Using the cyclic invariance of trace we can move the first exponential factor to the right of the operator $\hat{A}$, which gives the combination $e^{i\hat{H}_0 t/\hbar} \hat{A} e^{-i\hat{H}_0 t/\hbar} = \hat{A}(t)$, the operator in the interaction picture. Writing out the commutator and using the cyclic invariance once more gives

$$\langle \hat{A}(t) \rangle = -\frac{i}{\hbar} \int_{-\infty}^{t} dt' \, \mathrm{Tr} \left\{ \hat{\rho}_0 \left[\hat{A}(t), \hat{\tilde{H}}_1(t') \right] \right\} = -\frac{i}{\hbar} \int_{-\infty}^{t} dt' \, \langle \left[\hat{A}(t), \hat{\tilde{H}}_1(t') \right] \rangle_{\mathrm{eq}}. \tag{2.121}$$

The perturbation is usually of the form $\hat{\tilde{H}}_1(t) = \int dr \hat{B}(r,t) F^{\mathrm{ext}}(r,t)$ where $F^{\mathrm{ext}}(r,t)$ is an external force field (e.g. electric field) and $\hat{B}(r,t)$ is the operator to which $F^{\mathrm{ext}}(r,t)$ couples (in the case of the electric field $\hat{B}(r,t)$ would be charge density). In this case the expectation value of the operator $\hat{A}(t)$ is given to first order in $F^{\mathrm{ext}}(r,t)$ by

$$\langle \hat{A}(r,t) \rangle = \int_{-\infty}^{\infty} dt' \int dr' \, \chi_{AB}(r,t;r',t') F^{\mathrm{ext}}(r',t') \tag{2.122}$$

where the generalized response function (generalized susceptibility) is given by

$$\boxed{\chi_{AB}(r,t;r',t') = -\frac{i}{\hbar} \langle [\hat{A}(r,t), \hat{B}(r',t')] \rangle_{\mathrm{eq}} \theta(t-t')} \tag{2.123}$$

where we have allowed also the operator $\hat{A}$ to be spatially dependent. This is a fundamental result, known as the Kubo formula, that connects the response function to an expectation value of some operator evaluated in equilibrium. Many systems of interest are both translationally invariant and energy conserving, in which case $\chi_{AB}(r,t;r',t')$ only depends on the relative variables $r - r'$ and $t - t'$. In that case it is convenient to Fourier transform to get

$$\langle \hat{A}(k,\omega) \rangle = \chi_{AB}(k,\omega) F^{\mathrm{ext}}((k,\omega) \tag{2.124}$$

The causality condition — a response can only occur after the perturbation — implies that the generalized susceptibility satisfies the Kramers-Kronig relation

$$\boxed{\mathrm{Re}\chi_{AB}(\omega) = P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\mathrm{Im}\chi_{AB}(\omega')}{\omega' - \omega}}. \tag{2.125}$$

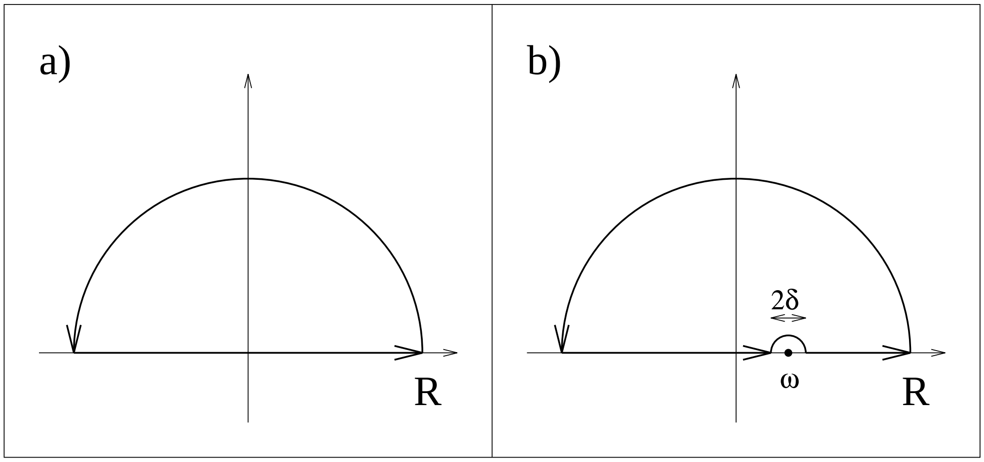


FIG. 10. Integration contours C_1 and C_2 used in proving the Kramers-Kronig relation.

Here P denotes the principal part of the integral, i.e., $P \int \frac{dx}{x-y} = \lim_{\delta \to 0^+} \left[\int^{y-\delta} \frac{dx}{x-y} + \int_{y+\delta} \frac{dx}{x-y} \right]$.

We will prove the Kramers-Kronig relation in two stages. First, let us write $\chi_{AB}(t)$ in terms of its Fourier transform as

$$\chi_{AB}(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \chi_{AB}(\omega) \tag{2.126}$$

and write the integral over the real axis as an integral over a closed contour. If $t < 0$, we can use the contour of Figure 10 since the over the arc vanishes as the radius of the arc tends to infinity (note that on the arc $\mathrm{Im}(z) > 0$, and therefore $\mathrm{Re}(-itz) < 0$ for $t < 0$). On the other hand, causality implies $\chi_{AB}(t < 0) = 0$, and we have

$$0 = \oint_{C_1} \frac{dz}{2\pi} e^{-izt} \chi_{AB}(z) = i \sum_i \mathrm{Res} \left(e^{-izt} \chi_{AB}(z) \right)_{z=z_i} \tag{2.127}$$

where I used residue theorem to evaluate the contour integral and denoted the poles of $\chi_{AB}(z)$ inside the contour by z_i. Since the right hand side must vanish for all $t < 0$, we conclude that $\chi_{AB}(z)$ cannot have poles inside the contour C_1, i.e. $\chi_{AB}(z)$ is an analytic function in the upper half plane.[9] This is true for all retarded response functions.

The second step is to consider the integral

$$I(\omega) = \oint_{C_2} d\omega' \frac{\chi_{AB}(\omega')}{\omega' - \omega}. \tag{2.128}$$

---

[9] This can be proven more carefully by examining the function $\chi(z) = \int_0^\infty dt\, e^{izt} \chi(t)$: for $\mathrm{Im}(z) > 0$ is necessarily convergent since the physical response function $\chi(t)$ is finite for all t.

Since the integrand has no poles inside C_2, the integral vanishes identically. If we assume that $\chi_{AB}(\omega)$ tends to zero as $|\omega| \to \infty$, the integral along the large arc vanishes, and we have

$$0 = \int_{-\infty}^{\omega-\delta} d\omega' \, \frac{\chi_{AB}(\omega')}{\omega' - \omega} + \int_{\omega+\delta}^{\infty} d\omega' \, \frac{\chi_{AB}(\omega')}{\omega' - \omega} + i \int_{\pi}^{0} d\theta \, \chi_{AB}(\omega + \delta e^{i\theta}). \tag{2.129}$$

Taking the limit $\delta \to 0^+$ gives

$$P \int_{-\infty}^{\infty} d\omega' \, \frac{\chi_{AB}(\omega')}{\omega' - \omega} = i\pi \chi_{AB}(\omega). \tag{2.130}$$

Separating both sides to real and imaginary parts proves the Kramers-Kronig relations. The relations are valid for all physical response functions which vanish as $\omega \to \infty$. Sometimes we need to subtract a constant from the usual response function to satisfy the condition — for instance, the Kramers-Kronig relations apply to $(\frac{1}{\epsilon(q,\omega)} - 1)$ rather than $\frac{1}{\epsilon(q,\omega)}$ since $\lim_{\omega\to\infty} \epsilon(q,\omega) = 1$.

Another important result, the fluctuation-dissipation theorem, relates the real and imaginary parts of the equilibrium correlation function $\langle \hat{A}(t)\hat{B}(0)\rangle$,

$$\boxed{\int_{-\infty}^{\infty} dt \, e^{i\omega t} \mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle = -i \tanh\left[\frac{\hbar\omega}{2k_B T}\right] \int_{-\infty}^{\infty} dt \, e^{i\omega t} \mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle} \tag{2.131}$$

We can prove the fluctuation-dissipation theorem by writing

$$\begin{aligned} & \langle \hat{A}(t)\hat{B}(0)\rangle \\ = & \ \tfrac{1}{Z}\mathrm{Tr}\left\{ e^{-\beta\hat{H}} e^{i\hat{H}t/\hbar} \hat{A} e^{-i\hat{H}t/\hbar} \hat{B} \right\} \\ = & \ \tfrac{1}{Z}\mathrm{Tr}\left\{ \hat{B} e^{i\hat{H}(t+i\beta\hbar)/\hbar} \hat{A} e^{-i\hat{H}(t+i\beta\hbar)/\hbar} e^{-\beta\hat{H}} \right\} \\ = & \ \tfrac{1}{Z}\mathrm{Tr}\left\{ e^{-\beta\hat{H}} \hat{B}\hat{A}(t+i\beta\hbar) \right\} \\ = & \ \langle \hat{B}(0)\hat{A}(t+i\beta\hbar)\rangle \\ = & \ \exp(i\beta\hbar\tfrac{\partial}{\partial t})\langle \hat{B}(0)\hat{A}(t)\rangle. \end{aligned} \tag{2.132}$$

This is the first intermediate result. The second intermediate result we get by using $\langle \hat{A}(t)\hat{B}(0)\rangle^* = \langle \hat{B}(0)^\dagger \hat{A}(t)^\dagger\rangle = \langle \hat{B}(0)\hat{A}(t)\rangle$ which implies

$$\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle = \frac{1}{2i}\langle(\hat{A}(t)\hat{B}(0) - \hat{B}(0)\hat{A}(t))\rangle \tag{2.133}$$

$$\mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle = \frac{1}{2}\langle(\hat{A}(t)\hat{B}(0) + \hat{B}(0)\hat{A}(t))\rangle \tag{2.134}$$

Using the first intermediate result we now obtain

$$\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle = \frac{1}{2i}\left[\exp(i\beta\hbar\frac{\partial}{\partial t}) - 1\right]\langle \hat{B}(0)\hat{A}(t)\rangle \tag{2.135}$$

$$\mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle = \frac{1}{2}\left[\exp(i\beta\hbar\frac{\partial}{\partial t}) + 1\right]\langle \hat{B}(0)\hat{A}(t)\rangle \tag{2.136}$$

Fourier transforming results in the replacement $i\frac{\partial}{\partial t} \to \omega$, and gives the Fourier transforms of both $\mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle$ and $\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle$ in terms of the Fourier transform of $\langle \hat{B}(0)\hat{A}(t)\rangle$, which yields equation (2.131).

Comparing the expressions of $\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle$ and $\chi_{AB}(t)$ shows that they are connected through

$$\chi_{AB}(t) = \frac{2}{\hbar}\theta(t)\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle \tag{2.137}$$

Setting $\hat{B} = \hat{A}$ and $t = 0$ shows that $\mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle$ is related to the fluctuations in the system; hence, equation (2.131) connects the fluctuations to the response function, whence the name fluctuation-dissipation theorem.

It is interesting to note that if we regard $\hbar$ as a small quantity in the equations (2.136) and expand $\exp(i\beta\hbar\frac{\partial}{\partial t}) \approx 1 + i\beta\hbar\frac{\partial}{\partial t}$, we obtain $\mathrm{Im}\langle \hat{A}(t)\hat{B}(0)\rangle \approx \frac{1}{2}\beta\hbar\mathrm{Re}\langle \hat{A}(t)\hat{B}(0)\rangle$, which gives $\chi_{AB}(t) \approx \beta\mathrm{Re}\langle \frac{\partial}{\partial t}\hat{A}(t)\hat{B}(0)\rangle\theta(t)$. Thus, in the classical limit $\hbar \to 0$ we find that the response function is given by the time derivative of the correlation function.

III. EQUILIBRIUM CRITICAL PHENOMENA

A. Introduction

Most phenomena in nature take place in a smooth, continuous fashion. Science has been very successful in describing smooth processes in terms of differential equations the solutions of which are usually continuous, analytic functions. Deviations from this kind of smooth behavior pose new challenges and require new methods. In the remaining part of this course we will familiarize ourselves with the topic of critical phenomena and their mathematical description.

Although smooth variation seems to be the rule in nature, exceptions to this rule are easy to find. The most familiar critical phenomena are solid-liquid transition (melting) and liquid-gas transition (vaporization or boiling). In both cases the system under consideration undergoes a violent change to a new state with different properties. Other examples of critical phenomena can be observed daily on city streets where the flowing state of traffic frequently undergoes a transition to a static state or a traffic jam. In addition, the ideas of critical phenomena have been applied to the financial markets where the stock market crashes can be thought of as phase transitions.

In most cases the behavior of a system is controlled by a few external control parameters, which typically include applied pressure, temperature, external magnetic field, or the density of vehicles on the streets. In the generic case the control parameters determine the state of the system quite unambiguously — water at a pressure of 100kPa and temperature of 300K is a liquid — but for some values of the control parameters the state of the system is not uniquely determined. An example of this is water at 100kPa and 373K in which case the system can be either in a high-density phase (liquid) or a low-density phase (vapor). Another example is a ferromagnet below Curie temperature at zero external magnetic field: all spins are aligned, but the direction of their alignment is arbitrary. The region in the parameter space where the ambiguity occurs is called the coexistence line (or coexistence manifold if there are more than two control parameters). Crossing the coexistence line corresponds to a phase transition. The coexistence line ends at the critical point which for water is ($p =$ 220MPa, T = 647K) and for a ferromagnet ($T{=}T_c$,$h{=}0$).

Consider what happens if we have a container of argon vapor at 100K and 100kPa and we start slowly decreasing the volume of the container (see Fig. 11): The pressure will increase, until (p, T) reaches the coexistence line. After that the pressure will not increase but instead more and more of the vapor turns into liquid, and the system is in a mixed state. Once all the vapor has turned into liquid, the pressure continues to increase (and quite rapidly so since the compressibility of the liquid is very small). Since the density as a function of pressure changes discontinuously as the system crosses the coexistence line, the phase transition is called discontinuous or first order. If, instead, we change the control parameters so that the system passes through the critical point, there is no discontinuity in the density as a function of pressure, and the transition is called continuous or second order. Study of critical phenomena can be defined as the study of continuous phase transitions, or physics in the vicinity of the critical point. In the following we will only occasionally make remarks on discontinuous phase transitions. The reason for this emphasis on continuous phase transitions is not that they are intrinsically more important or more interesting than

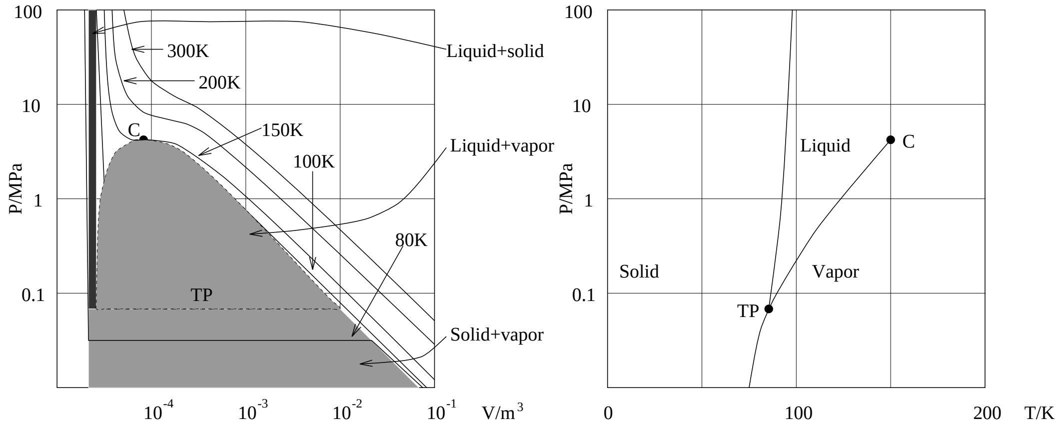


FIG. 11. Schematic phase diagram of argon: (a) volume-pressure diagram, (b) temperature-pressure diagram. Taken from [9].

discontinuous phase transitions. Instead, we focus on them because they are usually easier to describe, and have a number of common properties, which allows us to make some general statements about continuous phase transitions. To the best of my knowledge, there is currently no corresponding general theory of discontinuous phase transitions.

Critical phenomena that take place in different systems share a number of common features — the existence of an order parameter, power law divergences of many observables, diverging correlations lengths *etc.* — and we will focus on these common features rather than study any one system in great detail. It turns out that the properties of systems at criticality are to a great extent determined by a few parameters, and many systems that at first seem very different are closely related. In the following we will mostly study critical phenomena in magnetic systems, where the control parameters are temperature and the external magnetic field. The results are, however, directly applicable to many other systems such as the liquid-gas transition. The analogy between a magnetic phase transition and the liquid-gas transition is discussed in Appendix A.

B. Critical Exponents

One characteristic of critical phenomena is that many observables diverge near the critical point. Some quantities of particular interest are the specific heat[10] $C_h(T) = \frac{\partial E}{\partial T}$, magnetization $m(h,T) = -\left(\frac{\partial F}{\partial h}\right)_T$, and susceptibility $\chi(h,T) = \frac{\partial m}{\partial h}$. The critical exponents have conventional names, which are given below:

---

[10] Here E is the total (internal) energy, F is the free energy ($F = E - ST$), and h is the external magnetic field.

$$\boxed{\begin{array}{lcl} C(h=0,T) & \sim & C_\pm|t|^{-\alpha} \\ m(h=0,T) & \sim & m_\pm|t|^{\beta} \\ m(h,T=T_c) & \sim & \mathrm{sign}(h)|h|^{1/\delta} \\ \chi(h=0,T) & \sim & \chi_\pm|t|^{-\gamma}. \end{array}} \tag{3.1}$$

Here t is the reduced temperature $t = \frac{T-T_c}{T_c}$. The exponents characterizing the divergence of different quantities are experimentally observed to be the same on both sides of the transition, although proportionality constants may differ (i.e. we may have $C_+ \neq C_-$). Experimentally it is also observed that very different systems often have same critical exponents, and also the ratios of the proportionality constants above and below the transition temperature (e.g. C_+/C_-) are found to be same in many systems. Therefore, we can postulate that the critical exponents and ratios of the proportionality constants are independent of details like impurities, shape, size etc., of the system, and are determined by some fundamental underlying quantities. This generality makes critical exponents particularly useful in characterizing critical phenomena. One of the focal points of the following lectures is the calculation of critical exponents in a number of (usually approximate) ways, and in the process develop some understanding about the physics of phase transitions and critical phenomena.

C. Mean Field Theory

1. Order Parameter

Many phase transitions are connected with the appearance or disappearance of some kind of order. In the low-temperature solid phase atoms occupy well-determined positions in a regular lattice, and the structure is obviously quite well ordered. When the temperature increases, the material becomes liquid, and much of the order is lost. Similarly in magnetic systems: at low temperatures all spins point in particular directions (in a ferromagnet they all point in the same direction, in an antiferromagnet they point alternatingly in opposite directions), but once the temperature exceeds the Curie or Néel temperature, order is lost and spins point in random directions. This order is characterized by an *order parameter*, which in the words of the two-times physics Nobel prize winner John Bardeen is "the important thing in all important problems". A more conventional definition is that the order parameter is a quantity that has a different value in the two phases we are interested in, and it is conventionally chosen to have the value zero in the disordered high-temperature phase. In the case of a ferromagnet the order parameter is magnetization: at low temperatures the system magnetizes spontaneously, whereas at high temperatures the magnetization is zero. In the liquid-gas transition the order parameter is $\rho - \rho_c$ where ρ_c is the density at the critical point. There are often several alternative choices for the order parameter — for instance, in the liquid-solid transition we could choose an arbitrary function of $\rho - \rho_c$ or even $\rho - \rho_g$ (ρ_g is the density of the gaseous phase) — and some physical insight is required to make the most convenient choice.

In magnetic systems the order parameter is in the general case a three-dimensional vector with both a magnitude and a direction, whereas in the liquid-gas transition the order parameter is a scalar. In the transition between a normal metal and a superconductor the order parameter is the electron pair amplitude, which is a complex number and can therefore be presented as a two-dimensional vector. In more complicated phase transitions like the transition between the low- and high-temperature phases of a spin glass the order parameter may be even more complicated.

2. Symmetries and Spin Models

Symmetry is of fundamental importance in the analysis of critical phenomena. There are two kinds of symmetries present in a generic statistical physics problem: the symmetries of the Hamiltonian describing the system, and the symmetries of the equilibrium state. In the case of a ferromagnet, for instance, the Hamiltonian can often be written as $H_H = \frac{1}{2}\sum_{\vec{r},\vec{r}'} J(\vec{r},\vec{r}')\vec{S}(\vec{r}) \cdot \vec{S}(\vec{r}')$, indicating that spins at positions $\vec{r}$ and $\vec{r}'$ interact through a coupling constant $J(\vec{r},\vec{r}')$. Since in this Hamiltonian the interaction term is proportional to the scalar product $\vec{S}(\vec{r}) \cdot \vec{S}(\vec{r}')$, it only depends on the relative orientation of the two spins. Consequently, the Hamiltonian is invariant under global spin rotations — the energy is unchanged if all spins are rotated by the same amount about the same axis. The symmetry group of the Hamiltonian is therefore SU(2), the familiar group of spin rotations. If, however, we consider the effect of an external magnetic field $\vec{h}(\vec{r})$, we must add an extra term of the form $\sum_{\vec{r}} \vec{h}(\vec{r}) \cdot \vec{S}(\vec{r})$ to the Hamiltonian, and the symmetry of the Hamiltonian is reduced. Hamiltonians of this type (with or without the external field) are called Heisenberg models, and were first studied in 1928 by Werner Heisenberg. In most cases the Hamiltonian is studied on a lattice (i.e. $\vec{r}$, $\vec{r}'$ are sites on a lattice), and the coupling $J(\vec{r},\vec{r}')$ is taken to be non-zero only if the lattice sites $\vec{r}$ and $\vec{r}'$ are nearest neighbors. In that case we can write $H_H = \sum_{\langle i,j\rangle} J\vec{S}_i \cdot \vec{S}_j$ where $\sum_{\langle i,j\rangle}$ indicates a sum over nearest neighbor pairs. Note that each nearest neighbor pair occurs only once in the sum whereas in the unrestricted sum $\sum_{\vec{r},\vec{r}'}$ all pairs $(\vec{r},\vec{r}')$ occur twice — in the latter case we introduced the factor $\frac{1}{2}$ in the Hamiltonian to remove the double counting.

In this course we consider only classical Hamiltonians and ignore all quantum effects. For instance we treat spins as classical vectors, and also determine symmetry properties from a classical point of view. While our model Hamiltonian has symmetry group SU(2) if we treat the spins as quantum mechanical operators, in the classical approximation, when spins are treated as classical vectors, the symmetry group is the ordinary three-dimensional rotation group O(3). The two groups are closely related although not isomorphic, see e.g. J. J. Sakurai, *Modern Quantum Mechanics*, chapter 3. The classical approximation is usually well justified in the context of critical phenomena when we are interested in the long range properties.[11]

---

[11] The reason is that near the critical point the behavior of the system is controlled by a few modes whose energies vanish as $T \to T_c$. If $T_c > 0$, these modes behave in an essentially classical way (quantum statistical physics reduces to classical statistical physics at high temperatures), but if $T_c = 0$, quantum mechanical effects cannot be ignored. Hence, critical phenomena with $T_c = 0$ are

Occasionally in real materials the spin-orbit coupling implies that spins are aligned in particular lattice directions. If the spin-orbit coupling forces the spin vectors to lie in a fixed plane, we can treat the spin as a two-dimensional vector, and obtain the Hamiltonian for an XY model, $H_{XY} = \frac{1}{2}\sum_{\vec{r},\vec{r}'} J(\vec{r},\vec{r}')[S_x(\vec{r})S_x(\vec{r}') + S_y(\vec{r})S_y(\vec{r}')]$. This Hamiltonian is invariant under two-dimensional spin rotations and its symmetry group is therefore O(2). In some materials like $YFeO_3$ the spins are forced to lie in a specific lattice direction, and it is therefore natural to describe them with one-dimensional vectors or scalars. The model Hamiltonian is in that case the Ising Hamiltonian $H_I = \frac{1}{2}\sum_{\vec{r},\vec{r}'} J(\vec{r},\vec{r}')S(\vec{r})S(\vec{r}')$, and the symmetry group is the "one-dimensional rotation group" O(1), which is usually called Z(2). This simplest spin model was introduced by Wilhelm Lenz in 1925 and studied by Lenz's student Ernst Ising.

The reason why spin models occupy such a central position in the study of critical phenomena is that the spin language can be applied naturally to problems with different symmetry properties. In addition to the O(M) models (Ising, XY, Heisenberg) we can construct a large number of other spin models with different symmetry properties. The most common ones are p-state Potts model ($H_P = \sum_{\langle i,j\rangle} J\delta_{s_i,s_j}$ where $s_i \in \{1,2,\ldots,p\}$, permutation symmetry) and the p-state clock model ($H_c = \sum_{\langle i,j\rangle} J\cos(s_i - s_j)$ where $s_i = n\frac{2\pi}{p}$, $n \in \{1,2,\ldots,p\}$, discrete rotation symmetry) which interpolate between the Ising and XY models. With such a plethora of models it is not surprising that most critical phenomena can be mapped onto a spin model — all it takes is a careful analysis of the symmetry properties of the system we are interested in. It turns out, for instance, that the liquid-gas phase transition can be described by an Ising model with an external field, which is evident if we consider that for the liquid-gas transition the order parameter is the density difference $\rho - \rho_c$, which is a scalar, and the order parameter of an Ising model is also a scalar. In this course we will focus on the simplest model, the Ising model, and discuss the other models only when the difference between the discrete symmetry of an Ising model and the continuous symmetry of an XY or Heisenberg model is important.

Unlike the symmetry of the Hamiltonian, the symmetry of the equilibrium state depends on the temperature. At high temperatures — above the Curie temperature — entropic contributions to the free energy dominate, and the equilibrium state is isotropic. Each spin is equally likely to point in any direction. In equilibrium at low temperatures all spins point to the same direction, but the direction is arbitrary if there is no external field. Our intuition and experience tells that system has non-zero magnetization; however, if there is no external magnetic field, the magnetization can point in any direction, and therefore if we evaluate the ensemble-averaged magnetization $\langle\vec{m}\rangle$ it vanishes! In each individual sample, however, the symmetry is spontaneously broken and the magnetization is non-zero. It only averages to zero if we consider a large number of independent samples (an ensemble) since the direction of $\langle\vec{m}\rangle$ varies from sample to sample.

Mathematically, a complication arises from the fact that the limits of taking the system size to infinity and the external magnetic field to zero do not commute: the state of an infinite system depends on how the system was prepared. If at some time in the distant past the

---

known as quantum phase transitions.

system was subjected to a uniform external magnetic field, however weak, all spins aligned themselves in the direction of the field, and they remain aligned in that direction even after the external field has been removed. Consequently, the equilibrium state may not reflect the current Hamiltonian but may be determined by the system's memory. In finite systems the length of the memory is finite since it only costs a finite amount of energy to rotate spins to a new direction, and if we wait long enough, thermal fluctuations smear out long-term memory. In infinite systems, in contrast, the energy cost to rotate spins independently so that they all point in a new common direction is infinite, and thermal fluctuations do not remove memory effects. Hence, if our Hamiltonian $H = \sum_{\vec{r},\vec{r}'} J(\vec{r},\vec{r}')\vec{S}(\vec{r}) \cdot \vec{S}(\vec{r}')$ describes a *finite* system, the average spin at any site vanishes, $\langle \vec{S}(\vec{r}) \cdot \vec{u} \rangle = 0$ where $\vec{u}$ is a unit vector in an arbitrary direction. Since the magnetization is proportional to the average of the spin, we have the result that the average magnetization of a finite system, subject to no external magnetic fields in the finite past, is zero at all non-zero temperatures. Therefore, a purist would have to admit that there are no phase transitions in finite systems — any evidence to the contrary that you may have accumulated is only due to the fact that you have made observations only over a finite range of time! If the system is infinite and $T < T_c$ there exists, however, a direction $\vec{u}$ such that $\langle \vec{S}(\vec{r}) \cdot \vec{u} \rangle \neq 0$. Mathematically this corresponds to the physical choice of the order of the limits of a vanishing external field and infinite system size, $\lim_{\vec{h}\to 0} \lim_{N\to\infty}$. The average $\langle \ldots \rangle$ is now a thermal average performed over an ensemble of systems with identical histories, that is, the limit $\vec{h} \to 0$ is taken the same way for all members of the ensemble.[12] This meticulous definition of what we mean by the innocent-looking symbols $\langle \ldots \rangle$ is necessary to bring our mathematical description into agreement with our physical intuition.

3. Ising Model

Now we are in the position to make our first attempt at a quantitative description of a critical phenomenon. The simplest approximation is known as a mean field theory, and is based on the assumption of spatial homogeneity. We shall apply this assumption to the Ising model, and obtain simple expressions for the ordering temperature, critical exponents and the overall behavior of the model. Our estimates are only approximate, and sometimes even qualitatively wrong, but due to its simplicity the mean field is often a good starting point. In the following sections we will discuss how to improve our estimates.

The nearest neighbor spin-half Ising Hamiltonian in the presence of an external magnetic field reads

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_j s_j \tag{3.2}$$

where we have chosen the signs so that $J > 0$ corresponds to a ferromagnetic model (energy is

---

[12] The restriction of the ensemble is necessary to guarantee ergodicity, i.e. to force time-averages and ensemble-averages to be equal.

low if adjacent spins point in the same direction).[13] Here h is a quantity that is proportional to the external magnetic field but has the units of energy. Since we are dealing with a spin-$\frac{1}{2}$ Ising Hamiltonian, the spins s_j at each lattice site can assume only two values, $s_j = \pm 1$. Due to its simplicity the Ising model has been applied to a great number of problems in various fields — apart from magnetism, the Ising model has been used in models for neural networks, protein folding and social imitation, just to mention a few. Since 1969 more than 12000 scientific publications have made use of the Ising model, and the literature keeps growing at a rate of 700–900 publications per year [12].

We can re-write the Hamiltonian as a sum of two terms the first of which depends on a particular spin (which we call s_0), while the second term is independent of s_0,

$$H = -s_0 \left[J \sum_{\langle 0,j \rangle} s_j + h \right] + \widetilde{H}(\{s_j\}_{j \neq 0}). \tag{3.3}$$

The quantity in the square brackets can be viewed as an effective magnetic field $h_0'(\{s_j\}_{j \in \langle 0,j \rangle}) = \left[J \sum_{\langle 0,j \rangle} s_j + h \right]$ acting on the spin s_0. The effective magnetic field is a combination of the externally applied field h and interactions with neighboring spins. Note that h_0' is an observable which is subject to similar fluctuations as any other observable in statistical physics; in other words, h_0' has some average value $\langle h_0' \rangle$ in the ensemble, and in each member of the ensemble the actual value of h_0' may differ from the average value. To emphasize this we write the Hamiltonian as

$$H = -s_0 \langle h_0' \rangle - s_0 \left[h_0'(\{s_j\}_{j \in \langle 0,j \rangle}) - \langle h_0' \rangle \right] + \widetilde{H}(\{s_j\}_{j \neq 0}) \tag{3.4}$$

The average value of the spin s_0 is calculated as

$$\langle s_0 \rangle = \frac{\mathrm{Tr}\left(s_0 e^{-\beta H(\{s_j\})} \right)}{\mathrm{Tr}\left(e^{-\beta H(\{s_j\})} \right)} \tag{3.5}$$

where the trace indicates that we have to sum over all possible spin configurations $\{s_i\}$. To simplify this expression we make the drastic approximation that the effective field h_0' is equal to its thermal average $\langle h_0' \rangle$. Hence, we ignore thermal fluctuations between members of the ensemble, or, equivalently, we ignore the temporal fluctuations of the spins.[14] For simplicity we denote this thermal average by h', $h' = \langle h_0' \rangle$. This innocent-looking approximation uncouples the spin s_0 from the other spins, and renders the part of the Hamiltonian that depends on s_0 non-interacting. After these approximations we have arrived at the Hamiltonian

---

[13] We use the convention that a lower case spin variable s_j is an integer, usually ±1, whereas an upper case spin variable S_j denotes a physical spin with units of Js and either a half-integer or integer magnitude.

[14] Note that due to assumed ergodicity ensemble average is equal to time average.

$$H = -h's_0 + \widetilde{H}(\{s_j\}_{j\neq 0}) \tag{3.6}$$

so that the average value of s_0 is now easily calculated to be

$$\langle s_0 \rangle = \frac{e^{\beta h'} - e^{-\beta h'}}{e^{\beta h'} + e^{-\beta h'}} = \tanh[\beta h']. \tag{3.7}$$

Translational invariance implies that $\langle s_i \rangle = \langle s_0 \rangle = m$ for all sites i — the average spin is position-independent. The effective magnetic field is consequently given by $h' = \left[J \sum_{\langle 0,j \rangle} \langle s_j \rangle + h \right] = qJm + h$, where q is the number of nearest neighbors and m is the magnetization. Therefore, we have the self-consistency condition

$$\boxed{m = \tanh[\beta(qJm + h)].} \tag{3.8}$$

Let us first look at the solutions of this equation at a zero external magnetic field, in which case we have $m = \tanh[\beta qJm]$. Both sides of the equation are plotted in Fig. 12. This equation always has the solution $m = 0$. Furthermore, if m is a solution of the equation then so is $-m$, and we can restrict our analysis to $m \geq 0$. For very large m the hyperbolic tangent goes to 1, and the left hand side is much larger than the right hand side. For small m we can approximate the $\tanh[\beta qJm]$ by its Taylor expansion near $m = 0$, and we find that at small m the right hand side of our equation is approximately βqJm. Thus, if $\beta qJ > 1$ then for small m the right hand side is larger than the left hand side, and since for large m the right hand side is smaller than the left hand side, we conclude that for $\beta qJ > 1$ the equation has several (=3) solutions. By calculating the free energy of the two solutions $m = 0$ and $m = m^* > 0$ we can conclude that the latter solution represents the equilibrium state.[15] Thus, we have found that the Ising model develops a spontaneous magnetization if $\beta qJ > 1$, which gives us the mean field estimate for the critical temperature $\boxed{T_c = qJ/k_B}$.

Note that in the mean field approximation the critical temperature only depends on the number of nearest neighbors q. Therefore, we have the estimate that in one dimension the critical temperature is $T_{c,\text{mf}}^{(1d)} = 2J/k_B$ $(q = 2)$, in two dimensions on a square lattice $T_{c,\text{mf}}^{(2d,\text{sq})} = 4J/k_B$, and in three dimensions on a simple cubic lattice $T_{c,\text{mf}}^{(3d,\text{sc})} = 6J/k_B$. The best estimates from more accurate theories are $T_c^{(1d)} = 0$, $T_c^{(2d,\text{sq})} = 2.27J/k_B$ and $T_c^{(3d,\text{sc})} = 4.51J/k_B$ — thus, the mean field overestimates the critical temperature but appears to be more accurate in higher dimensions. Intuitively it is quite clear that a mean field theory has to overestimate the tendency towards an ordered state, and hence the temperature T_c above which entropy dominates — after all, our approximation amounted to disregarding fluctuations and hence reducing entropy.

We can derive mean field estimates for the critical exponents α, β, γ, and δ. Let us first consider α. Since α is the critical exponent of the specific heat, which is defined

---

[15] Left as an exercise.

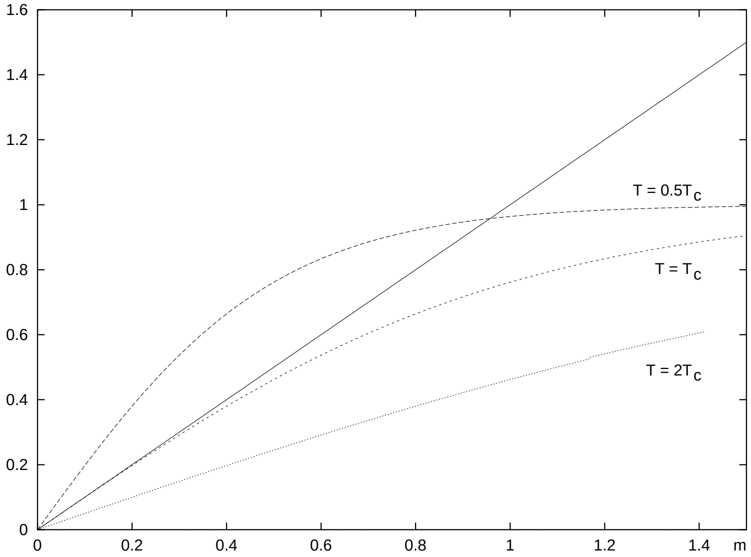


FIG. 12. Self-consistency condition for the mean field theory. The solid line is the left hand side of Eq. (3.8), and the dashed lines are the right hand sides of Eq. (3.8) at different temperatures. A non-trivial solution ($m \neq 0$) is found only for $T < T_c$.

as the derivative of energy with respect to temperature, we must start by calculating the expectation value of energy

$$E = \langle H \rangle = -J \sum_{\langle i,j \rangle} \langle s_i s_j \rangle = -\frac{1}{2} NqJm^2 \tag{3.9}$$

(the factor $\frac{1}{2}$ compensates for double counting). The specific heat is now obtained through

$$C_h(T) = \frac{\partial E}{\partial T} = -\frac{1}{2} NqJ \frac{\partial m^2(T)}{\partial T} \tag{3.10}$$

Thus, to get the specific heat behavior near T_c we need to know the behavior of the magnetization $m(T)$ near the critical point.

By examining the behavior of the self-consistency condition $m = \tanh[\beta qJm]$ for $\beta \approx \beta_c = (qJ)^{-1}$ we conclude that the solution $m^*(\beta)$ is small near the critical point, and consequently we may approximate both sides of the self-consistency condition by their Taylor series and obtain

$$m = \beta qJm - \frac{1}{3}(\beta qJ)^3 m^3. \tag{3.11}$$

The roots of this third order equation are $m = 0$ and

$$m = \pm m^* = \pm\sqrt{\frac{3(\beta qJ - 1)}{(\beta qJ)^3}} = \pm\frac{1}{\beta qJ}\sqrt{3}\sqrt{\frac{\beta qJ - 1}{\beta qJ}} \tag{3.12}$$

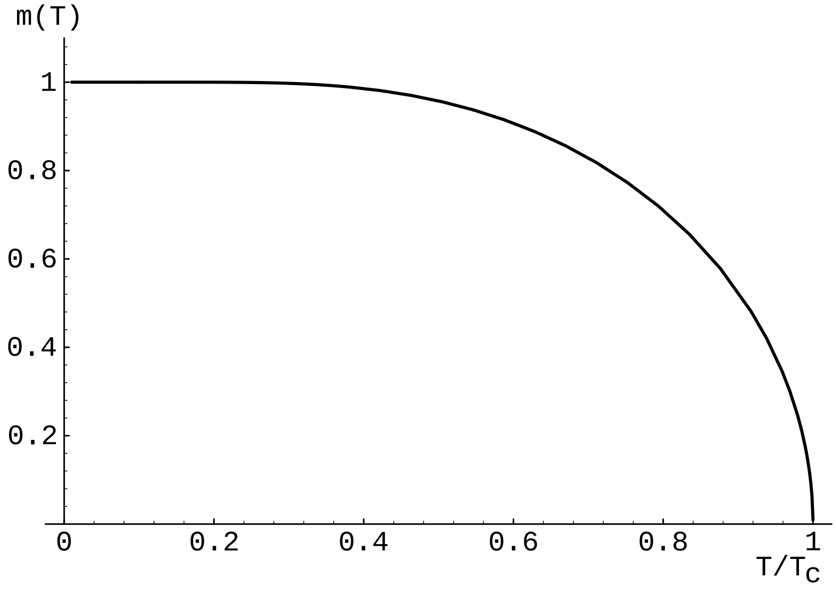


FIG. 13. Magnetization of the Ising model in the mean field approximation as a function of the temperature. Note the square root singularity near T_c.

Above the critical temperature only the root $m = 0$ is real, but for $T \lesssim T_c$ or $\beta \gtrsim \beta_c$ we have another real root given by

$$m^* = \sqrt{3}\frac{T}{T_c}\sqrt{1 - \frac{T}{T_c}}. \tag{3.13}$$

In terms of the reduced temperature $t = \frac{T-T_c}{T_c}$ we have

$$m(T) \approx \begin{cases} \sqrt{3}|t|^{1/2}, & t \lesssim 0 \\ 0, & t > 0. \end{cases} \tag{3.14}$$

The mean field magnetization is plotted in Fig. 13 as a function of the temperature.

We have thus obtained the critical exponent $\beta = \frac{1}{2}$ in the mean field approximation.[16] Substituting this into the expression for the specific heat we obtain

$$c_H(T) = \frac{C_h(T)}{N} \approx \begin{cases} \frac{3}{2}k_B|t|^0, & t \lesssim 0 \\ 0, & t > 0. \end{cases} \tag{3.15}$$

In the mean field approximation the specific heat does not diverge at the critical temperature, it merely has a discontinuity, and therefore the specific heat exponent is $\alpha = 0$.

The remaining two exponents γ and δ describe the system's behavior at small magnetic fields, and we therefore need to consider the self-consistency condition in the presence of an external field h. We leave it as an exercise to conclude that

---

[16] Note that traditionally the magnetization exponent and the inverse temperature $\frac{1}{k_B T}$ are both called β. The context should make it always clear which β we mean.

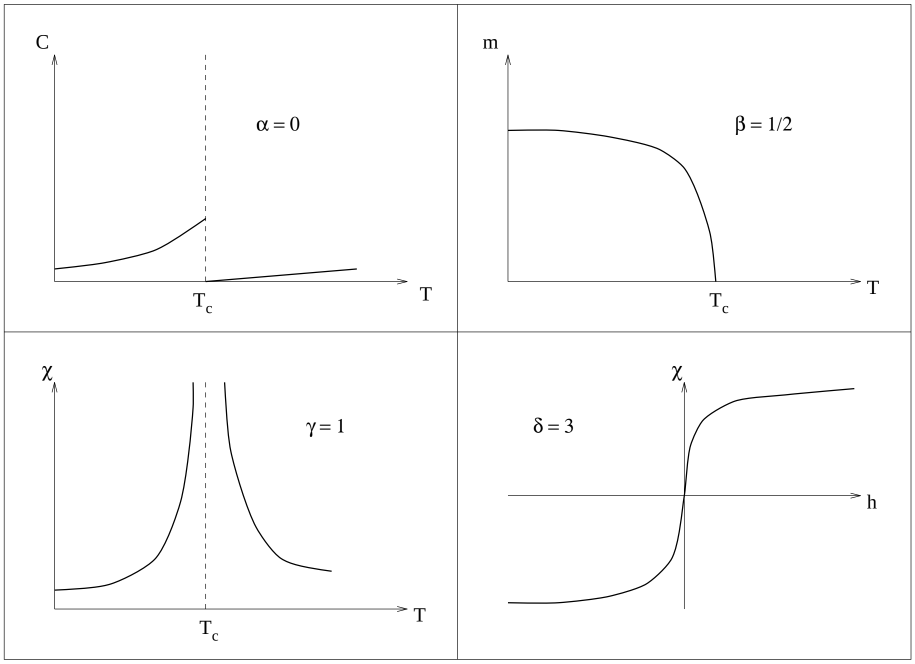


FIG. 14. Mean field predictions for the dependence of the specific heat, magnetization, and susceptibility on the external parameters.

$$\chi(T,h=0) \approx \begin{cases} \frac{1}{2k_BT_c}|t|^{-1}, t \lesssim 0 \\ \frac{1}{k_BT_c}|t|^{-1}, t \gtrsim 0 \end{cases} \tag{3.16}$$

so that the susceptibility exponent is $\gamma = 1$. Finally, as another exercise, you can show that $m(T = T_c, h) \sim \mathrm{sign}(h)|h|^{1/3}$ and therefore in the mean field approximation $\delta = 3$. As a summary the mean field behavior of the specific heat, magnetization, and susceptibility are shown in Fig. 14.

While the results obtained from mean field theories are often even qualitatively wrong, the approach is still very widely used. It often predicts trends correctly (e.g. higher critical temperature if the number of nearest neighbors is large), and it can serve as a reference theory which is useful in defining new concepts. Mean field theories are remarkably simple since they use an approximation that reduces an original interacting model to a non-interacting one, usually combined with a self-consistency condition. A mean field theory can also serve as a starting point for systematic improvements which is the topic of the next chapter.

4. Bethe Approximation

In the previous section we reduced the original problem containing many degrees of freedom to a mean field problem with only one degree of freedom. The reduction was based on an approximation the validity of which is not clear although we argued that a mean field theory can be expected to be more accurate if each spin is coupled to a large number of other spins. This approach can be generalized to a cluster method, where we reduce the original problem with many degrees of freedom to a problem with a few degrees of freedom. We expect this method to be more accurate than the simple mean field theory, but since we now have to solve a more complicated problem, some of the simplicity of the previous section is lost. In connection with spin problems the cluster approximation is known as Bethe approximation after Hans Bethe.

Our starting point is again the Ising Hamiltonian

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_j s_j \tag{3.17}$$

where we now isolate the terms containing a particular spin s_0 and its nearest neighbors s_1, $s_2, \ldots, s_q$. The interactions of the remaining spins with s_j, $j \in \{1, 2, \ldots, q\}$ are described by an effective magnetic field h'. This gives us the Bethe approximation for the part of the Hamiltonian that depends on s_0,

$$H_{\text{Bethe}} = -J s_0 \sum_{j=1}^{q} s_j - h s_0 - h' \sum_{j=1}^{q} s_j . \tag{3.18}$$

Note that the effective magnetic field h' is in part due to the external field h and in part due to the interactions between the spins s_j and other spins not explicitly included in the Hamiltonian. This Hamiltonian must be accompanied with a consistency condition that determines the effective field h'. To derive such a condition let us first calculate the average value of the central spin s_0. It is given by

$$\begin{aligned} \langle s_0 \rangle &= \frac{1}{Z} \sum_{s_0=\pm 1} \sum_{s_1=\pm 1} \cdots \sum_{s_q=\pm 1} s_0 e^{-\beta H} \\ &= \frac{1}{Z} \sum_{s_0=\pm 1} s_0 \left[e^{K s_0 + \beta h'} + e^{-K s_0 - \beta h'} \right]^q e^{\beta h s_0} \\ &= \frac{2^q}{Z} \left[e^{\beta h} \cosh^q(K + \beta h') - e^{-\beta h} \cosh^q(-K + \beta h') \right] , \end{aligned} \tag{3.19}$$

where $K = \beta J$. Similarly, we can calculate the average value of one of the spins s_j, $j \neq 0$, which gives

$$\begin{aligned} \langle s_{j \neq 0} \rangle = \tfrac{2^q}{Z} &\left[e^{\beta h} \cosh^{q-1}(K + \beta h') \sinh(K + \beta h') \right. \\ &\left. + e^{-\beta h} \cosh^{q-1}(-K + \beta h') \sinh(-K + \beta h') \right] . \end{aligned} \tag{3.20}$$

If the coupling J is ferromagnetic, $J > 0$, the equilibrium state of the original Hamiltonian is translationally invariant, and we must have $\langle s_0 \rangle = \langle s_{j \neq 0} \rangle$, which gives the consistency condition

$$\begin{aligned} & e^{\beta h} \cosh^q(K + \beta h') - e^{-\beta h} \cosh^q(-K + \beta h') \\ = {} & e^{\beta h} \cosh^{q-1}(K + \beta h') \sinh(K + \beta h') + e^{-\beta h} \cosh^{q-1}(-K + \beta h') \sinh(-K + \beta h'). \end{aligned} \tag{3.21}$$

This expression is in general quite complicated but it simplifies significantly if we consider that case of no external field, $h = 0$. Then we have

$$
\begin{aligned}
& \cosh^q(K+\beta h') - \cosh^q(-K+\beta h') \\
& = \cosh^{q-1}(K+\beta h')\sinh(K+\beta h') + \cosh^{q-1}(-K+\beta h')\sinh(-K+\beta h') \\
\Leftrightarrow \quad & \cosh^{q-1}(K+\beta h')\,[\cosh(K+\beta h') - \sinh(K+\beta h')] \\
& = \cosh^{q-1}(K-\beta h')\,[\cosh(K-\beta h') - \sinh(K-\beta h')] \\
\Leftrightarrow \quad & \frac{\cosh^{q-1}(K+\beta h')}{\cosh^{q-1}(K-\beta h')} = e^{2\beta h'} \\
\Leftrightarrow \quad & \frac{\cosh(K+\beta h')}{\cosh(K-\beta h')} = e^{\frac{2}{q-1}\beta h'}.
\end{aligned}
\tag{3.22}
$$

We can simplify this expression by noticing that the effective field h' is due to the interaction of a spin s_j with $(q-1)$ of its nearest neighbors (the q^{th} nearest neighbor, s_0, has been treated exactly), and therefore $h' = J(q-1)m$ where $m = \langle s_j \rangle$ is the magnetization. Substituting this into the above expression yields

$$\frac{\cosh[K(1+(q-1)m)]}{\cosh[K(1-(q-1)m)]} = e^{2Km}. \tag{3.23}$$

By examination we notice that this equation always has the solution $m = 0$. Also, if m is a solution, so is $-m$, and we can restrict our analysis to $m \geq 0$. If m is very large, the left hand side is approximately one, and the right hand side is very large. Thus, if K is such that for a small positive m the left hand side is larger than the right hand side, the equation must have a non-trivial solution m^*. To determine the condition for K we Taylor expand this equation for small m and obtain the condition

$$1 + m2K(q-1)\tanh(K) > 1 + m2K \tag{3.24}$$

which gives us the critical value K_c such that there exists a non-trivial m^* if $K > K_c$,

$$\boxed{\tanh K_c = \tfrac{1}{q-1}.} \tag{3.25}$$

For a one-dimensional Ising model we have $q = 2$ and the critical point is given by $\tanh K_c = 1$, which implies $K_c = \infty$ or $T_c = 0$, which is the exact solution for this case. For a two-dimensional Ising model on a square lattice $q = 4$, and $K_c \approx 0.347$ or $T_c \approx 2.89J/k_B$, which is rather close to the exact value of (approximately) $2.27J/k_B$ and a significant improvement over the simple mean field value of $4J/k_B$. For a three-dimensional Ising model on a simple cubic lattice we have $q = 6$ and $K_c \approx 0.203$ or $T_c \approx 4.93J/k_B$, which compares quite favorably with the estimate $4.51J/k_B$ from more sophisticated theories. Thus, the Bethe approximation yields much more accurate critical temperatures than the simple mean field theory. However, the Bethe approximation gives same critical exponents as the simple mean field theory as you will discover in the home problems.

5. Landau-Ginzburg Theory

a. Landau Theory. So far the mean field theory we have used is rather *ad hoc*, and must be constructed specifically for every model that we wish to consider. It would be desirable to develop a unified approach that can be generalized to different types of models without much difficulty. Such an approach was introduced by Lev Landau, and is known as the Landau theory of phase transitions.

The basic quantities of interest in Landau theory are the free energy and the order parameter. The theory is based on the assumption that the free energy, or some other suitable thermodynamic potential, can be written as a series expansion with the order parameter as the expansion parameter. In magnetic problems it is most convenient to consider the Gibbs free energy $\min_{\vec{m}} G(\vec{h}, \vec{m}, T)$ where the external variables are the magnetization $\vec{h}$ and the temperature T. Note that the magnetization takes a value in such a way that $G(\vec{h}, \vec{m}, T)$ is minimized, and therefore $\vec{m}$ is not a free parameter. For now we set $\vec{h} = 0$ and for compactness we also suppress the explicit minimization with respect to $\vec{m}$: it is understood that the physical value of the free energy corresponds to a magnetization $\vec{m}$ such that $G(\vec{m}, T)$ is a minimum. We begin by writing down the most general power series expansion for $G(\vec{m}, T)$ which is consistent with the symmetries of the problem. If there is no external field, the Gibbs free energy is independent of the direction of the magnetization, so for a model where $\vec{m}$ is a vector in a three-dimensional space we have the expansion

$$G(\vec{m}, T) = a(T) + \frac{1}{2}b(T)m^2 + \frac{1}{4}c(T)m^4 + \frac{1}{6}d(T)m^6 + \dots \tag{3.26}$$

Here $a(T)$, $b(T)$, $c(T)$, and $d(T)$ are temperature-dependent coefficients. Note that due to spherical symmetry no odd terms (in m) can occur: a term proportional to e.g. $\vec{m}$ would require the introduction of an extra vector and could only appear as a combination $\vec{h} \cdot \vec{m}$, which would break the assumed spherical symmetry. An external magnetic field does break the symmetry, and therefore odd terms can appear if the external magnetic field is nonzero. The Landau theory is by construction a mean field theory since we have not included any effects of a spatially varying magnetization.

It is by no means guaranteed that the Landau expansion is meaningful. First of all, by writing $G(\vec{m}, T)$ as a power series in $\vec{m}$ we tacitly assume that $G(\vec{m}, T)$ is an analytic function of $\vec{m}$ (this excludes non-analytic terms like $|\vec{m}|$). This is not the case at the critical point as we will see in the subsequent sections. Secondly, even if $G(\vec{m}, T)$ is an analytic function, the expansion is only useful if it converges sufficiently fast. Without knowing anything about the coefficients $a(T)$, $b(T)$, $c(T)$ etc. that is a bold assumption, and the best we can do is to hope that the expansion converges at least for sufficiently small $|\vec{m}|$. Thus, the Landau theory may not be accurate in the vicinity of the critical point, or if the magnetization is not small. However, it turns out to be a good starting point.

Let us now assume that all coefficients in equation (3.26) vanish except for $a(T)$, $b(T)$ and $c(T)$. We know that the equilibrium state corresponds to the minimum of free energy, so let us find the value of m that minimizes $G(\vec{m}, T)$. If $c(T) < 0$ the minimum occurs for $m \to \infty$ which is not in the range of convergence of the expansion, so for consistency we must require $c(T) > 0$ (if in a particular system $c(T) < 0$, we must include more terms so that the coefficient of the highest power is positive). In that case there are two possible extrema: $m = 0$ and $m = m^* = \sqrt{-\frac{b(T)}{c(T)}}$. The latter one is real only if $b(T) < 0$. The values of $G(\vec{m}, T)$ at these points are, respectively, $G(0, T) = a(T)$ and $G(m^*, T) = a(T) - \frac{1}{4}\frac{b^2(T)}{c(T)}$. Thus, for $b(T) > 0$ the equilibrium magnetization is $\vec{m} = 0$ but for $b(T) < 0$ the equilibrium

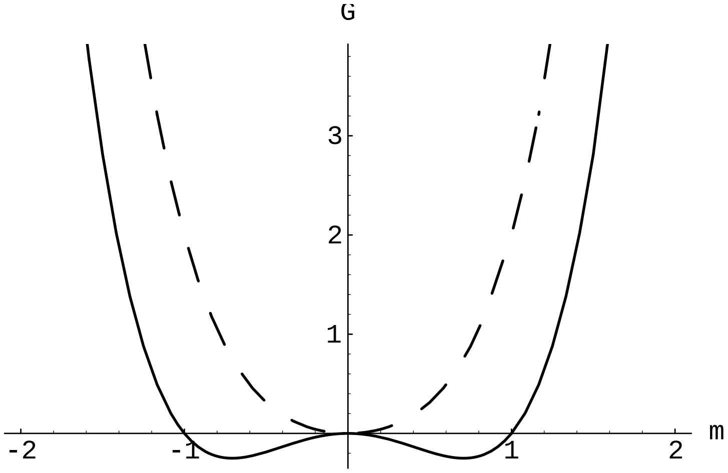


FIG. 15. Free energy as a function of $\vec{m}$ above and below the critical temperature. For purposes of illustration I have taken $\vec{m}$ to be a scalar. The solid curve is for $T < T_c$ and the dashed one for $T > T_c$. For $T < T_c$ the free energy is minimized for $m = \pm m^* \neq 0$, and the system is spontaneously magnetized.

magnetization is $|\vec{m}| = \sqrt{-\frac{b(T)}{c(T)}}$. In order for this result to agree with what we have found previously we must require $b(T)$ to be an increasing function of temperature so that the magnetization vanishes at high temperatures. We conclude that, for $c(T) > 0$, a phase transition between an ordered and disordered phase corresponds to a zero of the coefficient $b(T)$, i.e. $b(T_c) = 0$. For $T < T_c$ the non-magnetized state $m = 0$ corresponds to a local maximum of the free energy, as depicted in Fig. 15.

We still do not know anything about the temperature dependence of the coefficients $a(T)$, $b(T)$ and $c(T)$ except that $b(T_c) = 0$ and $c(T_c) > 0$. Let us assume for simplicity that $a(T)$ and $c(T)$ are independent of temperature. The magnetization near the critical point is then $m = 0$ for $T > T_c$ and $m = \sqrt{-b(T)/c}$ for $T < T_c$. From the previous sections on mean field theory we know that $m \sim |T - T_c|^{1/2}$ for $T \lesssim T_c$; this behavior is reproduced in the Landau theory if we assume

$$b(T) = b_0(T - T_c) \tag{3.27}$$

where b_0 is a constant. With $b(T)$ given by this form, the magnetization $m(T)$ is small near the critical point, and we can therefore expect the expansion in terms of m to be a good one near T_c.

Let us now analyze the Landau theory with $a(T) = a$, $b(T) = b_0(T - T_c)$ and $c(T) = c$. We already know that this gives $\beta = \frac{1}{2}$ for the magnetization exponent. We can calculate the heat capacity exponent using $c_h = T\left(\frac{\partial S}{\partial T}\right)_h$ and $S = -\left(\frac{\partial G}{\partial T}\right)_h$, which gives

$$c_h(T) = \begin{cases} \frac{b_0^2}{2c}T, & T < T_c \\ 0, & T > T_c \end{cases} \tag{3.28}$$

so that at the critical point the specific heat does not diverge but merely has a discontinuity. This is in agreement with what we found earlier.

Let us now take a step back and consider what happens if $c(T)$ is not positive. In that case we must include the term $\frac{1}{6}d(T)m^6$ and assume that $d(T) > 0$. Minimizing $G(m,T)$ with respect to m gives three local extrema, $m = 0$ and $m^2 = \frac{1}{2d}(-c \pm \sqrt{c^2 - 4bd})$. Evaluating the free energy at a temperature T_1 such that $b(T_1) = 0$ shows that the minimal free energy is obtained at $m^2 = -\frac{c}{d}$. Thus, the equilibrium magnetization is non-zero already at a temperature where $b(T)$ vanishes provided that $c(T)$ is negative at that temperature. Let us now examine the appearance of equilibrium magnetization. We keep our previous assumption that $b(T)$ is an increasing function of temperature and consider temperatures such that $b(T) > 0$. Then $m = 0$ is a local minimum as is $m^2 = (m^*)^2 = \frac{1}{2d}(-c + \sqrt{c^2 - 4bd})$, and the free energies at these magnetizations are given by $G(0,T) = a$ and

$$G(m^*, T) = a + \frac{c^3 - 6bcd - (c^2 - 4bd)\sqrt{c^2 - 4bd}}{24d^2}. \tag{3.29}$$

At the critical temperature these two free energies must be equal, which implies $b(T_c) = \frac{3c^2(T_c)}{16d(T_c)} > 0$; if $b(T) > b(T_c)$ the free energy minimum occurs at $m = 0$, and if $b(T) < b(T_c)$ the free energy minimum is at $m = m^*$. Since we assume that $b(T)$ is an increasing function of temperature, this is equivalent to saying that the magnetization vanishes above T_c and is non-zero below T_c. Substituting the critical temperature value of $b(T)$ into the expression for m^* we find that the magnetization obeys

$$m(T) \to \begin{cases} \sqrt{-\frac{3c(T_c)}{4d(T_c)}}, & T \to T_c^- \\ 0, & T \to T_c^+ \end{cases} \tag{3.30}$$

($T_c^{\pm}$ indicate temperatures that are infinitesimally smaller (respectively larger) than T_c). The magnetization as a function of temperature is plotted in Fig. 16 for the simple case when $b(T) = b_0(T - T_0)$ and the other coefficients a, c, and d are temperature independent.

Thus, we have found that if the coefficient $c(T)$ is negative at the critical temperature, the magnetization jumps suddenly from zero to a non-zero value as the temperature is lowered below T_c. This kind of discontinuous phase transition is commonly known as a first order phase transition. A phase transition is second order if the order parameter is continuous across the critical temperature; this is the case in Landau theory if $c(T_c) > 0$. Since in a first order transition the order parameter is not necessarily small near the transition point, any expansion in terms of the order parameter is subject to criticism, which is the basic reason why we mostly concentrate on continuous transitions.

In first order transitions the order is lost suddenly as the temperature exceeds the critical temperature. This corresponds to an instantaneous increase of entropy, and a corresponding latent heat $\Delta Q = T_c \Delta S$. Thus, if a system that undergoes a first order phase transition

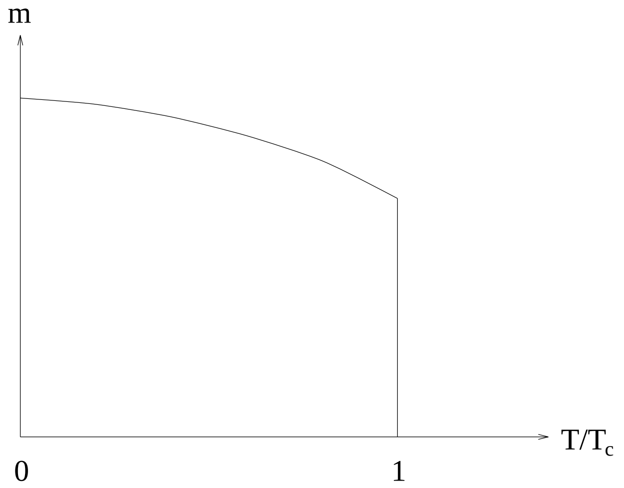


FIG. 16. The magnetization as a function of temperature in the Landau model with $c < 0$, corresponding to a discontinuous (or first order) phase transition.

is heated from a temperature $T < T_c$ to the critical temperature, the phase transition is associated with an absorption of heat. A common example of a first order transition is boiling as which we discussed in the introduction.

The strength of the Landau theory is its simplicity. It is readily generalizable to systems with different symmetries, and it gives reasonable results. The theory must be complemented by some more detailed model-specific arguments regarding the coefficients $a(T)$, $b(T)$ etc.. The limitations of Landau's theory are mostly due to the simple form of the expansion. We shall try to improve it in the next section.

b. Fluctuations. The Landau theory is a type of a mean field theory since it excludes all effects due to the spatial variation of the order parameter. This limitation can be relaxed by considering the Landau-Ginzburg (or Ginzburg-Landau) theory which is widely used especially in the study of superconductivity. Instead of regarding the free energy as a function of the order parameter m, we now consider it as a *functional* of a spatially varying order parameter $\vec{m}(\vec{r})$.

A function is a mapping from a number to a number, and similarly a functional is a mapping from a function to a number. The functional that you are most familiar with is in all likelihood a definite integral: $I_1[f] = \int_a^b dx f(x)$ can be thought of as a mapping from the function $f(x)$ onto a number (i.e. the value of the integral). A more complicated functional is for instance $I_2[f] = \int_a^b dx[f'(x)]^2$. Note that by convention the argument of a functional is enclosed in brackets $[\ldots]$ whereas the argument of a function is enclosed in parentheses $(\ldots)$.

The starting point for the Landau-Ginzburg theory is to write the free energy as a functional of the order parameter. Now it is more convenient to work with another thermodynamic potential, A, where the free variables are magnetization and temperature. This Helmholtz free energy is obtained from the Gibbs free energy by the Legendre transformation $A(\vec{m}, T) = G(\vec{h}_{\mathrm{ext}}, T) + \vec{h}_{\mathrm{ext}} \cdot \vec{m}$. We consider only the case when the external magnetic field vanishes, $\vec{h}_{\rm ext} = 0$, and write the Helmholtz free energy as the functional

$$A[\vec{m}(\vec{r}), T] = \int d^d r \mathcal{A}(\vec{m}(\vec{r}), \vec{\nabla} \cdot \vec{m}(\vec{r}), \nabla^2 \vec{m}(\vec{r}), \ldots, T) \tag{3.31}$$

where the Helmholtz free energy density $\mathcal{A}$ has the most general form allowed by the symmetry of the problem and $\vec{m}(\vec{r}) \propto \langle \vec{s}(\vec{r}) \rangle$ is the magnetization. For simplicity we will now restrict our analysis to the case when the order parameter $m(\vec{r})$ is a scalar, and consider only the simplest version of the Landau-Ginzburg theory with

$$A[m(\vec{r}), T] = \int d^d r \left[\frac{1}{2} K (\vec{\nabla} m(\vec{r}))^2 + f(m(\vec{r}), T) \right] \tag{3.32}$$

where $f(m, T)$ is the free energy density in the Landau theory of the previous section (then denoted by $G(m, T)$). We have added an extra term $\frac{1}{2} K (\vec{\nabla} m(\vec{r}))^2$ which assigns a free energy cost to a spatially varying order parameter.

If the magnetization is constant everywhere in space, we know that the magnetic field is obtained as a derivative of the Helmholtz free energy with respect to the magnetization,

$$\delta A = h \delta m. \tag{3.33}$$

Now that we allow the magnetization to vary in space, it is natural to allow also the magnetic field to be spatially varying. The generalization of (3.33) to the spatially varying case is

$$\delta A = \int d^d r \, h(\vec{r}) \delta m(\vec{r}). \tag{3.34}$$

Thus, we need to find out how a small change of magnetization at location $\vec{r}$ changes the free energy — the magnetic field at $\vec{r}$ is defined as the corresponding linear response coefficient. Formally we write this as

$$h(\vec{r}) = \frac{\delta A}{\delta m(\vec{r})} \tag{3.35}$$

and define $h(\vec{r})$ to be the *functional derivative* of A with respect to $m(\vec{r})$.

As an example of how to calculate the functional derivative we will now consider the specific case

$$A[m(\vec{r}), T] = \int d^d r \left[\frac{1}{2} K(T) (\vec{\nabla} m(\vec{r}))^2 + a(T) + \frac{1}{2} b(T) m^2(\vec{r}) + \frac{1}{4} c(T) m^4(\vec{r}) \right] \tag{3.36}$$

and calculate the magnetic field $h(\vec{r})$. We set $m(\vec{r}) = m_0(\vec{r}) + \delta m(\vec{r})$ and expand everything to first order in $\delta m(\vec{r})$. That gives

$$\begin{aligned} & A[m_0(\vec{r}) + \delta m(\vec{r}), T] = A[m_0(\vec{r}), T] \\ & + \int d^d r \left[K(T) \left(\vec{\nabla} \delta m(\vec{r}) \cdot \vec{\nabla} m_0(\vec{r}) \right) + \left(b(T) m_0(\vec{r}) + c(T) m_0^3(\vec{r}) \right) \delta m(\vec{r}) \right]. \end{aligned} \tag{3.37}$$

We integrate the first term by parts, using $\vec{\nabla} \delta m \cdot \vec{\nabla} m_0 = \vec{\nabla} \cdot \left(\delta m \vec{\nabla} m_0 \right) - \delta m \nabla^2 m_0$, to obtain

$$
\begin{aligned}
&A[m_0(\vec{r}) + \delta m(\vec{r}), T] = A[m_0(\vec{r}), T] \\
&+ \int d^d r \delta m(\vec{r}) \left[-K(T) \nabla^2 m_0(\vec{r}) + b(T) m_0(\vec{r}) + c(T) m_0^3(\vec{r}) \right] \\
&+ \int d^d r \vec{\nabla} \cdot \left(\delta m(\vec{r}) \vec{\nabla} m_0(\vec{r}) \right) .
\end{aligned}
\tag{3.38}
$$

Gauss's theorem allows us to convert the last integral to an integral over the surface of system, and assuming that $\delta m(\vec{r})$ vanishes on the surface we obtain the result

$$
h(\vec{r}) = b(T) m(\vec{r}) + c(T) m^3(\vec{r}) - K(T) \nabla^2 m(\vec{r}) \tag{3.39}
$$

where we dropped the subscript 0. The equation (3.39) is a non-linear differential equation connecting the magnetic field and magnetization, and it can be viewed as an equation of motion for $m(\vec{r})$. Note that if the magnetization is constant everywhere the functional derivative reduces to the usual derivative. Thus, the functional derivative is calculated by considering a small deviation $\delta m(\vec{r})$ to $m(\vec{r})$, and expanding the functional $A[m(\vec{r})]$ so that the change δA assumes the form (3.34). This is easy (similar to calculating the usual derivative) if A only depends on $m(\vec{r})$, but if it depends on some derivatives of $m(\vec{r})$, e.g. $\vec{\nabla} m(\vec{r})$ or $\nabla^2 m(\vec{r})$, we must integrate by parts so that δA only depends on δm and not on the derivatives of δm.

The functional derivative has similar properties as the usual derivative, although sometimes the fact that the derivative is taken at a specific point in space leads to complications. For instance the usual property $\frac{dx}{dx} = 1$ is replaced by

$$
\frac{\delta f(x)}{\delta f(x')} = \delta(x - x') \tag{3.40}
$$

where $\delta(x)$ is the Dirac delta function — note that here the functional $I[f]$ is particularly simple and given by the value of f at the point x, $f(x)$. The chain rule $\frac{dx}{dy} = \frac{dx}{dz}\frac{dz}{dy}$ is generalized to

$$
\frac{\delta y(x)}{\delta m(x')} = \int dx'' \frac{\delta y(x)}{\delta h(x'')} \frac{\delta h(x'')}{\delta m(x')} . \tag{3.41}
$$

The functional derivative is not just a mathematical curiosity but actually quite a convenient tool as we will see in the next section.

c. Susceptibilities. Since in the Landau-Ginzburg theory both the magnetization and the magnetic field depend on the spatial coordinate $\vec{r}$, we can ask the question how does changing the magnetic field at position $\vec{r}$ affect the magnetization at position $\vec{r}\,'$? This response coefficient is a generalization of the usual susceptibility $\chi = \frac{\partial m}{\partial h}$; we define the space-dependent susceptibility as

$$
\chi(\vec{r}, \vec{r}\,') = \frac{\delta m(\vec{r})}{\delta h(\vec{r}\,')} . \tag{3.42}
$$

The space-dependent inverse susceptibility is defined as the matrix inverse of $\chi(\vec{r}, \vec{r}\,')$, i.e. $\chi^{-1}(\vec{r}, \vec{r}\,')$ is a function satisfying

$$\int d^d r'' \chi(\vec{r},\vec{r}'')\chi^{-1}(\vec{r}'',\vec{r}') = \delta^{(d)}(\vec{r}-\vec{r}') \tag{3.43}$$

where $\delta^{(d)}(\vec{r})$ is the d-dimensional Dirac δ-function. Using the chain rule of functional differentiation (3.41) with $y = m$ and the rule (3.40) we find

$$\chi^{-1}(\vec{r},\vec{r}') = \frac{\delta h(\vec{r})}{\delta m(\vec{r}')} \tag{3.44}$$

or

$$\chi^{-1}(\vec{r},\vec{r}') = \frac{\delta}{\delta m(\vec{r}')}\left[b(T)m(\vec{r}) + c(T)m^3(\vec{r}) - K(T)\nabla^2 m(\vec{r})\right]. \tag{3.45}$$

The first two terms are easy to functionally differentiate with respect to $m(\vec{r}')$ but the last term looks less familiar. That can be done by a simple trick: we write $m(\vec{r}) = \int d^d r'' \delta^{(d)}(\vec{r}-\vec{r}'')m(\vec{r}'')$ and interchange the order of integration over $\vec{r}''$ and functional differentiation with respect to $m(\vec{r}')$. That gives us

$$\chi^{-1}(\vec{r},\vec{r}') = \left[b(T) + 3c(T)m^2(\vec{r}) - K(T)\nabla^2\right]\delta^{(d)}(\vec{r}-\vec{r}'). \tag{3.46}$$

Now we have obtained an expression for the inverse susceptibility using functional differentiation. An alternative, more elementary derivation for the susceptibility would be to regard Eq. (3.39) as a non-linear differential equation for $m(\vec{r})$, write $m(\vec{r}) = m_0 + \delta m(\vec{r})$, expand the equation to the first order in δm, and solve the resulting differential equation for $\delta m(\vec{r})$.

The expression (3.46) looks rather complicated since the right hand side contains a second derivative of the delta function. However, it simplifies considerably if we Fourier transform both sides. We make the approximation that the fluctuations, i.e. the variations in $m(\vec{r})$, are small, and replace $m^2(\vec{r})$ by its average value m^2. We write $\delta^{(d)}(\vec{r}-\vec{r}') = \int \frac{d^d q'}{(2\pi)^d} e^{i\vec{q}'\cdot(\vec{r}-\vec{r}')}$ and $\chi(\vec{r}-\vec{r}') = \int \frac{d^d q'}{(2\pi)^d} e^{i\vec{q}'\cdot(\vec{r}-\vec{r}')}\chi(\vec{q}')$, multiply both sides by $e^{-i\vec{q}\cdot(\vec{r}-\vec{r}')}$ and integrate over $(\vec{r}-\vec{r}')$. That yields

$$\chi^{-1}(\vec{q}) = b(T) + 3c(T)m^2 + K(T)q^2. \tag{3.47}$$

Consequently, the susceptibility is given by

$$\chi(\vec{q}) = \frac{1}{b(T) + 3c(T)m^2 + K(T)q^2} \tag{3.48}$$

— in real space $\chi(\vec{r},\vec{r}')$ is the convolution of $\chi^{-1}(\vec{r},\vec{r}')$ so in Fourier space the two functions are reciprocals of each other. The expression for the susceptibility can be re-written as

$$\chi(\vec{q}) = \frac{\xi^2(T)}{K(T)}\frac{1}{1 + (q\xi(T))^2} \tag{3.49}$$

where

$$\xi(T) = \sqrt{\frac{K(T)}{b(T) + 3c(T)m^2}} \tag{3.50}$$

is a quantity that has the units of length. For reasons that will become clear in a moment $\xi(T)$ is called the *correlation length.* A useful connection between the correlation length and the free energy density is

$$K\xi^{-2}(T) = \left(\frac{\partial^2 f}{\partial m^2}\right)_{m=\overline{m}} \tag{3.51}$$

where $\overline{m}$ is the equilibrium magnetization that minimizes $f(m)$.

The susceptibility in real space is now obtained by Fourier transforming $\chi(\vec{q})$. We get

$$\chi(\vec{r}) = \frac{\xi^2(T)}{K(T)} \int \frac{d^d q}{(2\pi)^d} \frac{e^{i\vec{q}\cdot\vec{r}}}{1 + (q\xi(T))^2}. \tag{3.52}$$

Moving to spherical coordinates and changing variables to $s = \xi(T)q$ gives

$$\chi(\vec{r}) = \frac{1}{K(T)\xi^{d-2}(T)} \int \frac{d\Omega}{(2\pi)^d} \int_0^\infty ds\, s^{d-1} \frac{e^{is\frac{r}{\xi}\cos\theta}}{1+s^2} = \frac{1}{K(T)\xi^{d-2}(T)} Y_d\left(\frac{r}{\xi(T)}\right). \tag{3.53}$$

In three dimensions we have

$$\begin{aligned} Y_3(w) &= \int \frac{d\Omega}{(2\pi)^3} \int_0^\infty ds\, s^2 \frac{e^{isw\cos\theta}}{1+s^2} \\ &= \frac{1}{(2\pi)^2} \int_{-1}^{1} d(\cos\theta) \int_0^\infty ds\, s^2 \frac{e^{isw\cos\theta}}{1+s^2} \\ &= \frac{1}{(2\pi)^2} \int_0^\infty ds\, \frac{s}{iw} \frac{e^{isw} - e^{-isw}}{1+s^2} \\ &= \frac{1}{2(2\pi)^2} \int_{-\infty}^\infty ds \frac{s}{iw} \frac{e^{isw} - e^{-isw}}{1+s^2} \\ &= \frac{1}{2(2\pi)^2} 2\pi i \left[\frac{1}{2iw} e^{-w} + \frac{1}{2iw} e^{-w}\right] \\ &= \frac{1}{4\pi} \frac{1}{w} e^{-w} \end{aligned} \tag{3.54}$$

where on the next to last line I calculated the integral using the residue theorem. Substituting this into the previous expression gives the susceptibility in three dimensions,

$$\chi(\vec{r}) = \frac{1}{K(T)\xi^{(}T)} \frac{1}{4\pi} \frac{1}{r/\xi(T)} e^{-\frac{r}{\xi(T)}} = \frac{1}{4\pi K(T)} \frac{e^{-\frac{r}{\xi(T)}}}{r} \tag{3.55}$$

The susceptibility is a response function that tells how the magnetization responds to a change in a magnetic field. From the fluctuation-dissipation theorem we know that a response function is related to a correlation function. This connection is quite general and is not limited to Landau or Landau-Ginzburg theories. In the case of susceptibility the relevant correlation function is the magnetization-magnetization correlation function

$$C(\vec{r}, \vec{r}\,') = \langle (m(\vec{r}) - \langle m(\vec{r})\rangle)\,(m(\vec{r}\,') - \langle m(\vec{r}\,')\rangle)\rangle\,. \tag{3.56}$$

To prove the connection between $C(\vec{r}, \vec{r}\,')$ and $\chi(\vec{r}, \vec{r}\,')$ let us start from the partition function

$$Z(h,T) = \mathrm{Tr}\exp\left[-\beta\left(H_0 - \int d^d r\, h(\vec{r})m(\vec{r})\right)\right] \tag{3.57}$$

where H_0 is the Hamiltonian in the absence of the external magnetic field h, and the trace Tr means that we have to sum over all possible magnetizations $m(\vec{r})$. The thermal average of the magnetization at point $\vec{r}_0$ is given by

$$\langle m(\vec{r}_0)\rangle = \frac{1}{Z(h,T)}\mathrm{Tr}\left\{m(\vec{r}_0)\exp\left[-\beta\left(H_0 - \int d^d r h(\vec{r})m(\vec{r})\right)\right]\right\} \tag{3.58}$$

Taking the functional derivative of this expression with respect to $h(\vec{r}_1)$ gives

$$\begin{aligned} &\frac{\delta\langle m(\vec{r}_0)\rangle}{\delta h(\vec{r}_1)} \\ &= \beta\frac{1}{Z(h,T)}\mathrm{Tr}\left\{m(\vec{r}_0)m(\vec{r}_1)\exp\left[-\beta\left(H_0 - \int d^d r h(\vec{r})m(\vec{r})\right)\right]\right\} \\ &- \beta\langle m(\vec{r}_0)\rangle\langle m(\vec{r}_1)\rangle \\ &= \beta\left(\langle m(\vec{r}_0)m(\vec{r}_1)\rangle - \langle m(\vec{r}_0)\rangle\langle m(\vec{r}_1)\rangle\right) \\ &= \beta C(\vec{r}_0,\vec{r}_1). \end{aligned} \tag{3.59}$$

Thus, we have the correlation function in the Landau-Ginzburg theory in three dimensions,

$$C(\vec{r}) = \frac{k_B T}{4\pi K(T)}\frac{e^{-\frac{r}{\xi(T)}}}{r}. \tag{3.60}$$

This is an important result. It tells how spins, or magnetizations, in different parts of the sample are correlated. Notice that the correlations become weaker as the spins are further apart, as we would expect. The rate of this decay is controlled by the correlation length $\xi(T)$ which we found to be $\xi(T) = \sqrt{\frac{K(T)}{b(T)+3c(T)m^2}}$. In the case of a second-order transition[17] the magnetization goes to zero as $T \to T_c$, and so does $b(T)$, so the correlation length diverges at the critical temperature T_c. This divergence is usually algebraic, and therefore we define ν as the critical exponent which describes the divergence of the correlation length,

$$\boxed{\xi(T) \sim |T - T_c|^{-\nu}.} \tag{3.61}$$

In the Landau-Ginzburg theory $\nu = \frac{1}{2}$. The correlation length is the most important length scale in critical phenomena: if two regions of the system are separated by a distance longer than ξ then they can be thought of as being independent, but if the distance between them is less than ξ, the two regions influence each other significantly. Since right at the critical point the correlation length is infinite, the exponential in the numerator in (3.60) is equal to 1 at

---

[17] To analyze the first order transition we would again have to include the higher order term in the expansion.

T_c. Therefore, at the critical point the correlation function decays only algebraically rather than exponentially. In the Landau-Ginzburg theory the decay is like r^{-1} but in general the exponent could be different. Therefore we define yet another critical exponent (the last one!), η, which describes how the correlation function decays at the critical point,

$$\boxed{C(\vec{r}) \sim \frac{e^{-\frac{r}{\xi(T)}}}{r^{d-2+\eta}}.} \tag{3.62}$$

In the Landau-Ginzburg theory, and in mean field theories in general, $\eta = 0$.

Due to the approximation $m^2(\vec{r}) \approx m^2$ in the derivation of the susceptibility, our analysis of the Landau-Ginzburg theory is not exact, and must be regarded as a first correction to a mean field theory. A more thorough treatment would take us too deep into field theory, which is beyond the scope of this course.

d. Goldstone Modes. In the previous section we studied a system described by a scalar order parameter which was allowed to fluctuate. Similar methods can be applied to systems where the order parameter is a multicomponent vector $\vec{m}(\vec{r})$. In that case there are two different kinds of fluctuations: fluctuations that change the length of the order parameter vector (longitudinal fluctuations), and fluctuations that only change the direction of magnetization without affecting the length $|\vec{m}(\vec{r})|$ (transverse fluctuations). The fluctuations that we studied in the previous section were clearly longitudinal ones, and many of our results carry over to the analysis of longitudinal fluctuations in systems with multicomponent order parameters.

The transverse fluctuations are, however, a different matter. We are not going to go through the analysis here, but the end result is that the correlation function between transverse fluctuations is given by

$$C_\perp(\vec{q}) = \begin{cases} \frac{k_B T}{b(T)+K(T)q^2}, & T > T_c \\ \frac{k_B T}{K(T)q^2}, & T < T_c \end{cases} \tag{3.63}$$

Thus, below the critical temperature the transverse correlation function, and therefore the transverse susceptibility, is very large for small q. That means that a weak uniform transverse magnetic field is sufficient to change the direction of the magnetization. This is due to the fact that in the absence of an external magnetic field the free energy is spherically symmetric, but below the critical temperature the symmetry is spontaneously broken: the system has "chosen" a particular direction in which it has magnetized. Even a weak external magnetic field can rotate all the spins so that they align with the external field. The free energy of a model with a two-component order parameter is depicted in Fig. 17.

This means that there exists a very low-energy excitation in the system, an excitation the energy of which goes to zero as $q \to 0$. Such an excitation is guaranteed to exist in all systems with a broken continuous symmetry and no long range forces. These excitations that are manifestations of a spontaneously broken continuous symmetry are known as Goldstone modes (after Jeffrey Goldstone). There are many examples of Goldstone modes: in spin systems they are spin waves, in the liquid-solid transition they are transverse phonons (what continuous symmetry is broken when a liquid freezes?).

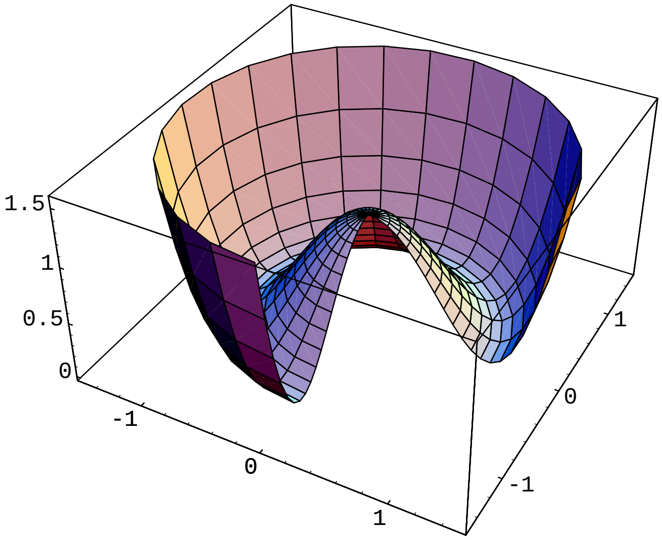


FIG. 17. Free energy as a function of two-component order parameter (m_x, m_y). A Goldstone mode corresponds to a rotation in the order parameter space, i.e. the direction of $\vec{m}$ varies smoothly as a function of $\vec{r}$ while the length $|\vec{m}(\vec{r})|$ is fixed to a value corresponding to the "trough" of the Mexican hat.

e. Ginzburg Criterion. Did we really have to go through the difficulty of introducing fluctuations to the Landau theory? In other words, are fluctuations important? That question was addressed by Ginzburg, and the corresponding criterion, which determines if fluctuations are important, is known as the Ginzburg criterion.

He argued that fluctuations are not necessary if the expectation value of $[m(\vec{r}) - \langle m(\vec{r})\rangle]^2$ is smaller than $\langle m^2(\vec{r})\rangle$. Since both these quantities could be spatially varying, it convenient to average them over some region of space. The most important length scale in critical phenomena is the correlation length ξ so it is natural to perform the averages over a correlation volume $V_\xi = \xi^d$. Thus, denoting this average $\mathrm{Avg}((\delta m)^2)$, we evaluate

$$\begin{aligned}
&\mathrm{Avg}((\delta m)^2) \\
&= \frac{1}{V_\xi^2}\int_{V_\xi} d^d r \int_{V_\xi} d^d r' \langle [m(\vec{r}) - \langle m(\vec{r})\rangle][m(\vec{r}\,') - \langle m(\vec{r}\,')\rangle]\rangle \\
&= \frac{1}{V_\xi^2}\int_{V_\xi} d^d r \int_{V_\xi} d^d r' C(\vec{r}, \vec{r}\,') \\
&= \frac{1}{V_\xi}\int_{V_\xi} d^d r C(\vec{r}, 0) \\
&= \xi^{-d}\int_0^\xi dr\, r^{d-1} \int d\Omega \frac{k_B T}{K(T)\xi^{d-2}} Y_d\left(\frac{r}{\xi}\right) \\
&= \xi^{-d+2}\frac{k_B T}{K(T)}\int_0^1 ds\, s^{d-1}\int d\Omega Y_d(s) \\
&= \xi^{-d+2}\frac{k_B T}{K(T)} A_d
\end{aligned} \tag{3.64}$$

where A_d is some dimensionless constant that depends on the dimensionality d. Similarly we get (for $T < T_c$)

$$\mathrm{Avg}(m^2) = -\frac{b}{c} \tag{3.65}$$

Thus, according to the Ginzburg criterion fluctuations are unimportant if

$$\xi^{-d+2}\frac{k_B T}{K(T)}A_d < -\frac{b}{c} \tag{3.66}$$

It is useful to replace the various quantities in terms of the measurable observables. For $T < T_c$ we have $\xi^2(T) = -\frac{K}{2b(T)} = \frac{K}{b_0(T_c - T)}$, and furthermore we have that the specific heat discontinuity at T_c is given by $\Delta c_h = \frac{b_0^2 T}{2c}$, see (3.28). Using these expressions to eliminate T and c we obtain the condition $\xi^{-d+4}(T) < \left(\frac{b_0 T_c}{K}\right)^2 \frac{k_B A_d}{\Delta c_h}$. Noticing that $\sqrt{b_0 T_c / K}$ has the units of length we can define it as the bare coherence length ξ_0 which is independent of temperature. With this change in notation we have arrived at the Ginzburg criterion stating that fluctuations are unimportant if

$$\left(\frac{\xi(T)}{\xi_0}\right)^{d-4} > \frac{k_B A_d}{\Delta c_h \xi_0^d}. \tag{3.67}$$

This can be expressed also as

$$\left(\frac{|T - T_c|}{T_c}\right)^{(4-d)/2} > \frac{k_B A_d}{\Delta c_h \xi_0^d}. \tag{3.68}$$

This condition is quite different depending on the sign of $(d-4)$. If $d > 4$, which cannot be realized in physical systems, the left hand side goes to infinity as T approaches T_c. Thus, in dimensions larger than four mean field theory works well in the vicinity of a critical point. In the more physical case $d < 4$ the left hand side goes to zero as T approaches T_c, and mean field theory works only sufficiently far from the critical point.

The crossover dimension above which a mean field theory becomes exact near a critical point is called the upper critical dimension d_c. It depends on the model we are studying. In general the condition for the validity of the mean field theory can be expressed as

$$d > d_c = \frac{2-\alpha}{\nu}. \tag{3.69}$$

In the simple Landau theory we have $\alpha = 0$ and $\nu = 1/2$, hence $d_c = 4$, as obtained above.

The temperature when the fluctuations become important is known as the Ginzburg temperature T_G. Setting both sides of (3.68) equal we have

$$|T_c - T_G| = T_c \left(\frac{k_B A_d}{\Delta c_h \xi_0^d}\right)^{2/(4-d)} \tag{3.70}$$

The bare correlation length ξ_0 plays a crucial role in determining how close to T_c we must go to see deviations from Landau theory. The specific heat jump is of the order of k_B per particle, and the dimensionless constant A_d is of the order of unity, so denoting the interparticle distance by a we get $|T_c - T_G| \approx T_c(a/\xi_0)^{2d/(4-d)}$. For instance in the case of the superfluid transition in He^4 the bare correlation length is set by the thermal wavelength at the transition temperature, which is about 0.1nm, and the interparticle separation is also about 0.1nm, so that $|T_c - T_G| \approx T_c$. Consequently, it is easy to see deviations from the mean field behavior in the case of the superfluid transition ($T_c \approx 2$K). In superconductivity ξ_0 is set by the size of a Cooper pair which is approximately 100nm and a is about 4nm, so that $|T_c - T_G| \approx (4/100)^6 T_c \approx 4 \times 10^{-9} T_c$, which is usually not observable since we cannot adjust the temperature accurately enough.

D. Beyond Mean Field Theories

1. One-Dimensional Ising Model

a. Transfer Matrix Method. The simplest spin model is the one-dimensional Ising model, which we have already encountered in the section on mean field theories. It turns out that the one-dimensional Ising model can be solved *exactly* using a transfer matrix technique, which can be applied to many other one-dimensional problems as well. We shall now familiarize ourselves with this technique and use the Ising model as an example.

The Hamiltonian for an Ising chain is given by

$$H = -J \sum_{i=1}^{N-1} s_i s_{i+1} - J_S s_N s_1 \tag{3.71}$$

where the last term is added so that we can consider both an open chain (by choosing $J_S = 0$) and a closed ring (by choosing $J_S = J$). Here $s_i = \pm 1$ is the Ising spin variable at site i. The quantity of interest is the partition function $Z(T) = \mathrm{Tr} e^{-\beta H}$, where the trace means that we have to sum over all possible states of the system, i.e. over all sets $\{s_i\}$ of the spin variables. To simplify the notation we write $K = \beta J$ and $K_S = \beta J_S$ so that the partition function assumes the form

$$Z(T) = \mathrm{Tr} e^{K \sum_{i=1}^{N-1} s_i s_{i+1} + K_S s_N s_1}. \tag{3.72}$$

Since the exponential of a sum is a product of exponentials, we have

$$Z(T) = \sum_{\{s_i\}} e^{K s_1 s_2} e^{K s_2 s_3} e^{K s_3 s_4} \ldots e^{K s_{N-1} s_N} e^{K_S s_N s_1} \tag{3.73}$$

Now we define the transfer matrices $\hat{T}$ and $\hat{T}^{(S)}$ with elements $\hat{T}_{s_i s_j} = e^{K s_i s_j}$ and $\hat{T}^{(S)}_{s_i s_j} = e^{K_S s_i s_j}$ so that we can write the partition function as

$$\begin{aligned} Z(T) &= \sum_{\{s_i\}} \hat{T}_{s_1 s_2} \hat{T}_{s_2 s_3} \ldots \hat{T}_{s_{N-1} s_N} \hat{T}^{(S)}_{s_N s_1} \\ &= \sum_{s_1} \sum_{s_2} \cdots \sum_{s_N} \hat{T}_{s_1 s_2} \hat{T}_{s_2 s_3} \ldots \hat{T}_{s_{N-1} s_N} \hat{T}^{(S)}_{s_N s_1}. \end{aligned} \tag{3.74}$$

The right hand side is now a matrix product, and it can be written as $\sum_{s_1} (\hat{T}^{N-1} \hat{T}^{(S)})_{s_1 s_1} = \mathrm{Tr}(\hat{T}^{N-1} \hat{T}^{(S)})$. Thus, we have

$$Z(T) = \mathrm{Tr}(\hat{T}^{N-1}\hat{T}^{(S)}). \tag{3.75}$$

This expression is easiest to evaluate if we first diagonalize $\hat{T}$ by finding a unitary matrix $\hat{U}$ such that $\hat{D} = \hat{U}\hat{T}\hat{U}^\dagger$ where $\hat{D}$ is a 2-by-2 diagonal matrix whose elements are the eigenvalues of $\hat{T}$. Since J_S is either 0 or J, the same transformation $\hat{U}$ diagonalizes also the matrix $\hat{T}^{(S)}$. Therefore, we have

$$Z(T) = \lambda_+^{N-1}\lambda_+^{(S)} + \lambda_-^{N-1}\lambda_-^{(S)} \tag{3.76}$$

where $\lambda_\pm$ and $\lambda_\pm^{(S)}$ are the eigenvalues of $\hat{T}$ and $\hat{T}^{(S)}$, respectively.

To find the eigenvalues explicitly we have to diagonalize the matrix

$$\hat{T} = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix} \tag{3.77}$$

which gives $\lambda_\pm = e^K \pm e^{-K}$. Since $\lambda_+ > \lambda_-$, it is useful to write

$$Z(T) = \lambda_+^{N-1}\lambda_+^{(S)} \left[1 + \left(\frac{\lambda_-}{\lambda_+}\right)^{N-1} \frac{\lambda_-^{(S)}}{\lambda_+^{(S)}} \right] \tag{3.78}$$

The free energy $F(T) = -k_B T \log(Z(T))$ is now given by

$$\begin{aligned} F(T) &= -k_B T (N-1)\log(\lambda_+) - k_B T \log(\lambda_+^{(S)}) - k_B T \log\left[1 + \left(\frac{\lambda_-}{\lambda_+}\right)^{N-1} \frac{\lambda_-^{(S)}}{\lambda_+^{(S)}} \right] \\ &\approx -k_B T (N-1)\log(\lambda_+) - k_B T \log(\lambda_+^{(S)}) - k_B T \left(\frac{\lambda_-}{\lambda_+}\right)^{N-1} \frac{\lambda_-^{(S)}}{\lambda_+^{(S)}} \end{aligned} \tag{3.79}$$

The first term on the right hand side is extensive, i.e proportional to N, the second term is a boundary term independent of N, and the third term vanishes for large N for $T \neq 0$ (at $T = 0$ we have $\lambda_+ = \lambda_-$). Consequently, the free energy per spin in the large N limit is given by

$$\frac{F(T)}{N} = f(T) = -k_B T \log(\lambda_+) = -k_B T \log(2\cosh(K)) \tag{3.80}$$

Note that the free energy is determined by the largest eigenvalue of the transfer matrix.

We can also calculate the magnetization at site i by evaluating (for simplicity we now consider a ring, i.e $\hat{T}^{(S)} = \hat{T}$)

$$\begin{aligned}
\langle s_i\rangle &= \tfrac{1}{Z}\mathrm{Tr}s_ie^{-\beta H}\\
&= \tfrac{1}{Z}\mathrm{Tr}\hat{T}_{s_1s_2}\ldots\hat{T}_{s_{i-1}s_i}s_i\hat{T}_{s_is_{i+1}}\ldots\hat{T}_{s_Ns_1}\\
&= \tfrac{1}{Z}\mathrm{Tr}\hat{T}_{s_1s_2}\ldots\hat{T}_{s_{i-1}s_i}\hat{\sigma}_z\hat{T}_{s_is_{i+1}}\ldots\hat{T}_{s_Ns_1}\\
&= \tfrac{1}{Z}\mathrm{Tr}\left(\hat{T}^{i-1}\hat{\sigma}_z\hat{T}^{N+1-i}\right)\\
&= \tfrac{1}{Z}\mathrm{Tr}\left(\hat{T}^{N}\hat{\sigma}_z\right)\\
&= \tfrac{1}{Z}\mathrm{Tr}\left[\begin{pmatrix}\lambda_+^N & 0\\ 0 & \lambda_-^N\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\tfrac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\right]\\
&= \tfrac{1}{Z}\mathrm{Tr}\begin{pmatrix}0 & \lambda_+^N\\ \lambda_-^N & 0\end{pmatrix}\\
&= 0,
\end{aligned} \tag{3.81}$$

where we used the explicit form of the diagonalizing transformation U. Thus, we find that the magnetization vanishes. We can also evaluate the correlation function $C(r) = \langle s_is_{i+r}\rangle$ by the same technique (details are left as an exercise) and we find

$$C(r) = \langle s_is_{i+r}\rangle \approx \left(\frac{\lambda_-}{\lambda_+}\right)^r + \left(\frac{\lambda_-}{\lambda_+}\right)^{N-r} \tag{3.82}$$

which can be written as (for $r \ll N$)

$$C(r) = e^{-r/\xi} \tag{3.83}$$

where the correlation length $\xi(T)$ is given by

$$\xi(T) = \frac{1}{\log(\lambda_+/\lambda_-)}. \tag{3.84}$$

Thus, the correlation length is determined by the ratio between the largest and the second largest eigenvalue of the transfer matrix. In the case of an Ising ring we have $\lambda_+ = 2\cosh(K)$ and $\lambda_- = 2\sinh(K)$ so that $\xi(T) \approx \frac{1}{2}e^{2K} = \frac{1}{2}e^{J/k_BT}$. Thus, the correlation length diverges as T approaches zero meaning that $T_c = 0$. Usually the divergence is a power law, $\xi(T) \sim (T - T_c)^{\nu}$, but in the case of an Ising ring we have found an essential singularity (divergence that is faster than any power law).

We have now established quite rigorously that at any non-zero temperature the correlation function $C(r)$ of a one-dimensional Ising model goes to zero as r grows to infinity. That means that spins separated by a large distance are uncorrelated: knowing the value of s_i does not give any information about the value of s_{i+r} if r is very large. Another way of saying this is that the one-dimensional Ising model has no long range order at any non-zero temperature. There is a physical way to understand why this must be the case: Let us consider the free energy at a low temperature. The ground state, or the state with lowest energy, is one where all spins are aligned, i.e $s_i = 1$ for all i or $s_i = -1$ for all i. This state clearly has long range order. The lowest lying excited state is one where $s_i = +1$ for $i \leq W$ and $s_i = -1$ for $i > W$ for some W, in other words, there is a domain boundary between the sites W and $(W + 1)$. The free energy cost of creating a domain wall is given by $\Delta F = \Delta E - T\Delta S$, where ΔE and ΔS are the changes in energy and entropy due to the appearance of a domain wall. The energy of a domain wall is $E_0 + 2J$, where E_0 is the ground state energy. The entropy is given by $k_B \ln(\Omega)$ where Ω is the number of ways to create the domain wall. If there are N sites, we can choose the position of a domain wall in $N-1$ ways, and therefore $\Delta F = 2J - k_B T \log(N-1)$. If the temperature is non-zero and N is large, the entropic contribution dominates, and $\Delta F < 0$. Therefore, since the equilibrium state is one with minimum free energy, domain walls are spontaneously generated at all positive temperatures, and the long range order is lost.

In two dimensions the situation is more complicated since the domain walls are lines rather than points. The energy cost is given by $2JL$, where L is the length of the domain wall. To estimate the number of ways of creating a domain wall of length L, i.e. $\Omega(L)$, we say that if the domain wall has reached a particular point in the lattice, it can continue to any neighboring point except the one it came from; thus, if the number of nearest neighbors is q, the number of choices is at each point along the domain boundary is approximately $(q-1)$, and therefore $\Omega(L)$ is roughly $(q-1)^L$. Consequently, $\Delta F = L(2J - k_B T \log(q-1))$. If the temperature is higher than $T_c \approx 2J/(k_B \log(q-1))$ domain walls are spontaneously generated, and long range order is lost. For a two-dimensional Ising model on a square lattice $q = 4$, and $T_c \approx 1.82 J/k_B$. The exact value is $T_c = [2/\log(1+\sqrt{2})]J/k_B \approx 2.27 J/k_B$, so our rough estimate is off by 20%, which is better than could be expected. Note that in deriving our estimate for $\Omega(L)$ we did not take into account complications arising from self-intersecting domain walls. Their exact treatment poses formidable obstacles, which are briefly discussed in the following sections.

2. Two-Dimensional Ising Model

a. Series Expansions. The two-dimensional Ising model is considerably more difficult to solve than its one-dimensional relative. However, an exact solution was found by the Norwegian physicist Lars Onsager in 1944. The two-dimensional Ising model will serve as a useful test case for many approximate approaches since it encompasses many of the features of more complicated theories and allows us to judge the accuracy of the approximations by comparing the results with the known exact solution.

One approximation strategy starts from the partition function $Z(T) = \mathrm{Tr} e^{-\beta H}$. If we are interested in the low-temperature properties of $Z(T)$, the most important contributions come from the ground state and the low-lying excited states. Consequently, we may write $Z(T)$ as a series where successive terms correspond to higher and higher excited states. The coefficients of the various terms reflect the number of ways of creating excitations with the appropriate energy. Since this multiplicity is often a combinatorial function of the excitation energy, the convergence properties of low-temperature expansions are frequently rather poor. However, by including more terms we may obtain a rather accurate picture of the low-temperature properties.

The low-temperature expansion can be used to obtain the low-temperature specific heat

$c_h = -\frac{\partial}{\partial T}(E/N) = \frac{J^2}{k_B T^2}\frac{\partial^2}{\partial K^2}\frac{\log(Z(K))}{N}$ where $K = \beta J$ and J is the Ising coupling constant. In the case of an Ising model the low-temperature expansion for the specific heat assumes the form of a power series $c_h = \sum_{l=0} a_l e^{-l2K}$. Since the specific heat diverges at the critical point, we can assume that near the critical point it has the form $c_h = A[1 - e^{2(K_c - K)}]^{-\alpha} = A[1 + \alpha e^{2K_c}e^{-2K} + \frac{\alpha(\alpha-1)}{2}e^{4K_c}e^{-4K} + \ldots]$. By comparing the two expansions we can determine the three unknowns α, A, and K_c. In particular, we have

$$\frac{a_l}{a_{l-1}} = e^{2K_c}\left[1 + \frac{\alpha - 1}{l}\right] \tag{3.85}$$

so plotting the ratio of consecutive coefficients a_l and fitting it to a straight line *vs.* l^{-1} gives both the critical temperature through K_c and the critical exponent α. In practice, however, this does not work quite so easily. The reason is that the form $A[1 - e^{2(K_c - K)}]^{-\alpha}$ gives only the leading divergence at the critical point, and there could be corrections that are either smooth functions of K near K_c or diverge less rapidly than the leading contribution. Both these corrections influence the series expansion near the critical point, and make the whole approach less rigorous. Surprisingly enough, the results are nevertheless quite good.

If, on the other hand, we are interested in the high-temperature properties, the inverse temperature β is small, and we may expand the exponential function $e^{-\beta H}$ as a power series, and obtain the partition function by calculating traces of powers of H, $Z = \sum_n \frac{(-\beta)^n}{n!}\mathrm{Tr}H^n$. The remaining traces are often easier to evaluate than the original partition function.

b. Duality. In two dimensions the series expansion method is more powerful than could be expected from general considerations. It actually allows us to determine the *exact* critical temperatures of Ising models on many different lattices. We do not get the exact partition function, or the exact free energy, so our solution does not rival the *tour de force* of Onsager.

We start by doing the low-temperature expansion of the d-dimensional Ising model on a hypercubic lattice (= square lattice if $d = 2$). The Hamiltonian is

$$H = -J\sum_{\langle ij\rangle} s_i s_j \tag{3.86}$$

where the sum runs over nearest neighbors. For simplicity we have set the external magnetic field to zero. The ground state is clearly one where all spins are aligned ($s_i = +1$). The number of nearest neighbor pairs, henceforth called bonds, is Nd — each spin has $2d$ nearest neighbors, so the number of bonds is $N \times 2d \times \frac{1}{2}$ where the last factor eliminates double counting. In ground state each bond contributes a $(-J)$ to the energy, and the ground state energy is therefore $E_0 = -NdJ$. The lowest-lying excited state is one where exactly one spin is flipped, see Fig 18. That affects $2d$ bonds (neglecting boundary effects), each of which contributes a $+J$ to the energy, giving an excitation energy of $2d(J - (-J)) = 4dJ$. The number of ways to create this excitation is N since each one of the N spins could be flipped. The next lowest-energy excitation is one where two neighboring spins are flipped. This affects $2(2d) - 1$ bonds; all but one of them contribute a $+J$ to the energy, the bond between the two flipped spins contributes a $-J$ as in the ground state. Consequently, the excitation energy is $(2(2d)-2)\times 2J = 4(2d-1)J$. This excitation can be created in $\frac{1}{2}N \times 2d$ ways (N spins, each has $2d$ nearest neighbors, $\frac{1}{2}$ because the spins could be chosen in either order). Therefore, the first few terms of the low-temperature expansion are

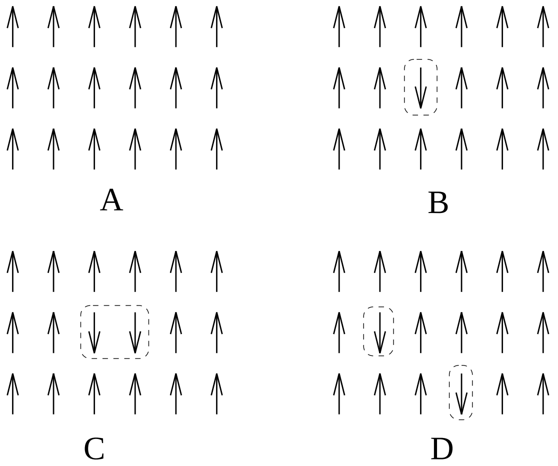


FIG. 18. A few low-lying excited states of the two-dimensional ferromagnetic Ising model on a square lattice: (A) the ground state, (B)-(D) domains of reversed spins with energies $8J$, $12J$, and $16J$ above the ground state energy, respectively.

$$Z_{\text{low}-\text{T}}(T) = e^{NdK}\left[1 + Ne^{-4dK} + Nde^{-4(2d-1)K} + \frac{1}{2}N(N-2d-1)e^{-8dK} + \cdots\right] \quad (3.87)$$

where the last term comes from configurations with two flipped non-neighboring spins.

We can think of the excited states as islands, or domains, of flipped spins; the excitation energy is $2J$ times the number of bonds between opposite spins. The only bonds that connect opposite spins occur at the boundary of the island of flipped spins, and the number of such bonds is simply given by the surface area of the island. In two dimensions, the islands are two-dimensional, and their boundaries are one-dimensional — "surface area" in this case means simply the perimeter of the island. The low-temperature expansion can now be written in the somewhat formal way as

$$Z_{\text{low}-\text{T}}(T) = e^{NdK}\sum_{\text{islands}} e^{-2K\times(\text{surface area of the island})} \quad (3.88)$$

This is our final expression for the low-temperature expansion.

Let us now consider the high-temperature expansion. We start by writing the partition function as

$$Z(T) = \sum_{\{s_i\}} e^{K\sum_{\langle ij\rangle} s_i s_j} = \sum_{\{s_i\}}\prod_{\langle ij\rangle} e^{Ks_i s_j} \quad (3.89)$$

Since the product $s_i s_j$ can only assume the values ± 1, we can use the equality $e^{Ks_i s_j} = \cosh(K) + \sinh(K)s_i s_j = \cosh(K)[1 + ts_i s_j]$ where $t = \tanh(K)$. At high temperatures $K = \beta J$ is small, and consequently $\tanh(K)$ is small, and we can use t as our expansion parameter. Since the factor $\cosh(K)$ occurs for each bond, we have

$$Z(T) = [\cosh(K)]^{\text{number of bonds}}\sum_{\{s_i\}}\prod_{\langle ij\rangle}[1 + ts_i s_j]. \quad (3.90)$$

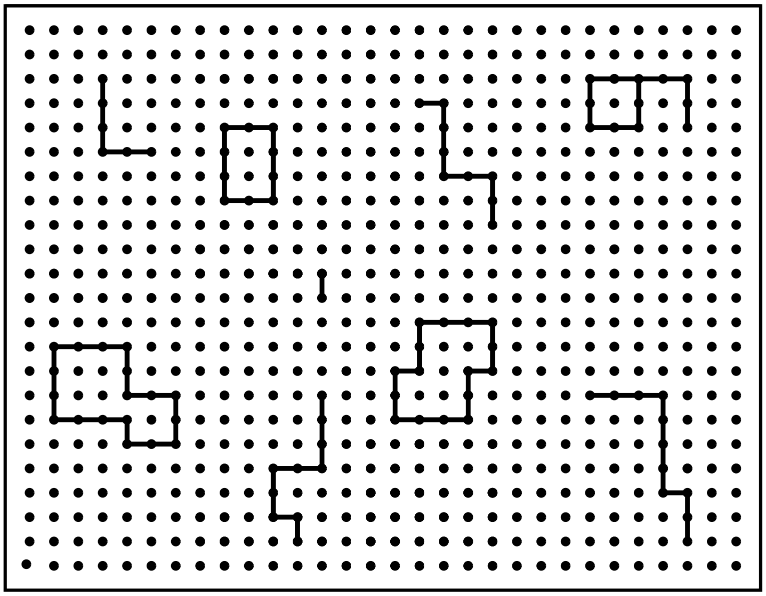


FIG. 19. Graphical representation of terms arising from the high-temperature expansion of the partition function of the two-dimensional Ising model on a square lattice

Our next step is to expand the product as a power series of t, and sum over all spin configurations. In expanding the product we face a choice every time we include a new term: we can choose to multiply the result thus far by either 1 or ts_is_j. It is useful to indicate this choice graphically, see Fig. 19, and join the points i and j with a line if we picked the term ts_is_j. If no lines go through the site i then we have chosen not to include any terms that contain the factor s_i, and therefore summing over the allowed values of s_i yields $\sum_{s_i=\pm 1} s_i^0 = 2$. If exactly one line goes through the site i, we have picked exactly one factor s_i, and the sum gives $\sum_{s_i=\pm 1} s_i^1 = 0$. In general, if an even number of lines go through the site i the sum over spins gives $\sum_{s_i=\pm 1} s_i^{2n} = 2$, and if an odd number of lines go through the site, the sum gives $\sum_{s_i=\pm 1} s_i^{2n+1} = 0$. Therefore, only those terms survive the sum over spin configurations in which each spin s_j appears an even number of times (considering 0 as an even number). Graphically that means that only closed graphs contribute: if the graph consisting of the lines we have drawn ends at a site i, then the spin s_i must appear an odd number of times in that particular term, and the contribution vanishes due to spin sum. Each line in a closed graph corresponds to a factor of t, and therefore the contribution from a particular closed graph is proportional to $t^{\text{number of bonds in the graph}}$. Finally, we have a factor of 2 arising from each spin sum. Therefore, we have another formal-looking result

$$Z_{\text{high-T}}(T) = [\cosh(K)]^{\text{number of bonds}} 2^{\text{number of sites}} \sum_{\text{closed graphs}} t^{\text{bonds in the graph}} \tag{3.91}$$

Now we have two formal expansions for the partition function, one at low temperatures and another at high temperatures, and their usefulness may not be evident at this point. However, consider the observation that *in two dimensions the perimeter of each island is a closed graph, and the interior of each closed graph is an island.* Therefore, the sums in the two expansions have exactly same structures, meaning that if we view them as power series the coefficients of all terms are the same in the two series! Consequently, we have for the two-dimensional Ising model on a square lattice the result

$$\begin{aligned} Z_{\text{low}-\text{T}}(T) &= e^{2NK} g(e^{-2K}) \\ Z_{\text{high}-\text{T}}(T) &= [\cosh(K)]^{2N} 2^N g(\tanh(K)) \end{aligned} \tag{3.92}$$

We do not know much about the function $g(x)$ but we know that it is the same function in the two expansions. Therefore, we can relate the low-temperature expansion of an Ising model on a square lattice to the high-temperature expansion of another Ising model (different coupling constant K). If we can relate the low-temperature expansion of one model to the high-temperature expansion of another model, we say that the two models are dual to each other. The Ising model on a square lattice is dual to itself, or self-dual. Another example of duality is that the Ising model on a hexagonal lattice is dual to an Ising model on a triangular lattice. Duality is not restricted to two dimensions, however, often the dual of an interesting model is not interesting itself and not much is gained in the process.

Self-duality of the two-dimensional Ising model on a square lattice allows us to relate the low-temperature properties of a model with coupling constant K to the high-temperature properties of a model with coupling constant $\widetilde{K}$ where $e^{-2K} = \tanh(\widetilde{K})$ or more simply

$$\boxed{\sinh(2K)\sinh(2\widetilde{K}) = 1} \tag{3.93}$$

Let us assume that there exists a critical coupling constant K_c such that the free energy is non-analytic at K_c. Since the factors multiplying the unknown function $g(x)$ in the expressions (3.92) are manifestly analytic, the non-analyticity must come from the unknown function $g(x)$. That means that if the free energy is non-analytic at coupling constant K_c, it must also be non-analytic at coupling constant $\widetilde{K_c}$. If there is only one value of the coupling constant at which the free energy is non-analytic, we must have $\widetilde{K_c} = K_c$. Therefore, assuming that the two-dimensional Ising model on a square lattice has exactly one critical point, it must be given by $\sinh^2(2K_c) = 1$ or more explicitly

$$\boxed{K_c = \frac{1}{2}\ln(1+\sqrt{2})} \tag{3.94}$$

which is the exact critical point of the two-dimensional Ising model on a square lattice. This result was obtained by Kramers and Wannier in 1941, three years before the complete solution by Onsager.

c. Onsager Solution. The two-dimensional Ising model was the first interacting model in two dimensions to be solved exactly. The exact solution is quite complicated, and we will not reproduce it here. There are several ways to find the solution; I will only sketch one method here.

An obvious starting point to obtain the partition function of the two-dimensional Ising model is either one of the two expansions we derived in the previous section — since the difficulty (determining the function $g(x)$) is the same in both cases, it does not matter if we start from the low-temperature or high-temperature expansion. For definiteness, let us consider the high-temperature expansion. The problem is then to evaluate the sum

$$g(t) = \sum_{\text{closed graphs}} t^{\text{bonds in the graph}} \tag{3.95}$$

We can start by approximating the sum over closed graphs by a sum over all closed random walks, see Fig. 20. This results in an overestimate of $g(t)$ since (i) graphs with intersections are overcounted, and (ii) a random walk can traverse the line segment between two neighboring sites several times whereas each line segment can appear at most once in a graph. The first problem can be dealt with if we assign a factor -1 to random walks with intersections; since the graph in Fig. 20(b) corresponds to three random walks, two of which are non-intersecting and one intersects itself, the total contribution from the random walks is 1+1+(-1) = 1, and we have eliminated the overcounting. The second problem — restricting random walks to at most one passage of each line segment — is more complicated since it requires that the random walker remembers where it has been, which results in a non-Markovian process.

It turns out, however, that this problem can be overcome, and in the end of a very long day we get the result

$$\frac{\log(Z)}{N} = \log[2\cosh^2(K)] + \frac{1}{2}\int \frac{dq_x dq_y}{(2\pi)^2} \log\left[(1+t^2)^2 - 2t(1-t^2)(\cos(q_x)+\cos(q_y))\right] \tag{3.96}$$

where the integral runs over the first Brillouin zone and $t = \tanh(K)$. The critical point corresponds to a singularity (non-analyticity) of Z which arises when the argument of the logarithm in the integrand vanishes. That can only happen for $q_x = q_y = 0$ (otherwise the argument of the logarithm is always positive), which gives the condition

$$0 = (1+t_c)^2 - 4t_c(1-t_c^2) = (1-t_c^2-2t_c)^2 \tag{3.97}$$

for the critical value of t. The solution is $t_c = -1+\sqrt{2}$ or $K_c = \frac{1}{2}\log(1+\sqrt{2})$, which agrees with the solution we found by using duality.

To examine the singular part of Z in more detail we expand the integrand for small q near t_c and obtain

$$\begin{aligned}\left(\frac{\log(Z)}{N}\right)_{\text{sing}} &= \frac{1}{2}\int \frac{d^2q}{(2\pi)^2}\log[q^2+4(\delta t)^2] \\ &= \frac{1}{8\pi}\int_0^{\Lambda} 2q\,dq \log[q^2+4(\delta t)^2] \\ &= \frac{1}{8\pi}\left[(q^2+4(\delta t)^2)\{\log[q^2+4(\delta t)^2]-1\}\right]_{q=0}^{q=\Lambda}\end{aligned} \tag{3.98}$$

where $\delta t = t - t_c$ and we kept only the singular terms at small q. We also approximated the Brillouin zone by a disk of radius Λ; since the imporatnt contribution comes from small wavevectors, this approximation is of little importance. After the substitution we find

$$\left(\frac{\log(Z)}{N}\right)_{\text{sing}} = -\frac{1}{\pi}|\delta t|^2 \log|\delta t|. \tag{3.99}$$

It is more useful to write this in terms of the temperature T and the deviation from the critical temperature T_c. The deviation $(T - T_c)$ is linearly proportional to $\delta t =$

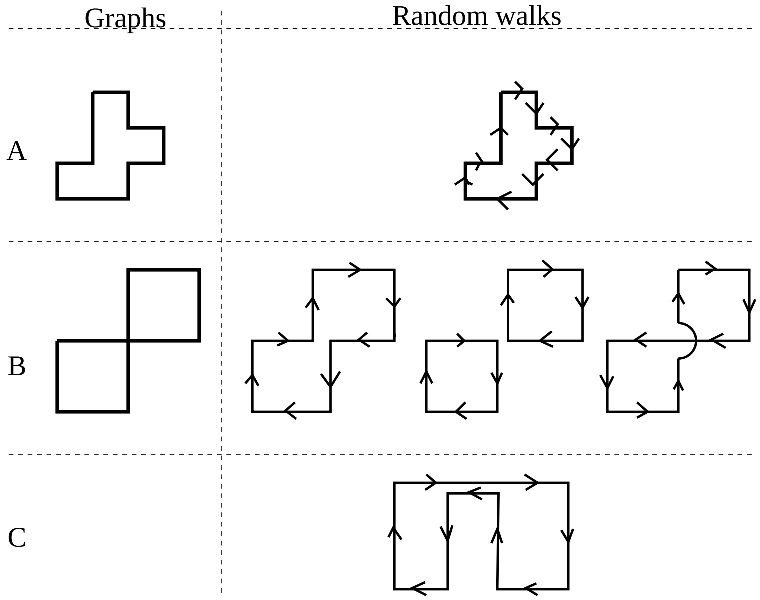


FIG. 20. Some graphs contributing to the high-temperature expansion of the two-dimensional Ising model on a square lattice, and the corresponding random walks. The graph (b) corresponds to three different random walks, and a factor of -1 for each self-intersection is needed to eliminate overcounting. The random walk in (c) has no corresponding graph, and should not be included in the expansion.

$(\tanh(K) - \tanh(K_c)) = (\tanh(J/(k_B T)) - \tanh(J/(k_B T_c)))$, so we have $\left(\frac{\log(Z(T))}{N}\right)_{\text{sing}} \sim -|T-T_c|^2 \log|T-T_c|$. The specific heat is proportional to the second derivative of $\log(Z)$ with respect to T, which gives

$$c_h \sim -\log\left|\frac{T-T_c}{T_c}\right| \tag{3.100}$$

Thus, the specific heat is only logarithmically divergent at the critical point, and therefore the specific heat exponent $\alpha = 0$. The other critical exponents can also be calculated exactly, and we have $\beta = \frac{1}{8}$, $\gamma = \frac{7}{4}$, $\delta = 15$, and $\nu = 1$. Just for curiosity, the spontaneous magnetization is given exactly by the simple expression [5]

$$m(T,0) = \begin{cases} 0, & T > T_c \\ \left[1 - \sinh^{-4}(2K)\right]^{1/8}, & T < T_c \end{cases} \tag{3.101}$$

E. General Properties of Critical Phenomena

1. Scaling Hypothesis

In the previous sections we have encountered several ways to analyze the behavior of an arbitrary system near criticality. The starting points in the various analyses have been expressions for the free energy (or equivalently the partition function) near the critical point, and end results have been critical exponents and critical temperatures. Now we will attempt to do the reverse: given a set of critical exponents, what is the proper expression for the free energy? Obviously the answer cannot be unique since many different systems have similar critical exponents, but nevertheless we will acquire some useful information.

Let us consider a system that is described with two independent thermodynamic variables, which we can take to be temperature and the external magnetic field. The free energy per particle (or spin) is a function of these two variables, $f = f(t, h)$. In the Landau theory we had $f_L = \min_m \left[\frac{t}{2}m^2 + um^4 - hm\right]$ which has the limiting behaviors

$$f_L \sim \begin{cases} -\frac{t^2}{u}, h = 0 \\ -\frac{h^{4/3}}{u^{1/3}}, t = 0. \end{cases} \tag{3.102}$$

We can write this in the form

$$f \sim t^2 g\left(\frac{h}{t^\Delta}\right) \tag{3.103}$$

where $g(x)$ is an arbitrary function satisfying $\lim_{x\to 0} g(x) = \text{constant}$ and $\lim_{x\to\infty} x^{-4/3} g(x) = \text{constant}$ (where the two constants are not necessarily equal). The first constraint on $g(x)$ comes from the $h = 0$ behavior, and the second constraint can be used to satisfy the $t = 0$ behavior of $f(t, h)$: for small t we have $f(t, h) \sim t^2 \left(\frac{h}{t^\Delta}\right)^{4/3} \sim t^{2-4\Delta/3} h^{4/3}$ and to satisfy the behaviors (3.102) we must have $\Delta = 3/2$. Thus, we have found that the Landau free energy can be written as $f(t, h) = t^2 g(h/t^{3/2})$ where $g(x)$ is an unknown function with a known limiting behavior.

The homogeneity, or scaling, assumption states that for any phase transition, the singular part of the free energy has the form

$$f_{\text{sing}} \sim t^{2-\alpha} g\left(\frac{h}{t^\Delta}\right) \tag{3.104}$$

where Δ is called the gap exponent. This assumption has far reaching consequences. First, using $\langle E\rangle = -\frac{1}{Z}\frac{\partial}{\partial\beta}Z = F(T) - T\frac{\partial F}{\partial T}$ we obtain $C_h(T) = -T\frac{\partial^2 F}{\partial T^2}$. Near the critical point this implies $C_h(T) \sim -\frac{\partial^2 f}{\partial t^2}$, which gives $C_{h=0} \sim t^{-\alpha}$ — thus, the parameter α in (3.104) is nothing but the specific heat exponent. The magnetization is given by $m = \frac{\partial f}{\partial h} \sim t^{2-\alpha-\Delta} g'(h/t^\Delta)$. In the limit of vanishing external field this gives (assuming that $g'(0)$ is regular) $m \sim t^{2-\alpha-\Delta}$ and the magnetization exponent is given by $\beta = 2 - \alpha - \Delta$. In the other limit, $t \to 0$, we must assume that the derivative $g'(x)$ behaves as x^p for large x. That gives us $m \sim t^{2-\alpha-\Delta}\left(h/t^\Delta\right)^p$; since the magnetization in a finite field near the critical point must be finite (i.e. not ∞ and not 0), we must have $p = (2 - \alpha - \Delta)/\Delta$, which gives $m \sim h^{1/\delta}$ where $\delta = \frac{\Delta}{2-\alpha-\Delta} = \frac{\Delta}{\beta}$. The susceptibility is defined as $\chi = \frac{\partial m}{\partial h} \sim t^{2-\alpha-2\Delta} g''(h/t^\Delta)$ and for $h = 0$ we get $\chi \sim t^{2-\alpha-2\Delta}$ or $\gamma = 2\Delta - 2 + \alpha$.

Thus, the other critical exponents can be expressed in terms of the two independent exponents α and Δ. Since the exponent Δ is not directly measurable it is useful to derive

d	n	α	β	γ	δ	ν	η	$\alpha+2\beta+\gamma$	$1+\gamma/\beta$
3	1	0.12	0.31	1.25	5	0.64	0.05	1.99	5.03
3	2	0.00	0.33	1.33	5	0.66	0.00	1.99	5.03
3	3	-0.14	0.35	1.40	5	0.70	0.04	1.96	5.00
2	1	0	1/8	7/4	15	1	1/4	2	15
MFT	MFT	0	1/2	1	3	1/2	0	2	3

TABLE I. Scaling laws and critical exponents for Ising, XY, and Heisenberg models ($n = 1, 2, 3$) in two and three dimensions, and the classical mean field theory predictions [4]

relations between the measurable exponents. These relations are known as scaling relations, two of which are

$$\boxed{\begin{array}{lcl} \alpha+2\beta+\gamma & = & 2 \\ \delta-1 & = & \frac{\gamma}{\beta}. \end{array}} \tag{3.105}$$

The first one of these is known as the Rushbrooke scaling law, and the second one is the Widom scaling law.

In the form (3.104) the scaling hypothesis suffers from one difficulty: since the exponents α and Δ need not be integers, the analytic structure of the *Ansatz* can be rather complicated. It is therefore often better to work with the more general form

$$f(t,h) \sim |t|^{2-\alpha_\pm} g_\pm\left(\frac{h}{|t|^{\Delta_\pm}}\right) \tag{3.106}$$

where the subscript ± indicates that the form could be different for $t > 0$ and $t < 0$. By requiring that the two expansions agree for $t = 0$, $h \neq 0$, you can conclude that $\alpha_- = \alpha_+$ and $\Delta_- = \Delta_+$, and all critical exponents are the same on both sides of the transition. The amplitudes, however, may differ, as we already saw in the mean field analysis.

The scaling relations were derived from the rather artificial-looking homogeneity or scaling hypothesis. There is no proof for the hypothesis, but it compares quite favorably with the best known estimates (or exact values) of critical exponents for the various $O(n)$ models as listed in the Table III E 1. Thus, it appears that the hypothesis works, but at this point we do not understand why. The hypothesis will be made more plausible in the following chapters on block spins and renormalization but at this point we can merely take it as an experimental fact and as a useful starting point for the phenomenology of phase transitions.

We are now in position to give the first plausibility argument supporting the scaling hypothesis of the previous section. Let us consider an arbitrary spin model on a lattice, $H = H[\{S_i\}]$, and consider possible critical phenomena associated with this Hamiltonian. We focus our attention on the free energy per spin, $f(t, h)$, which is a function of both the reduced temperature t and the magnetic field h.

One characteristic property of critical points is that the correlation length ξ diverges — hence, the spins are correlated over large distances. Therefore, we follow Leo Kadanoff's intuition from 1967 and construct a set of block spin variables $S_j^{(b)}$ which correspond to several physical spins. The block spins must be the same kind of objects as the original spins S_i, *i.e.* if the original Hamiltonian describes Ising spins $S_i = \pm 1$, also the block spins must be Ising, $S_j^{(b)} = \pm 1$. We introduce a many-to-one mapping from the original spins to the block spins by going from the original lattice with lattice constant a to a new lattice with lattice constant Λa (where Λ is a real number) — in other words, we coarse-grain the system by a factor Λ, see Fig. 21. In the Ising case we set the block spin on the coarse-grained lattice to $S_j^{(b)} = +1$ if the majority of the original Λ^d spins within the coarse-grained unit cell was up, and we set the block spin to -1 if the majority of the original spins was down — if there were as many down spins as up spins we can choose the block spin arbitrarily.

So far we haven't really done anything except defined a new set of variables. The rationale is that if the system is close to criticality, the correlation length ξ is much larger than either a or Λa, and furthermore ξ is assumed to be the only relevant length scale in the problem. That implies that as long as the blocks are much smaller than ξ, the internal structure of the blocks is irrelevant (and, to a good approximation, non-existent). Therefore, the free energy of the block spin system can be expected to have the same form as the free energy of the original spin system — in other words, as long as our Hamiltonian describes physics on a length scale that is much shorter than ξ, it does not matter what the length scale actually is: the form of the Hamiltonian is scale invariant. The parameters t' and h' that describe the block spin system are, however, in general different from the original parameters. The reason is that when we analyze a lattice Hamiltonian and calculate a quantity like ξ that has the dimensions of length, our result always has the form (a number)×(lattice constant), and since we have increased the lattice constant by a factor of Λ in the coarse-graining process, the dimensionless number in the expression for ξ must be reduced. Since this number only depends on the reduced temperature and the magnetic field, at least one of them must change in the coarse-graining process. We do not know how t and h change so we make the simplest possible assumption and write

$$\begin{cases} h' &= \Lambda^x h \\ t' &= \Lambda^y t. \end{cases} \tag{3.107}$$

This assumption satisfies the constraints that (i) if the external magnetic field h is zero, so is the new field h' (ii) if the original system is critical ($h = t = 0$), so is the block spin system.

Since we argue, following Kadanoff, that the block spin representation is just another description of the underlying physical system, the total free energy of the system must be

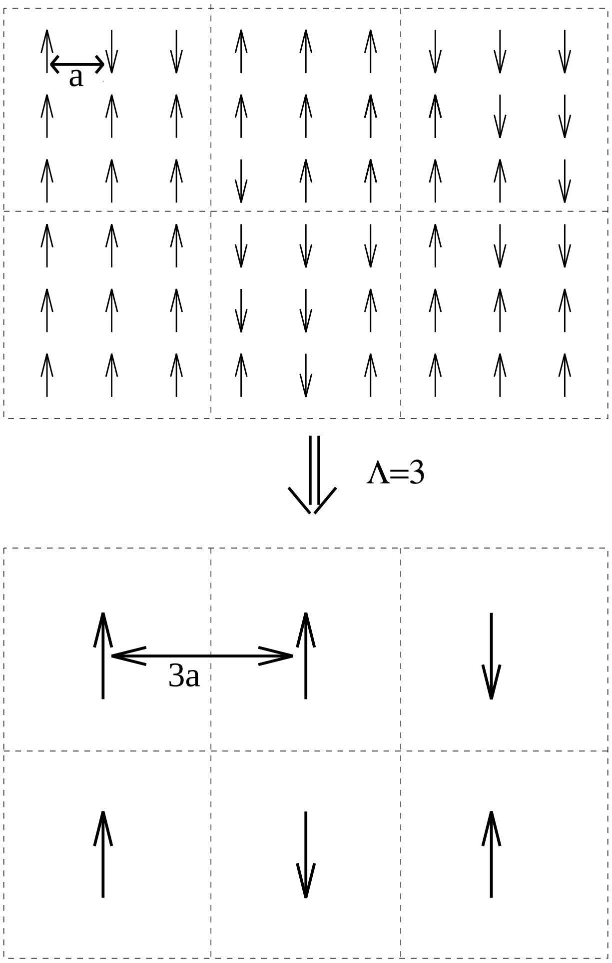


FIG. 21. Coarse-graining of the original lattice and the corresponding block spins

the same regardless of whether we calculate it in the original representation or in the block spin description. Since the number of block spins is Λ^{-d} times the number of the original spins, the free energy per spin in the two descriptions must satisfy

$$f(t,h) = \Lambda^{-d} f(\Lambda^y t, \Lambda^x h). \tag{3.108}$$

Taking into account the fact that the lattice constants in the two descriptions differ by the factor Λ we conclude that the correlation length, measured in units of lattice spacings, satisfies

$$\xi(t,h) = \Lambda \xi(\Lambda^y t, \Lambda^x h). \tag{3.109}$$

Let us now choose the rescaling factor Λ in a convenient way. By setting $\Lambda = |t|^{-1/y}$ in (3.108) we obtain $f(t,h) = t^{d/y} f(\pm 1, h/t^{x/y})$ which is precisely the scaling form used in the previous chapter if we identify $2-\alpha = d/y$ and $\Delta = x/y$ — note that we automatically find that the functional form may be different above and below the transition temperature. Similarly setting $\Lambda = |t|^{-1/y}$ in (3.109) yields for $h=0$ a scaling form for the correlation length $\xi(t, h=0) = t^{-1/y}\xi(\pm 1, 0)$. Eliminating y from these equations yields the hyperscaling law

$$\boxed{d\nu = 2-\alpha}, \tag{3.110}$$

which is also known as the Josephson scaling relation.

We can derive one more relation between the critical exponents. This time our starting point is the definition of the critical exponent η which describes the behavior of the correlation function $C(r) = \langle S(0)S(\vec{r})\rangle - \langle S(0)\rangle\langle S(\vec{r})\rangle$,

$$C(r) \sim \frac{e^{-r/\xi}}{r^{d-2+\eta}}. \tag{3.111}$$

From the fluctuation-dissipation theorem we know that the (static) susceptibility is given by the Fourier transform of the correlation function, which in the $k=0$ limit gives, by changing variables $\vec{r} \to \xi\vec{x}$,

$$\chi(t,h) = \frac{1}{k_B T}\int d^d r C(r) = [\xi(t,h)]^{2-\eta}\frac{S_{d-1}}{k_B T}\int_0^\infty dx \frac{e^{-x}}{x^{d-2+\eta}} \tag{3.112}$$

where S_{d-1} is the surface area of a d-dimensional sphere. Thus, using $\xi \sim |t|^{-\nu}$ we have the Fisher scaling relation $\boxed{\gamma = \nu(2-\eta)}$.

The Kadanoff block spin picture is intuitively appealing, and leads to experimentally verified relations between critical exponents, but unfortunately the block spin picture suffers from some major drawbacks. Firstly, even if we make sure that the block spin is the same kind of an object as the original spin, it is by no means obvious that the Hamiltonian describing the block spin system is similar to the original Hamiltonian, and consequently the validity of (3.108) is not clear. Secondly, while it is true that near the critical point spins are correlated over a large distance ξ, it is not obvious that the underlying lattice constant a is unimportant. These two objections are connected, and can be removed by a more detailed analysis as we will show in the sections on renormalization group. At this point, we can think of the block spin picture as a plausibility argument supporting the scaling hypothesis.

Let us now ignore the objections at the end of the previous section and apply the scaling hypothesis to an arbitrary experimental setup. It will prove to be a powerful tool to organize data and to extract information from a limited set of measurements or simulations.

We saw that the scaling laws followed from the scaling assumption, which itself could be motivated by assuming that the correlation length was the only important length scale near a phase transition. When we apply this idea to real experimental systems we must, however, be somewhat careful. All experiments are performed on finite systems, and therefore the system size sets another macroscopic length scale in addition to the correlation length. While the finiteness of the sample is of no consequence as long as the system is much larger than ξ, things may be different in the vicinity of a critical point where ξ becomes very large. Therefore, we must allow for the possibility that the scaling laws are modified due to the finiteness of the system. This is accomplished by writing them in the form

$$\begin{cases} \chi(L,t) &= |t|^{-\gamma} g_\chi\left(\frac{L}{\xi(t)}\right) \\ C(L,t) &= |t|^{-\alpha} g_c\left(\frac{L}{\xi(t)}\right) \end{cases} \tag{3.113}$$

where $g_\chi(x)$ and $g_c(x)$ are unknown functions that satisfy $\lim_{x\to\infty} g_{\chi,c}(x) =$ constant as well as $\lim_{x\to 0} x^{-\gamma/\nu} g_\chi(x) =$ constant and $\lim_{x\to 0} x^{-\alpha/\nu} g_c(x) =$ constant. The first condition guarantees that if the system size is much larger than the correlation length the system behaves as if it were infinite. The second condition, which specifies the small-x behavior of the unknown functions, guarantees that the susceptibility or specific heat of a finite system remains finite even if t approaches 0 (that is, even as T approaches T_c) — remember that $\xi(t) \sim |t|^{-\nu}$. In a finite system thermodynamic quantities do not diverge at the critical point but instead have local maxima[18] at $t = 0$.

At first it appears that the finiteness of experimental systems introduces an extra complication in the study of critical phenomena. We can, however, turn this weakness into a strength. The way to do this is to re-write *e.g.* the expression for the susceptibility as

$$L^{-\gamma/\nu}\chi(L,t) = \left(L^{1/\nu}|t|\right)^{-\gamma} g\left[\left(L^{1/\nu}|t|\right)^{\nu}\right] \tag{3.114}$$

where I used $\xi(t) \sim |t|^{-\nu}$ and multiplied both sides by $L^{-\gamma/\nu}$. The beauty of this equation is that the right hand side depends only on the combination $L^{1/\nu}|t|$. Therefore, if we collect data on systems of different sizes, and plot $L^{-\gamma/\nu}\chi(L,t)$ versus $L^{1/\nu}|t|$, all our data falls on one curve! This is known as the data collapse method, and it is extremely useful in organizing both real experimental data and "data" obtained by computer simulations of finite systems. In practice we do not usually know the quantities T_c, γ, and ν, but instead we must work backwards and find a set of parameters $\{T_c, \gamma, \nu\}$ such that all our data indeed does collapse on a single line as indicated in Fig. 22.

---

[18] Consider for instance that classically the heat capacity is approximately $\frac{1}{2}k_B\times$(number of degrees of freedom) which is finite if the number of degrees of freedom is finite.

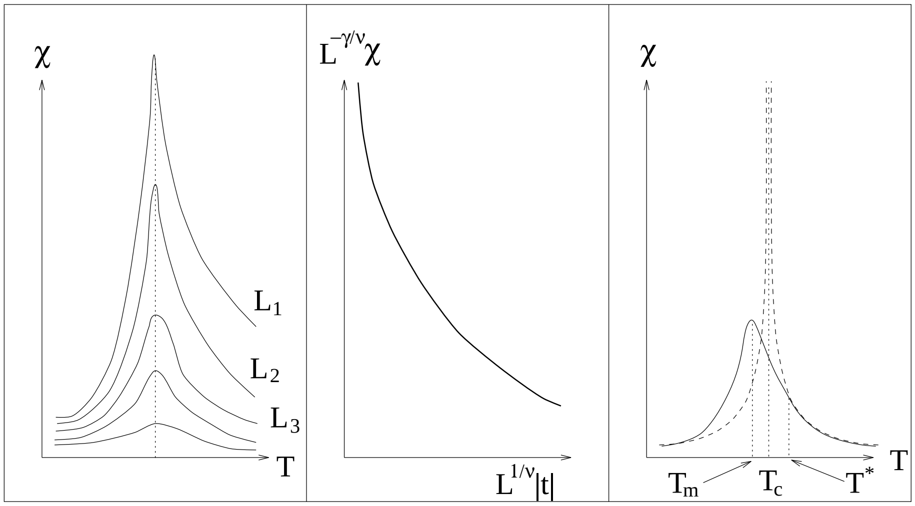


FIG. 22. Finite size scaling: (a) susceptibility data on samples of different sizes, (b) the data collapse method giving critical exponents ν and γ and the critical temperature T_c, and (c) a more complicated case with a size-dependent pseudocritical temperature $T_m(L)$ and rounding temperature $T^*(L)$. In (c) the dashed line is the infinite system susceptibility.

Often we encounter an additional complication when we try to apply finite-size scaling ideas to a specific system: the susceptibility $\chi(L,T)$ may peak at different temperatures depending on the system size, see Fig. 22(c). Since in infinite systems the peak of $\chi(T)$ occurs at T_c (where χ is divergent), it is natural to call the temperature $T_m(L)$ where susceptibility assumes its maximum a pseudocritical temperature. Obviously $T_m(L) \to T_c$ as $L \to \infty$. The pseudocritical temperature approaches T_c as a power law [10]

$$\frac{|T_m(L) - T_c|}{T_c} \sim L^{-\lambda} \tag{3.115}$$

where λ is known as a shift exponent. Another, related quantity is the rounding temperature $T^*(L)$ which is defined as the temperature where the finite system susceptibility differs significantly from the infinite system susceptibility. Obviously again $T^*(L) \to T_c$ as $L \to \infty$, and we therefore hypothesize that

$$\frac{|T^*(L) - T_c|}{T_c} \sim L^{-\theta} \tag{3.116}$$

Since we have assumed that ξ is the only relevant length scale in the problem, it is natural to assume that the rounding temperature corresponds to a temperature where the correlation length is approximately equal to the system size, $\xi(T^*(L)) \approx L$. If $\xi(T) \sim |T - T_c|^{-\nu}$, we have $|T^*(L) - T| \sim L^{-1/\nu}$ and therefore $\theta = 1/\nu$. Similar arguments can be applied to the pseudocritical temperature and yield $\lambda = 1/\nu$. Therefore, assuming $T_m(L) - T_c = T_c\kappa L^{-1/\nu}$ where κ is some unknown proportionality constant, the scaling law may be modified to give

$$L^{-\gamma/\nu}\chi(L,T) = \left|L^{1/\nu}t - \kappa\right|^{-\gamma} g\left(L^{1/\nu}|t|\right). \tag{3.117}$$

a. Dimensional Crossover. Various critical properties, including critical exponents, depend on the dimensionality of the system we are studying, but what determines the dimensionality? The simple answer is that fundamentally all experimental systems are three dimensional, but it is not obvious that this is the correct answer in the present context. Dimensionality enters the discussion of a statistical physics problem essentially as a scaling parameter which determines how the number of (interacting) objects increases with distance: in one dimensional systems the number of objects within a range l of a given site increases linearly with l, in two dimensions it increases as l^2, and in three dimensions as l^3. Seen this way, the dimensionality may depend on the length scale l that we are interested in. In critical phenomena the interesting length scale is set by the correlation length ξ and therefore the dimensionality must be determined by considerations at that length scale.

Let us consider a piece of magnetic material with dimensions $L_x \times L_y \times L_z$, and ask how the susceptibility varies with temperature. For definiteness we assume $L_z > L_y > L_x$. At very high temperatures, well above the critical temperature, the correlation length $\xi(T)$ is very short. Consequently, the most part ("bulk") of the sample is further than a distance ξ from all boundaries, and the system is effectively three-dimensional. As the temperature is lowered, the correlation length grows, and at some temperature, call it T_2, the correlation length is approximately equal to the shortest dimension of the sample, L_x. If we now take a look at a narrow columnar region of the sample in the x-direction, its length L_x is approximately ξ, and therefore spins at the two ends of the column are correlated. Consequently, spins within the column "act as one" and we can think of the three-dimensional sample as consisting of a two-dimensional system of columns ("spaghetti"): the effective dimensionality has been reduced from 3 to 2. Similarly, when the temperature is lowered further, ξ grows, and at some temperature T_1 we have $\xi(T) \approx L_y$. Now spins are correlated not only in the x-direction but also in the y-direction, and only free degree of freedom is in the z-direction. The system can be thought of as consisting of a one-dimensional collection of planes in the xy-direction ("lasagna"), and the effective dimensionality has been reduced from 2 to 1. Finally, when the temperature is reduced further, at temperature T_0 the correlation length ξ exceeds even the longest sample dimension L_z, and spins within the whole system are correlated. The effective dimensionality has been reduced to zero, and we must apply the finite size scaling ideas of the previous section. The changing of the effective dimensionality as a function of temperature is known as dimensional crossover, which is a well-studied effect, especially in layered magnets.

The susceptibility $\chi(T)$ varies according to a power law $\chi(T) = |T - T_c|^{-\gamma(d)}$, so plotted on a log-log plot χ versus temperature is a straight line. However, the exponent $\gamma(d)$, and therefore the slope of the line, depends on the dimensionality d. We have just concluded that d itself depends on temperature — hence, the slope of χ *vs.* $|T - T_c|$ on a log-log plot changes at temperatures T_2, T_1, and T_0, as indicated in Fig. 23. By determining the kink positions T_2, T_1, and T_0 we can get rough estimates for the temperature dependence of the correlation length.

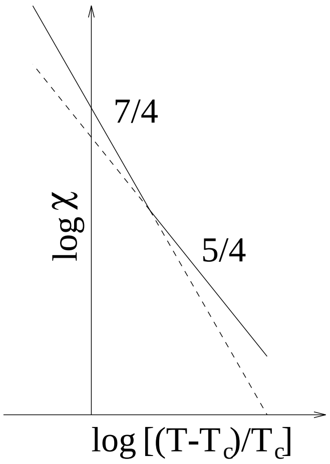


FIG. 23. Dimensional crossover. As the correlation length grows, the spins within the sample get more and more correlated, and the effective dimensionality of a non-isotropic sample is reduced. This manifests itself as kinks in the susceptibility *vs.* temperature plot.

4. Mermin-Wagner Theorem

In the discussion of the role of fluctuations in the Landau theory we concluded that fluctuations were less important in higher dimensions. In particular, we concluded that near the critical point fluctuations were unimportant if the dimensionality d exceeds the upper critical dimension d_c. The upper critical dimension depends on the model we are studying, but for the simple Landau theory we found $d_c = 4$. This discovery means that the mean field theory treatment of critical phenomena is exact if d is larger than d_c. Since the interesting theory has $d_c = 4$, and we live in a three-dimensional world, the consequence is that we cannot usually trust mean field theory but must find a way to include fluctuations. Attempting to do precisely that has kept physicists busy for the last 70 years.

A particular kind of fluctuation is due to Goldstone modes, the modes that are due to a broken continuous symmetry and have very small energy for small wavevector $\vec{q}$. If we calculate for instance the correlation function between two spins one of which is at the origin and the other at position $\vec{r}$ (i.e. $C(\vec{r}) = \langle \vec{s}(0) \cdot \vec{s}(\vec{r}) \rangle$), it is often easier to first calculate the Fourier transform of the correlation function, and then transform back from the wavevector space to the real space. The correlation function in the wavevector space is greatly affected by Goldstone modes with small wavevectors — they are easy to excite since their energy is so low. When we transform back to real space to get $C(\vec{r})$, we have to do an integral over wavevectors, which is of the type $\int_0^\infty dq\, q^{d-1}$. This means that contribution from small wavevectors becomes larger if we decrease the dimensionality d, and consequently the Goldstone modes are more important in lower dimensions. It turns out that below a certain dimension d_L the Goldstone modes are so important that there can be no long range order at any nonzero temperature; in other words, $\lim_{r \to \infty} C(r) = 0$ regardless of T if $d \leq d_L$. This dimensionality d_L is called the lower critical dimension. Physically, the lack of long range order means that if we know the direction of a spin at the origin, it gives us no information about the direction of a spin far away. An example of a state with long range order is the ferromagnetic ground state: if the spin at the origin points to direction $\hat{s}$, all other spins point to that direction, too, and $C(r) = 1$ for all r. A paramagnetic state, in contrast, has only short range order and $\lim_{r \to \infty} C(r) = 0$.

A careful analysis along the lines that I sketched above leads to the famous Mermin-Wagner theorem (also known as Mermin-Wagner-Coleman and Mermin-Wagner-Berezinskii theorem), which states that *models with discrete symmetry have no long range order in 1D, and models with continuous symmetry have no long range order in 1D or 2D.* Ising model is a model with discrete symmetry: at zero external field the Hamiltonian is invariant under the discrete transformation $s_i \to -s_i$, so consequently we know that the one-dimensional Ising model has no long range order for $T > 0$, which is in agreement with what we have found before. At two dimensions the Ising model does have an ordered phase, which is again in agreement with the Mermin-Wagner theorem. The Heisenberg model, in contrast, has a continuous symmetry, and therefore it cannot have an ordered phase in two dimensions.

5. *Universality Hypothesis*

In the preceding sections and in the homework sets we have studied a few different models and evaluated their critical exponents. We have found that sometimes quite different models can have same critical exponents. On the other hand, some rather similar-looking models (e.g. Ising model in 1D and 2D) have quite different critical properties. What do the critical properties depend on?

Let us start by collecting some evidence. We know that for instance that the one-dimensional Ising model does not have a phase transition, whereas the two-dimensional Ising model has one. Also, we know that mean field theory gives correct critical exponents for $d > 4$ but not for $d < 4$. Therefore, we can conclude that the dimensionality of space, d, is important. We have some more evidence to consider. We know that if the ordered state of the model has continuous symmetry, the system is guaranteed to have low-energy excitations (Goldstone modes), which decrease the tendency towards an ordered state — this was formulated more precisely in the Mermin-Wagner theorem. Thus, it appears that the symmetry properties of the model are important.

The universality hypothesis states that these are the *only* quantities that determine the critical properties of a model. More precisely: *"For systems with short range interactions, the critical properties* (e.g. *existence of phase transitions, critical exponents) are fully determined by the dimensionality of space (d) and the symmetry of the Hamiltonian."* This is known as the *universality hypothesis.* Note that it is a hypothesis — there is no proof — but so far there are no counterexamples (as far as I know). The qualifier "short range interactions" is necessary to take care of some pathological properties of systems with Coulomb interactions, which do not necessarily have Goldstone modes for $q > 0$ but only for $q = 0$.

The universality hypothesis amounts to a great simplification in the theory of critical phenomena. It states that to determine the critical properties of a model that we have never seen, we only need to analyze its symmetry properties (counting the spatial dimensions is trivial). For instance, we know that the liquid-gas transition is similar to an Ising magnet, and therefore we know that compressibility of a two-dimensional liquid diverges with exponent $\frac{7}{4}$ at the critical point (compressibility in the liquid-gas system corresponds to susceptibility in the spin model, and the number $\frac{7}{4}$ comes from the Onsager solution). Therefore, we say that boiling in two dimensions belongs to the *universality class* of the two-dimensional Ising model. By universality class we mean a class of models that have similar critical properties. If we can determine that two models belong to the same universality class, i.e. that they have similar symmetries and the dimensionality is the same, we can study the one that is more convenient. This is the basic reason why we are all the time dealing with spin models — they are versatile enough to cover many universality classes and it is more convenient to have one just one terminology.

6. Topological Excitations

Now we will, at least for a while, leave the relatively simple world of the spin-$\frac{1}{2}$ Ising model, and consider a more general continuous model.[19] A continuous model, where the direction of the spin may change smoothly, is in some respects qualitatively different from an Ising model, where each spin is ±1 and therefore a smooth variation is impossible. Intuitively we can understand that the discreteness of the Ising model adds some rigidity to the system, and may therefore reduce the role of fluctuations and increase the tendency towards an ordered state. This is indeed what the Mermin-Wagner theorem tells us: a discrete model (like the Ising model) may have long-range order already in two dimensions, whereas a continuous model (like Heisenberg model) can have long-range order only in dimensions larger than two.

There is, however, one type of rigidity that occurs in continuous models. In a continuous model for a ferromagnet the low-energy configurations are such that locally spins point in a common direction, but this common direction changes slowly from one part of the system to another. A relevant question is how these low-energy configurations are connected to each other, that is, is there a low-energy path connecting two low-energy configurations? If such a path does exist, then we expect the state of the system to fluctuate between various low-energy configurations, but if there is no way to get from one low-energy state to another without passing through some intermediate states with high energies, we expect that fluctuations are reduced.

How do we determine if such a low-energy path exists? Let us be more specific and consider an $O(n)$-model $H = H[\vec{S}(\vec{r})]$ where the spin $\vec{S}(\vec{r})$ is an n-dimensional vector and the Hamiltonian is invariant under a global rotation in the spin space. We consider the case where the space variable $\vec{r}$ is in the usual d-dimensional Euclidean space. A particular configuration of the system is specified by the direction of the spin everywhere in space, that is, by $\vec{S}(\vec{r})$. Some of these configurations have low energies, and are therefore important at low temperatures. Let us denote two such low-energy configurations $\vec{S}_0(\vec{r})$ and $\vec{S}_1(\vec{r})$. These configurations are connected by a low-energy path if we can smoothly deform the configuration $\vec{S}_0(\vec{r})$ in such a way that it coincides with $\vec{S}_1(\vec{r})$ — here “low-energy” implies that the

---

[19] This section is intended for those of you who are more mathematically inclined. It is not required for understanding the following sections. Those of you who yearn for more may turn to the entertaining article by David Mermin, reference [11].

deformation must be smooth. We can formulate this task more precisely in the following way: *Does there exist a continuous one-parameter mapping* $\vec{S}(\alpha;\vec{r})$ *such that* $\vec{S}(0;\vec{r}) = \vec{S}_0(\vec{r})$ *and* $\vec{S}(1;\vec{r}) = \vec{S}_1(\vec{r})$*?* Note that the mapping must be continuous both in α and in $\vec{r}$. The study of this question falls into the field of topology. If the answer to our question is yes, *i.e.* if there exist a low-energy path connecting $\vec{S}_0(\vec{r})$ and $\vec{S}_1(\vec{r})$, then the configurations (or mappings) $\vec{S}_0(\vec{r})$ and $\vec{S}_1(\vec{r})$ are said to belong to the same homotopy class (be homotopic), otherwise they are said to belong to different homotopy classes. How many homotopy classes there are depends on the parameters d and n, or more generally on the physical space (the space to which $\vec{r}$ belongs) and on the order parameter space (the space to which $\vec{S}$ belongs).

This may sound very formal and need some clarification. Let us consider an $O(3)$ model, i.e. a Heisenberg model, where the spin at any point is a unit vector in a three dimensional (spin) space. The physical space can be taken to be the usual three-dimensional space, $\vec{r} \in \mathbb{R}^3$, and since the magnitude of the spin vector is fixed to be 1, the order parameter space is the two-dimensional surface of a three-dimensional unit sphere, $\vec{S}(\vec{r}) \in S_2$. The ground state is one where all spins point in the same direction, $\vec{S}(\vec{r}) = \hat{S}_0$. Some low-lying excitations are spin waves of the form $\vec{S}_{\vec{k}}(\vec{r}) = \sqrt{1-(\delta\vec{S})^2}\hat{S}_0 + \delta\vec{S}\cos(\vec{k}\cdot(\vec{r}-\vec{v}t))$ where $\hat{S}_0 \cdot \delta\vec{S} = 0$. The energy of these excitations is proportional to $(\delta\vec{S})^2 k^2$. Since we can reach a spin wave configuration from the ground state by continuously increasing $|\delta\vec{S}|$ from zero, and since the mapping is continuous everywhere in space, the spin wave configuration is homotopic to the ground state, and therefore connected to it by a low-energy path.

Another configuration that can be assumed to have a low energy is one that looks like a hedgehog ("*igelkott*"), all spins point away from a fixed point which we can call the origin, in other words $\vec{S}(\vec{r}) = \hat{r}$. This is a spin configuration which is smooth (neighboring spins point to almost the same direction) except near the origin; thus, its energy is larger than the ground state energy, but not necessarily very much larger since the energy cost is concentrated near one point. This region, where the spin configuration is not differentiable, is called the core.[20] The hedgehog configuration cannot be reached from the ground state following a low-energy path — if we attempt to construct such a mapping, for instance $\vec{S}(\alpha;\vec{r}) = (1-\alpha)\hat{S}_0 + \hat{r}\left(\sqrt{1-(1-\alpha)^2\sin(\hat{r}\cdot\hat{S}_0)} - (1-\alpha)\cos(\hat{r}\cdot\hat{S}_0)\right)$, the mapping is not continuous at the origin. Therefore, this hedgehog configuration is an example of a topological excitation — a low-energy excitation which is not homotopic to the ground state.

Physically, the continuous mapping is similar to "combing" the spins. It is impossible to comb the quills of an entire hedgehog in such a way that adjacent quills always point in (nearly) the same direction, and so that at the end of combing process all the quills of the hedgehog point in the same direction. However, any region that does not surround the core can be combed in this way. This is an interesting statement: we do not actually have to study the vicinity of the core (the inside of the hedgehog) to realize that there is

---

[20] The excitation is called a point defect if the core is a point, *i.e.* if the core has dimension zero, and a line defect, if the core is a line.

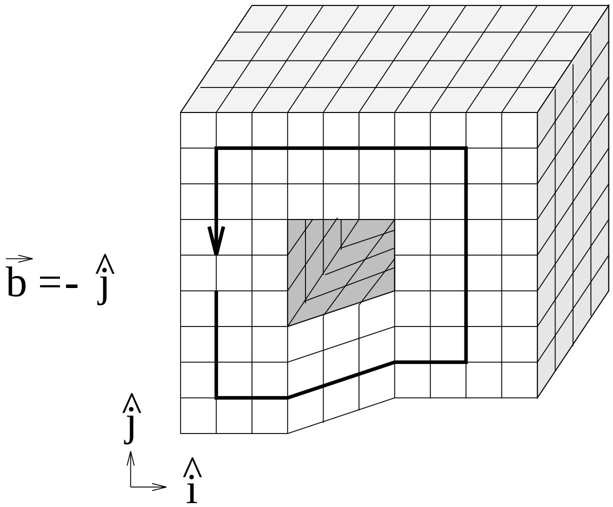


FIG. 24. A line dislocation in a crystalline solid and a corresponding topological invariant, the Burgers vector $\vec{b}$. The path indicated by the thick line would form a closed loop in an ideal solid but due to an edge dislocation it fails to close. The Burgers vector is the vector connecting the end and the beginning of the path.

something strange going on there, it is sufficient to study spin configurations on any closed surface that surrounds the origin. The topological structure of the configuration can be characterized by a set of so-called topological invariants, which can be determined from the far-field configuration. The forms of the invariants differ from system to system; the ones most commonly encountered in condensed matter physics are Burgers vector (in the study of dislocations in crystals), Chern number (in the theory of the fractional quantum Hall effect), and the Pontryagin index (in the study of spin models). The Burgers vector is depicted in Fig. 24. In our example of the Heisenberg model the topological invariant (the Pontryagin index) is given by

$$Q[\vec{S}(\vec{r})] = \frac{1}{4\pi} \oint_{\Gamma} d^3 r \epsilon_{\mu\nu} \vec{S}(\vec{r}) \cdot \partial_\mu \vec{S}(\vec{r}) \times \partial_\nu \vec{S}(\vec{r}) \tag{3.118}$$

where Γ is the boundary of the system, μ and ν are the Cartesian components x, y, z and $\epsilon_{\mu\nu}$ is an antisymmetric tensor, i.e. $\epsilon_{xy} = \epsilon_{yz} = \epsilon_{zx} = -\epsilon_{yx} = -\epsilon_{zy} = -\epsilon_{xz} = 1$, $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz} = 0$. For any spin configuration that is smooth (differentiable) on the surface Γ the invariant $Q[\vec{S}(\vec{r})]$ is an integer. If the configurations $\vec{S}_0(\vec{r})$ and $\vec{S}_1(\vec{r})$ are in the same homotopy class, their topological invariants are equal, $Q[\vec{S}_0(\vec{r})] = Q[\vec{S}_1(\vec{r})]$. For the ground state we have $Q = 0$ and for the hedgehog state $Q = 1$. The hedgehog configuration, and configurations homotopic to it, are more technically called skyrmions (after T. H. R. Skyrme, a British nuclear physicist).

In the next section we will study one type of topological excitation, a vortex, in more detail.

Let us now consider a special phase transition, known as the Kosterlitz-Thouless (or Kosterlitz-Thouless-Berezinskii) phase transition. It is a phase transition in the two-dimensional XY-model. The existence of a finite temperature phase transition in the two-dimensional XY-model is quite surprising since the Mermin-Wagner theorem states that in a model with continuous symmetry (like the XY-model) there is no long range order in two dimensions. Hence, the phase transition cannot be a transition between an ordered and disordered state, at least not in the conventional sense of the word.

Before we start our mathematical analysis, let us pause to think which systems could be described by the two-dimensional XY-model, and could therefore be candidates for observing the Kosterlitz-Thouless transition. Obviously we need a two-dimensional system, and a system where the order parameter is a two-dimensional vector. Since two-dimensional vectors can be represented as complex numbers, the order parameter may also be thought of as a complex number. There are at least three different systems that satisfy these requirements:

1. Systems consisting of weakly coupled superconductors — the superconducting order parameter is the Cooper pair amplitude $\psi = |\psi|e^{i\phi}$, which is a complex number. The magnitude $|\psi|$ is determined by external constraints like temperature and magnetic field, but the phase angle ϕ varies from one piece of superconducting material to another. The interaction energy between two superconductors depends on their phase difference. These types of systems are studied experimentally at Chalmers [Chen Chiidong, Ph. D. thesis, Chalmers University of Technology, School of Electrical and Computer Engineering, 1994].

2. Films of superfluid He^4. Below 2.172K He^4 is a superfluid with quite unusual properties, for instance the viscosity vanishes identically. The density and the velocity of the superfluid are obtained from superfluid amplitude Ψ which is a complex number, and is an order parameter for the superfluid transition. The superfluid transition in He^4 films was the first system in which the Kosterlitz-Thouless transition was studied, and the theory was found to be in a very good agreement with the experiments [D. J. Bishop and J. D. Reppy, *Phys. Rev. Lett.* **40**, 1727 (1978)].

3. Two-dimensional crystals. In two-dimensional crystals the melting occurs through a spontaneous generation of dislocations. The dislocations can be characterized by their two-dimensional Burgers vectors (see Fig. 24). The interaction energy between two dislocations is proportional to the scalar product of their Burgers vectors, which results in a Hamiltonian similar to the XY-model [B. I. Halperin and D. R. Nelson, *Phys. Rev. Lett.* **41**, 121 (1978)]. Experimentally two-dimensional melting has been studied in systems of layered liquid crystals, in electron systems on the surface of liquid helium, in systems of gases adsorbed on graphite, and several other systems.

Mathematically we are interested in the phase transitions associated with the Hamiltonian

$$H_{XY} = -J \sum_{\langle i,j \rangle} (S_i^{(x)} S_j^{(x)} + S_i^{(y)} S_j^{(y)}) = -J \sum_{\langle i,j \rangle} \cos(\theta_i - \theta_j) \tag{3.119}$$

where $|\vec{S}_i| = 1$ and θ_i is the polar angle of the spin at site i. First of all, the Hamiltonian H_{XY} has a continuous symmetry: changing $\theta_i \to \theta_i + \phi$ for all i leaves the Hamiltonian invariant. Secondly, the ground state is given by $\theta_i = \theta$, i.e. by all spins pointing in the same direction.

To analyze the low-energy properties of this system, we make the approximation that the direction of the spins varies slowly from one site to another, and therefore $\cos(\theta_i - \theta_j) \approx 1 - \frac{1}{2}(\theta_i - \theta_j)^2$. With this approximation the Hamiltonian becomes

$$H_{XY} \approx E_0 + \frac{1}{4} J \sum_{\vec{r},\vec{a}} [\theta(\vec{r}+\vec{a}) - \theta(\vec{r})]^2 \tag{3.120}$$

where $\vec{r}$ runs over all lattice sites and $\vec{a}$ over a set of vectors connecting adjacent sites. Note that this implies that we count each pair of nearest neighbors twice, which we have compensated by multiplying the sum by a half. The ground state energy E_0 is given by $E_0 = -\frac{1}{2}qNJ$ where q is the number of nearest neighbors. To find the low-lying excitations we Fourier transform this by setting

$$\theta(\vec{r}) = \frac{1}{\sqrt{N}} \sum_{\vec{k}} \theta_{\vec{k}} e^{i\vec{k}\cdot\vec{r}} \tag{3.121}$$

and obtain

$$H_{XY} \approx E_0 + \frac{1}{2} J \sum_{\vec{k}} \sum_{\vec{a}} \left[1 - \cos(\vec{k}\cdot\vec{a})\right] \theta_{\vec{k}} \theta_{-\vec{k}}. \tag{3.122}$$

Here the sum $\sum_{\vec{k}}$ runs over the first Brillouin zone. Let us now consider the specific case of a square lattice so that $\vec{a} \in a\{\hat{x}, -\hat{x}, \hat{y}, -\hat{y}\}$ where a is a lattice constant. Approximating $\cos(\vec{k}\cdot\vec{a}) \approx 1 - \frac{1}{2}(\vec{k}\cdot\vec{a})^2$ we get

$$H_{XY} \approx E_0 + Ja^2 {\sum_{\vec{k}}}' k^2 \theta_{\vec{k}} \theta_{-\vec{k}} \tag{3.123}$$

where $\sum'$ now runs only over half of the Brillouin zone. We have now succeeded in diagonalizing the Hamiltonian, and the low-energy excitations we have found are spin waves with wave vector $\vec{k}$ and energy $Ja^2k^2|\theta_k|^2$.

To determine if there is long range order we must calculate the spin-spin correlation function

$$C(\vec{r}) = \left\langle \vec{S}(\vec{r}) \cdot \vec{S}(0) \right\rangle = \mathrm{Re}\left\langle e^{i[\theta(\vec{r}) - \theta(0)]} \right\rangle. \tag{3.124}$$

To calculate this average we prove the classical version of the cumulant theorem: *If* $H = \epsilon A^2$ *where* $-\infty < A < \infty$, *then* $\langle e^{\alpha A} \rangle = e^{\frac{1}{2}\alpha^2 \langle A^2 \rangle}$. Proof:

$$\begin{aligned} \langle e^{\alpha A} \rangle &= \frac{\int_{-\infty}^{\infty} dA\, e^{-\beta\epsilon A^2} e^{\alpha A}}{\int_{-\infty}^{\infty} dA\, e^{-\beta\epsilon A^2}} = \frac{\int_{-\infty}^{\infty} dA\, e^{-\beta\epsilon\left[A - \frac{\alpha}{2\beta\epsilon}\right]^2 + \frac{\alpha^2}{4\beta\epsilon}}}{\int_{-\infty}^{\infty} dA\, e^{-\beta\epsilon A^2}} = e^{\frac{\alpha^2}{4\beta\epsilon}} \\ \langle A^2 \rangle &= \frac{\int_{-\infty}^{\infty} dA A^2 e^{-\beta\epsilon A^2}}{\int_{-\infty}^{\infty} dA e^{-\beta\epsilon A^2}} = \frac{1}{2\beta\epsilon}, \end{aligned}$$

Q.E.D.. Since our approximate Hamiltonian is quadratic in $\theta(\vec{r})$, we can apply the cumulant theorem and find

$$\begin{aligned} C(\vec{r}) &= \mathrm{Re}\,\left\langle e^{i[\theta(\vec{r})-\theta(0)]}\right\rangle \\ &= \mathrm{Re}\,e^{-\frac{1}{2}\langle[\theta(\vec{r})-\theta(0)]^2\rangle} \\ &= \mathrm{Re}\,e^{-\frac{1}{2N}\sum_{\vec{k},\vec{k}'}(e^{i\vec{k}\cdot\vec{r}}-1)(e^{i\vec{k}'\cdot\vec{r}}-1)\langle\theta_{\vec{k}}\theta_{\vec{k}'}\rangle} \end{aligned} \tag{3.125}$$

Using the equipartition theorem (or by straightforward integration) we have $Ja^2k^2\langle\theta_{\vec{k}}\theta_{\vec{k}'}\rangle = \delta_{\vec{k},-\vec{k}'}k_BT$ and therefore

$$C(\vec{r}) = e^{-\frac{k_BT}{NJa^2}\sum_{\vec{k}}\frac{1-\cos(\vec{k}\cdot\vec{r})}{k^2}}. \tag{3.126}$$

The remaining difficulty is how to calculate the sum over $\vec{k}$. For a square lattice the Brillouin zone is also a square, and is given by $-\pi/a < k_x, k_y < \pi/a$. We approximate this by a circle of radius π/a, and get

$$\frac{1}{N}\sum_{\vec{k}}\frac{1-\cos(\vec{k}\cdot\vec{r})}{k^2} \approx \frac{a^2}{(2\pi)^2}\int_0^{2\pi}d\phi\int_0^{r\pi/a}dk\frac{1-\cos(k\cos(\phi))}{k} \tag{3.127}$$

where I changed variables in the integral over k. This double integral is rather difficult and we have to contend ourselves with a reasonable estimate. In the integral over k the cosine term acts to remove the singularity at $k=0$, but apart from that it does not seem to be particularly important — the numerator varies between 0 and 2, and is in the average 1 except for small k when it is zero. Thus, let us replace the numerator by 1 and ignore the small-k singularity. This gives

$$\frac{1}{N}\sum_{\vec{k}}\frac{1-\cos(\vec{k}\cdot\vec{r})}{k^2} \approx \frac{a^2}{(2\pi)^2}\int_0^{2\pi}d\phi\log\left(\frac{\pi r}{a}\right) = \frac{a^2}{2\pi}\log\left(\frac{\pi r}{a}\right). \tag{3.128}$$

Substituting this into the expression for the spin-spin correlation function we obtain

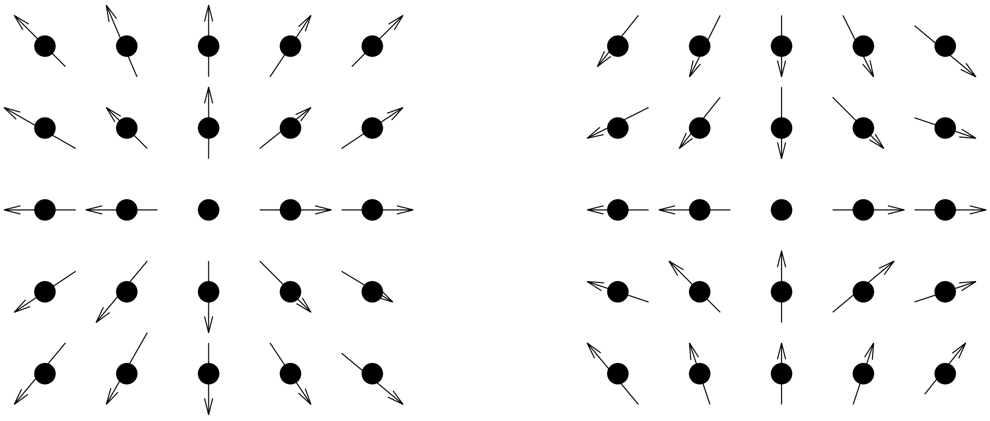
$$C(\vec{r}) = e^{-\frac{k_BT}{2\pi J}\log\left(\frac{\pi r}{a}\right)} = \left(\frac{\pi r}{a}\right)^{-\frac{k_BT}{2\pi J}}. \tag{3.129}$$

Thus, the spin-spin correlation function appears to be algebraically decaying at all temperatures. The usual form of a correlation function is

$$C(r) = \frac{e^{-r/\xi(T)}}{r^{d-2+\eta}} \tag{3.130}$$

so our result implies first of all that the exponent η is temperature dependent, $\eta(T) = \frac{k_BT}{2\pi J}$, but more importantly, the correlation length $\xi(T)$ is infinite at all temperatures! The second property suggests that the system is critical at all temperatures, which would be very strange indeed. Therefore, we must take a step back and consider the validity of our approximations.

The crucial approximation was that we only included spin waves. What if there are other low-energy excitations? They might destroy the algebraic order and give a usual,



n = 1

n = -1

FIG. 25. Two vortices: (a) winding number $n = 1$, (b) winding number $n = -1$. Note that even far from the vortex, the spin orientation is different from the ground state.

finite correlation length. The configurations that we have neglected are called vortices. A vortex in the origin is defined as a spin configuration which satisfies

$$\oint_{\Gamma} d\vec{l} \cdot \vec{\nabla}\theta(\vec{r}) = 2\pi n \tag{3.131}$$

where n is an integer different from zero and Γ is a closed contour that encircles the origin. The integer n is known either as a winding number, or vorticity, and it is a topological invariant — smooth deformations of the spin configuration do not change n. Thus, a vortex is a configuration in which the spin performs n complete revolutions along the path Γ and we can say that the spoin angles at the (arbitrary) starting point $\vec{r}^{+}$ and finishing point $\vec{r}^{-}$ on the path differ by $2\pi n$. While this leaves the exact energy $E \sim \cos(\theta(\vec{r}^{+}) - \theta(\vec{r}^{-}))$ invariant, it changes our approximation $E \sim [1 - \frac{1}{2}(\theta(\vec{r}^{+}) - \theta(\vec{r}^{-}))^2]$ by a large amount, which is why our approximation missed vortex excitations. One particular configuration that satisfies the condition (3.131) is given by $\theta(\vec{r}) = n\phi + \phi_0$ where ϕ is the polar angle of $\vec{r}$. Two vortex configurations are shown in Fig. 25.

To calculate the energy of a vortex we start from the continuum approximation for the Hamiltonian and evaluate

$$E = \frac{1}{2}J \int d^2r |\vec{\nabla}\theta(\vec{r})|^2 = \pi J n^2 \int_a^R \frac{dr}{r} = \pi J n^2 \log\left(\frac{R}{a}\right) \tag{3.132}$$

where a is the core radius of the vortex and R is the radius of the system (Note that in continuum $\vec{\nabla}\theta(\vec{r}) = \frac{n}{r}\hat{\phi}$ diverges at $r = 0$, and therefore a vortex must have a core that is not described by the simple form $\theta(\vec{r}) = n\phi + \phi_0$.). Thus, the vortex energy diverges logarithmically as the system size is increased. Therefore, at low temperatures there are no isolated vortices. What about at higher temperatures? To answer that question we must

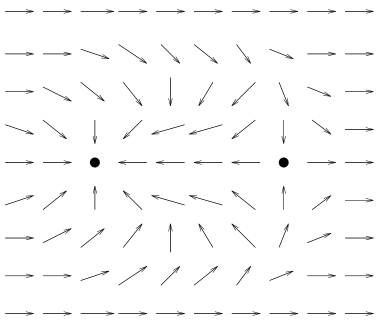


FIG. 26. A vortex-antivortex pair. Note that far from the pair the spin orientation is similar to that in the ground state.

calculate the free energy cost ΔF to create a vortex. Since $\Delta F = \Delta E - T\Delta S$, we must find the entropy change due to a vortex creation. If the vortex has a core size a, there are approximately $(R/a)^2$ places where we can create a vortex, and therefore the entropy due to a vortex creation is $\Delta S = k_B \log\left[(R/a)^2\right] = 2k_B \log(R/a)$ and the free energy cost is

$$\Delta F = \pi J n^2 \log\left(\frac{R}{a}\right) - 2k_B T \log\left(\frac{R}{a}\right) = \left[\pi J n^2 - 2k_B T\right] \log\left(\frac{R}{a}\right). \tag{3.133}$$

This is negative if $T > n^2 T_c = n^2 \frac{\pi J}{2k_B}$. Thus, if $T > T_c = \frac{\pi J}{2k_B}$ isolated vortices are spontaneously created.

Below T_c vortices ($n = 1$) and antivortices ($n = -1$) appear in pairs, as we can see by approximating the energy cost of having a vortex at position $\vec{r}_1$ and an antivortex at position $\vec{r}_2$. The gradient of the spin direction is then given by

$$\vec{\nabla}\theta(\vec{r}) = \left[\frac{1}{|\vec{r} - \vec{r}_1|} - \frac{1}{|\vec{r} - \vec{r}_2|}\right]\hat{\phi} \approx -\frac{1}{r^2}(\vec{r}_1 - \vec{r}_2)\cdot\hat{r}\hat{\phi} \tag{3.134}$$

so that the energy of a vortex-antivortex pair is

$$E \approx J \int_0^{2\pi} d\phi \int_a^R dr r^{-3}[(\vec{r}_1 - \vec{r}_2)\cdot\hat{r}]^2 \tag{3.135}$$

which stays finite as $R \to \infty$. The entropy is still logarithmically divergent, so at all temperatures $T > 0$ vortex-antivortex pairs are spontaneously created. A vortex-antivortex pair is depicted in Fig. 26 — note that the far-field configuration looks like a ground state.

Now we have a physical interpretation of the critical temperature T_c: it is the vortex unbinding temperature. For $T < T_c$ vortices and antivortices come in pairs so that the total vorticity of a pair vanishes, but for $T > T_c$ isolated vortices can occur. Mathematically the problem of vortices in the XY-model is similar to problem of interacting charges in a two-dimensional world, or the problem of magnetic interactions in a system of current-carrying wires. In the two-dimensional analog the vortex unbinding transition corresponds to a transition between charge-neutral two-dimensional molecules and a plasma of free charges.

Calculating the spin-spin correlation function in the presence of vortices is beyond the scope of this course. The result of such a calculation is that above T_c the correlation function assumes its usual form with a correlation length given by

$$\xi(T) \sim e^{b/|T-T_c|^{1/2}} \tag{3.136}$$

where b is a non-universal constant. Thus, the correlation length diverges faster than any power law as we approach the critical point *i.e.* the critical exponent ν is $\nu = \infty$). Right at the critical point the correlation function decays with exponent $\eta = k_B T_c/(2\pi J) = \frac{1}{4}$.

F. Renormalization Group Techniques

Renormalization group technique was developed by Kenneth Wilson in the early 1970s as a formal implementation of the scaling ideas of Kadanoff. It has become the standard tool for analyzing the critical properties of a wide range of models, and the ideas have been applied to many other fields as well (e.g. turbulence). The basic idea behind renormalization group is that near criticality the system looks approximately the same at different length scales (as long as the length scale is less than the correlation length), and by looking at how the system's appearance changes as we change the length scale, we should get some information about the system's critical properties.

We will start our study of the renormalization group technique by considering a special case, decimation, which can often be carried out exactly for one-dimensional models. In higher dimensions exact decimation is usually not possible and some approximations are needed.

1. Decimation

a. One-Dimensional Ising Model. Let us once again consider the infinite one-dimensional Ising model

$$H = -J \sum_{i=-\infty}^{\infty} s_i s_{i+1} - h \sum_{i=-\infty}^{\infty} s_i. \tag{3.137}$$

The partition function is given by

$$Z(T,h) = \sum_{\{s_i\}_{i=-\infty}^{\infty}} \exp\left[K \sum_{i=-\infty}^{\infty} s_i s_{i+1} + \beta h \sum_{i=-\infty}^{\infty} s_i\right]. \tag{3.138}$$

This would be easy to evaluate if the different spins were not coupled — in that case we could do sum over each spin independently — but the term $s_i s_{i+1}$ couples two neighboring spins, which makes the evaluation of Z more complicated. However, we can divide all spins into two subsets in such a way that spins in one subset interact only with spins in the other subset[21]: the neighbors of an even site are odd, and the neighbors of an odd site are even.

---

[21] Lattices for which such a division is possible are called bipartite, and they are frequently much simpler to deal with than more general lattices. The difference is particularly important in the case

To take advantage of this separation into mutually non-interacting lattices we write the partition function as

$$Z(T,h) = \sum_{\{s_{2n-1}\}_{n=-\infty}^{\infty}} \sum_{\{s_{2l}\}_{l=-\infty}^{\infty}} \exp\left[K \sum_{k=-\infty}^{\infty} (s_{2k-1}s_{2k} + s_{2k}s_{2k+1}) + \beta h \sum_{k=-\infty}^{\infty} s_{2k} + \beta h \sum_{k=-\infty}^{\infty} s_{2k+1}\right], \tag{3.139}$$

where we separated the contributions from even and odd sites. Let us now consider the sum over the spin at a particular even site $2l$. There are only three terms in the exponent that depend on s_{2l}, so the sum over s_{2l} gives

$$\sum_{s_{2l}=\pm 1} \exp\left[K s_{2l}(s_{2l-1} + s_{2l+1}) + \beta h s_{2l}\right] = 2\cosh\left[K(s_{2l-1} + s_{2l+1}) + \beta h\right].$$

We can do the same for all other even terms, and find that the partition function can be written as

$$Z(T,h) = \sum_{\{s_{2n-1}\}_{n=-\infty}^{\infty}} \prod_{n=-\infty}^{\infty} 2\cosh\left[K(s_{2n-1} + s_{2n+1}) + \beta h\right] e^{\frac{1}{2}\beta h(s_{2n-1}+s_{2n+1})}, \tag{3.140}$$

where I wrote the last term in a symmetric way for convenience.

Now we have expressed the partition function of the original system as a sum of only half of the degrees of freedom — only the odd spins. We can imagine that the partition function (3.140) could arise as a partition function for a system consisting of only the odd spins, but since it is not in the usual form $\mathrm{Tr}e^{-\beta H'}$, it is not immediately clear what Hamiltonian H' of the odd-spin-system would result in the partition function (3.140). Let us proceed by making the *Ansatz* that the Hamiltonian H' is also of the Ising type, i.e.

$$H' = -J' \sum_{n=-\infty}^{\infty} s_{2n-1}s_{2n+1} - \sum_{n=-\infty}^{\infty} (h' s_{2n-1} + g'). \tag{3.141}$$

Here J' describes the interaction between the odd spins, h' describes their coupling to the external magnetic field, and g' is included to allow for a shift in the zero of energy. The parameter g' does not affect the behavior of the model.

It is by no means clear that this *Ansatz* is correct, that is, that we can find constants J', h', and g' so that the Hamiltonian H' gives rise to the partition function (3.140). To proceed we write down the partition function for the Hamiltonian H',

$$Z'(T,h') = \sum_{\{s_{2n-1}\}_{n=-\infty}^{\infty}} \exp\left[K' \sum_{n=-\infty}^{\infty} s_{2n-1}s_{2n+1} + \beta \sum_{n=-\infty}^{\infty} (\frac{1}{2}h'(s_{2n-1} + s_{2n+1}) + g')\right]. \tag{3.142}$$

---

of an antiferromagnet: a bipartite lattice has a vanishing entropy as $T \to 0$ but a non-bipartite lattice has a large number of degenerate ground states and therefore does not obey the third law of thermodynamics, i.e. the entropy does not vanish as temperature goes to zero.

s_{2l-1}	s_{2l+1}	LHS	RHS
1	1	$2e^{\beta h}\cosh[\beta h + 2K]$	$e^{K'+\beta h'+\beta g'}$
-1	1	$2\cosh[\beta h]$	$e^{-K'+\beta g'}$
1	-1	$2\cosh[\beta h]$	$e^{-K'+\beta g'}$
-1	-1	$2e^{-\beta h}\cosh[\beta h - 2K]$	$e^{K'-\beta h'+\beta g'}$

TABLE II. Decimation equations (3.143). LHS = “left hand side”, RHS = “right hand side”.

We isolate the terms in (3.140) and in Z' that depend on spins s_{2l-1} and s_{2l+1}, and set them equal to each other — that is necessary if the partition functions are to agree. That gives the equations

$$\begin{aligned} &2\cosh\left[K(s_{2l-1}+s_{2l+1})+\beta h\right] e^{\frac{1}{2}\beta h(s_{2l-1}+s_{2l+1})} \\ &= \exp\left[K' s_{2l-1} s_{2l-1} + \tfrac{1}{2}\beta h'(s_{2l-1}+s_{2l+1}) + \beta g'\right] \end{aligned} \tag{3.143}$$

where $s_{2l-1} = \pm 1$ and $s_{2l+1} = \pm 1$. Thus, we have four equations and only three unknowns K', h' and g', and the correctness of our *Ansatz* appears to be in doubt. It is convenient to write out all the equations, which I have done in Table II. We see that the second and third equations collapse into one, so that there are only three independent equations and therefore we may expect to find a solution. Dividing the first equation by the fourth yields

$$h' = h + \frac{1}{2\beta}\log\left[\frac{\cosh[2K+\beta h]}{\cosh[-2K+\beta h]}\right]; \tag{3.144}$$

dividing the product of the first and fourth equations by the product of the second and third equations gives

$$K' = \frac{1}{4}\log\left[\frac{\cosh[2K+\beta h]\cosh[-2K+\beta h]}{\cosh^2[\beta h]}\right]. \tag{3.145}$$

A few remarks are now in order. We have obtained the result that if we perform a partial sum over half of the degrees of freedom, i.e. the even spins, the resulting partition function describes an Ising-type interaction between the remaining degrees of freedom. The form of the Hamiltonian is exactly the same as before but the coupling constant has changed, $K \to K'$ as has the effective magnetic field, $h \to h'$. Nothing prevents us from doing another partial sum over half of the remaining degrees of freedom (say, sites $4l+1$). If we were to do that, we would find that the partition function after the partial summation would describe an Ising-interaction between the remaining degrees of freedom (s_{4l+3}). The coupling constant of the resulting model is K'' and the effective magnetic field is h'', which describe the behavior of the model at a length scale that is four times the original length scale (the separation between the remaining degrees of freedom has increased by a factor of four). Thus, we can coarse-grain the original Ising chain by successively removing every second spin, and the result at each stage is another Ising model with a new coupling constant

$K^{(n)}$ and a new effective magnetic field $h^{(n)}$. The equations (3.144) and (3.145) that describe how the parameters of the Hamiltonian change when we remove some degrees of freedom are called renormalization group equations.

Let us analyze the renormalization group equations a little more carefully. First of all, we notice that if the original magnetic field h vanishes, so does the renormalized field h': no magnetic field is spontaneously generated in the rescaling process. This has to be the case since a zero magnetic field implies that there is no preferred direction for the spins (either up or down); if there is no preferred direction in the small length scale description, there cannot be one in the large length scale description either. Secondly, if the external magnetic field vanishes, we have $K' = \frac{1}{2} \log \cosh[2K]$. If $K' < K$, the interaction between the coarse-grained spins is weaker than between the adjacent spins. This means that if two spins are very far apart, interaction between them is very weak, and their directions are not correlated. If on the other hand $K' > K$, the interactions between coarse-grained spins is stronger than between adjacent spins and their directions are strongly correlated. In the intermediate case $K' = K$ the interaction strength does not depend on the separation between spins. The equation $K' = \frac{1}{2} \log \cosh[2K] = K$ has two solutions, $K = 0$ and $K = \infty$ which correspond to the intermediate case when the effective coupling is independent of the distance between the spins. Since $K = \beta J$, this implies that the interaction strength is scale independent if $T = 0$ or $T = \infty$. In the first case interactions are strong, and the system is ordered, whereas in the second (infinite temperature) case interactions are very weak and the system is disordered. In this particular model we have $K' < K$ for all $0 < K < \infty$, so unless the temperature is exactly zero, the interactions get weaker with coarse-graining and the system is disordered at sufficiently large length scales. This confirms our earlier result that the one-dimensional Ising model is disordered at all nonzero temperatures.

It is useful to describe the renormalization group equations graphically as I have done in Fig. 27. The starting points of the arrows indicate the original parameters (K, h), and the end points indicate the renormalized parameters (K', h') after one decimation procedure, i.e. on a length scale that is twice the original length scale. We see that upon successive decimations, the parameters flow towards one of four different points, depending on the original parameters (K, h): if $h > 0$, the flow is towards $(K^* = 0, h^* = +\infty)$, if $h < 0$ the parameters flow towards $(K^* = 0, h^* = -\infty)$, and if $h = 0$, they flow towards $(K^* = 0, h^* = 0)$ unless $K = \infty$ in which case the parameters remain $(K = \infty, h^* = 0)$. The flow diagram tells us that even a small magnetic field changes the system's large scale behavior qualitatively — the magnetic field becomes more and more important upon successive decimations — and therefore the magnetic field is called a relevant variable. We can imagine that some other perturbations get weaker at longer length scales, in which case they are called irrelevant. The concepts of renormalization group flow and relevant and irrelevant variables will be discussed in more detail in later sections.

b. Two-Dimensional Ising Model. Let us now try to apply decimation to the two-dimensional Ising model. First we must choose a lattice. In the discussion of the one-dimensional Ising model we relied on the fact that the lattice was bipartite — that allowed us to carry out the sum over half of the spins — so let us also choose a bipartite two-dimensional lattice. The simplest bipartite lattices in two dimensions are the square lattice and the hexagonal lattice; the triangular lattice is not bipartite as you will quickly see if you

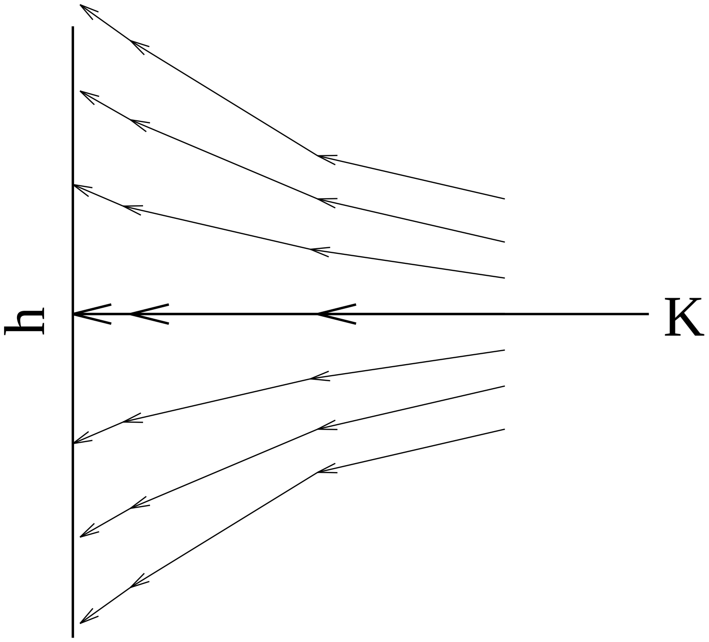


FIG. 27. Renormalization group flows for the one-dimensional Ising chain under decimation.

try to construct the antiferromagnetic ground state on it. Something interesting happens if you decimate the Ising model on the hexagonal lattice — as you will see in the home problems — but let us now focus on the square lattice.

The square lattice can be divided into two sublattices, call them A and B, so that all the neighbors of a spin in the A sublattice belong to the B sublattice and *vice versa*, see Fig. 28. Let us isolate the terms that depend on a particular spin on the A sublattice, s_{Aj}. Denoting the four neighbors of this spin by $s_{Bj,\text{up}}$, $s_{Bj,\text{down}}$, $s_{Bj,\text{left}}$ and $s_{Bj,\text{right}}$, we can do the sum over $s_{Aj} = \pm 1$ and obtain

$$\begin{aligned}&\textstyle\sum_{s_{Aj}=\pm 1} \exp\left[K s_{Aj}(s_{Bj,\text{up}} + s_{Bj,\text{down}} + s_{Bj,\text{left}} + s_{Bj,\text{right}}) + \beta h s_{Aj}\right] \\ &\quad = 2\cosh\left[K(s_{Bj,\text{up}} + s_{Bj,\text{down}} + s_{Bj,\text{left}} + s_{Bj,\text{right}}) + \beta h\right].\end{aligned} \tag{3.146}$$

We can repeat this for all sites in the A sublattice, and obtain the partition function

$$Z = \sum_{\{s_B\}} \prod_{\diamond} 2\cosh\left[K(s_{\diamond\uparrow} + s_{\diamond\rightarrow} + s_{\diamond\downarrow} + s_{\diamond\leftarrow}) + \beta h\right] e^{\frac{1}{4}\beta h(s_{\diamond\uparrow} + s_{\diamond\rightarrow} + s_{\diamond\downarrow} + s_{\diamond\leftarrow})} \tag{3.147}$$

where the product is over the diamond-shaped cells of the B sublattice (see Fig. 28), and $s_{\diamond\uparrow}$ etc. denote spins at the corners of the cells. Note that each spin belongs to four cells so that to avoid overcounting I have divided the magnetic field contribution by four in the last factor.

We will attempt to do the same trick that worked for the one-dimensional Ising model, that is, write the partition function (3.147) as a partition function of an Ising model for spins on the B sublattice. Denoting the new parameters by K', h' and g', we can write the Ising model partition function as

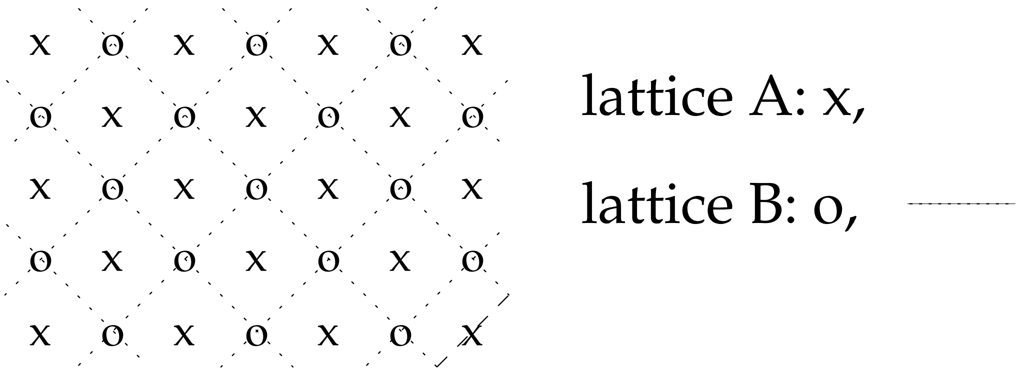


FIG. 28. Division of the square lattice to two sublattices A and B. Note that the sublattices are also square lattices but their unit cells are rotated by 45 degrees.

$$Z' = \sum_{\{s_B\}} \prod_{\diamond} \exp\left[K'(s_{\diamond\uparrow}s_{\diamond\rightarrow} + s_{\diamond\rightarrow}s_{\diamond\downarrow} + s_{\diamond\downarrow}s_{\diamond\leftarrow} + s_{\diamond\leftarrow}s_{\diamond\uparrow}) + \frac{1}{4}\beta h'(s_{\diamond\uparrow} + s_{\diamond\rightarrow} + s_{\diamond\downarrow} + s_{\diamond\leftarrow}) + \beta g'\right]. \tag{3.148}$$

To make decimation work, we must find K', h' and g' such that $Z' = Z$. However, consider the two configurations $\{s_{\diamond\uparrow}, s_{\diamond\rightarrow}, s_{\diamond\downarrow}, s_{\diamond\leftarrow}\} = \{1, 1, -1, -1\}$ and $\{s_{\diamond\uparrow}, s_{\diamond\rightarrow}, s_{\diamond\downarrow}, s_{\diamond\leftarrow}\} = \{1, -1, 1, -1\}$: they give the same contribution to Z but different contributions to Z' (unless $K' = 0$). Thus, we cannot find parameters K', h' and g' such that $Z' = Z$, and therefore the partition function after decimation does not describe a simple Ising model on the B sublattice.

We must conclude that exact decimation does not work for the two-dimensional Ising model on a square lattice. This we could have concluded already from the form (3.147): it couples spins $s_{\diamond\uparrow}$ and $s_{\diamond\downarrow}$ which are *not* nearest neighbors, and therefore contains interactions between next-nearest neighbors which are not included in the Ising Hamiltonian. This shows that the arguments that we used in the discussion of block spins to justify the scaling hypothesis are not strictly valid: the Hamiltonian may depend on the length scale at which we are interested. The systematic way to analyze these changes is known as the renormalization group, which will form the concluding topic of this course.

2. Renormalization Group

a. General properties. In the latter part of this course we have tried to analyze the critical properties of a number of models. We have seen that they can be characterized in terms of critical exponents, and consequently our task has been to attempt to calculate the exponents. Quite frankly, so far our success has been limited: apart from the mean field results — which we know to be unreliable in the physical three-dimensional world — we only know the exponents in a few special, exactly solvable cases. The problem has turned out to be quite complicated indeed.

In physics, however, there is a tradition of dealing with complications. The atomic spectrum appears to be quite complicated to the untrained eye but we know from quantum mechanics that it can be organized in terms of a few quantum numbers like the angular momentum. This organization is possible because of the rotational symmetry that allows us to classify all states (according to the irreducible representations of the symmetry group). Can we identify such a symmetry in the present problem, the analysis of critical phenomena? We have already done so: the Kadanoff block spin construction suggested that near the critical point the system is invariant under a scale transformation. Unfortunately, when we tried to decimate the two-dimensional Ising model we saw that this was not quite the case. The system is nevertheless approximately invariant under a scale transformation, and we can use that property to analyze its critical properties. The corresponding symmetry group is called the renormalization group (RG).

Let us consider a Hamiltonian $H(\{s_j\})$ which describes a system of spins $\{s_j\}$ on a d-dimensional lattice with lattice constant a. The Hamiltonian has some parameters like the external magnetic field, the coupling constant etc., which we will call $\{K_n\}$ — thus, we can have for instance

$$\mathcal{H} = -\beta H = -\beta(-J\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i) = K_1 \sum_{\langle i,j\rangle} s_i s_j + K_2 \sum_i s_i, \tag{3.149}$$

where we multiplied by $-\beta$ for convenience. The space $\{K_n\}$ is called the parameter space of the Hamiltonian. The parameters $\vec{K} = \{K_n\}$ determine all the properties of the system. In particular, they determine the probability that the spins $\{s_i\}$ assume particular values $\{\sigma_j\}$ — this is given by $P(\{s_j = \sigma_j\}) = \frac{1}{Z}\exp[-\beta H(\{\sigma_j\})]$. They also determine the correlation length $\xi = af(\vec{K})$ where $f(\vec{K})$ is some function of the parameters. Note that the correlation length is proportional to the lattice spacing a.

An RG transformation of the Hamiltonian consists of several steps. First we form block spins, in the way that was introduced by Kadanoff, by grouping Λ^d spins together to create block spins $\{s_j^{(b)}\}$. The distance between adjacent block spins is Λ times larger than the distance between the original spins so that the new lattice constant is $a' = \Lambda a$. Since the block spins are functions of the original spins, the probability distribution of the block spins must be related to the probability distribution of the original spins. Writing the probability distribution for the block spins as $P(\{s_j^{(b)} = \sigma'_j\} = \frac{1}{Z'}\exp[-\beta H'(\{\sigma'_j\})]$ we can define the block spin Hamiltonian $H'(\{s_j^{(b)}\})$. We assert that the block spin Hamiltonian H' is of the same form as the original Hamiltonian H, and only the parameters $\{K_n\}$ may change.[22] The parameters $\vec{K}' = \{K'_n\}$ of the block spin Hamiltonian are connected to the original parameters $\vec{K} = \{K_n\}$ through a transformation R_Λ,

---

[22] This is actually no restriction since the original Hamiltonian can include all kinds of terms, for instance we could write

$$\mathcal{H} = K_1 \sum_{\langle i,j\rangle} s_i s_j + K_2 \sum_i s_i + K_3 \sum_{i,j \text{ n.n.n}} s_i s_j + K_4 \sum_{\langle i,j,k,l\rangle} s_i s_j s_k s_l \tag{3.150}$$

where the last two terms run over pairs of next-nearest neighbors and and groups of four neighboring spins, respectively. If the last two terms are not present on the microscopic level we just set $K_3 = K_4 = 0$.

$$\vec{K}' = R_\Lambda \vec{K}. \tag{3.151}$$

The transformation R_Λ is called a renormalization group transformation. Note that R_Λ is a mapping from one vector in the parameter space to another — if R_Λ were a linear mapping it would be a matrix, but in general it is nonlinear; equation (3.151) is nothing but a shorthand notation for $K_j' = F_j(\Lambda; K_1, K_2, \dots, K_n)$. It is usually not possible to find R_Λ exactly, instead, we will have to resort to various approximate techniques some of which we will study in the subsequent sections. Furthermore, there is no unique way to rescale a particular problem so the transformations R_Λ are not uniquely determined by the microscopic Hamiltonian.

Equation (3.151) tells how the parameters describing the system depend on the length scale that we are considering. Each RG transformation corresponds to an additional level of coarse graining through the construction of block spins. The new, coarse-grained description ignores some short scale structure of the original microscopic description, but preserves the long wavelength properties of the model. If we successively increase the scale by a factor Λ, the parameters $\{K_n\}$ move along some trajectories in the parameter space. This motion is called the renormalization group flow.

If we perform two scale transformations R_{Λ_1} and R_{Λ_2} in succession, their combined effect is to change the scale by the factor $\Lambda = \Lambda_1 \Lambda_2$; thus, the transformations R_Λ form a semigroup satisfying

$$R_{\Lambda_1 \Lambda_2} = R_{\Lambda_1} R_{\Lambda_2}. \tag{3.152}$$

However, in general there is no way to reconstruct the small scale Hamiltonian if we know the system's large scale behavior and therefore the inverse transformation $(R_\Lambda)^{-1}$ does not exist — consequently, the operations R_Λ do not form a group and the term renormalization group is something of a misnomer.

b. Fixed points and renormalization group flow. There are usually a number of points in the parameter space that satisfy $R_\Lambda \vec{K}^* = \vec{K}^*$, *i.e.* points that are invariant under the renormalization group transformation R_Λ. These *fixed points* are particularly important. To see this, recall that the correlation length is given by $\xi = af(\vec{K})$ where a is the lattice constant and $f(\vec{K})$ is some function of the parameters $\{K_n\}$. Since the correlation function is a measurable quantity, it cannot depend on whether we describe the system in terms of the microscopic spins $\{s_i\}$ or block spins $\{s_j^{(b)}\}$. When we perform the RG transformation that takes us over from the microscopic description $\mathcal{H}$ to the coarse-grained description $\mathcal{H}'$, the parameters change from $\vec{K}$ to $R_\Lambda \vec{K}$, and the lattice constant changes from a to Λa. Thus, we have the result

$$\xi = af(\vec{K}) = \Lambda a f(R_\Lambda \vec{K}). \tag{3.153}$$

Now, if $\vec{K}$ is a fixed point $\vec{K}^*$ then $R_\Lambda \vec{K}^* = \vec{K}^*$ and we have $\xi^* = af(\vec{K}^*) = \Lambda a f(\vec{K}^*)$. This equation has only two solutions: either $\xi^* = 0$, or $\xi^* = \infty$. The first case corresponds to a completely uncorrelated phase whereas the second case is a critical point: the correlation length is divergent. Thus, all points $\vec{K}^*$ in the parameter space that are invariant under RG transformations correspond either to an uncorrelated phase, or to a critical point. The former are called trivial fixed points and the latter are called critical fixed points.

Each critical fixed point has its basin of attraction which consists of those points in the parameter space that flow towards the fixed point, that is, of points $\vec{K}$ such that $\lim_{\Lambda\to\infty} R_\Lambda \vec{K} = \vec{K}^*$. This basin of attraction is called the critical manifold. Let us consider a point $\vec{K}$ on the critical manifold, and perform successive RG transformations R_Λ so that $\vec{K}^{(n)} = (R_\Lambda)^n \vec{K}$. Since ξ is independent of n we have $\xi = af(\vec{K}) = \Lambda a f(\vec{K}^{(1)}) = \ldots = \Lambda^n a f(\vec{K}^{(n)}) = \ldots$. Since $\Lambda > 1$ and $f(\vec{K}^*) = \infty$ the right hand side diverges as $n \to \infty$, and therefore $\xi = \infty$. Thus, the correlation length is divergent for all parameter values $\vec{K}$ that lie in the basin of attraction of a critical fixed point, which justifies the term critical manifold.

Thus, we have determined that the RG analysis allows us, by determining the fixed points of the RG transformation, to determine the critical points and critical manifolds. However, we are mostly interested in the critical exponents since they are more universal, and thus far we do not know if RG can be used to calculate them. We introduced the RG concept since it appeared that there was a symmetry — scale invariance — that was at least approximately valid near the critical point, and we knew from experience that symmetries often lead to simplifications in physical problems [23]. In our earlier example, quantum mechanics, identifying the rotational symmetry allowed us to organize and understand atomic spectra in great detail. So far RG falls short of such great expectations. This suggests that we have not yet fully utilized the potential of the symmetry in the present problem. In particular, we have not found any "quantum numbers". Recalling that angular momentum quantum numbers appear in quantum mechanics as eigenvalues of rotation operators suggests that if we wish to find some counterpart of quantum numbers in the present problem, we should analyze the eigenvalue problem for the symmetry operations.

Since the symmetry in the present case is only approximate, and valid only near criticality when the correlation length is large, we proceed by analyzing the RG equation near the fixed points. Writing $\vec{K} = \vec{K}^* + \delta\vec{K}$ and applying the transformation R_Λ we get

$$\vec{K}' = \vec{K}^* + \delta\vec{K}' = R_\Lambda(\vec{K}^* + \delta\vec{K}) \approx \vec{K}^* + \left(\frac{\partial R_\Lambda}{\partial \vec{K}}\right)_{\vec{K}^*} \delta\vec{K} \tag{3.154}$$

where we linearized the RG transformation near the fixed point. In component notation we now have linearized RG equations

$$\delta K'_m = \sum_n \left(\frac{\partial K'_m}{\partial K_n}\right)_{K^*} \delta K_n \tag{3.155}$$

and we can define the corresponding eigenvalue problem. However, since we do not know anything about the matrix $\hat{M}^\Lambda_{mn} = \left(\frac{\partial K'_m}{\partial K_n}\right)_{K^*}$ (except that it is real), it is not clear that the matrix is diagonalizable or that its eigenvalues are real.

---

[23] Scale invariance implies that the equations describing critical phenomena are invariant under global scale transformations (stretching). It is believed nowadays that the invariance extends to *local* scale transformations as well (spatially varying stretching). This more general *conformal invariance* leads to great simplifications in particular for two-dimensional models.

For simplicity we assume that the matrix is diagonalizable. We can then write down the eigenvalue problem

$$\sum_n \hat{M}^{\Lambda}_{mn} e^n_\sigma = \lambda_{\Lambda,\sigma} e^m_\sigma \tag{3.156}$$

where e^n_σ is the n^{th} component of the σ^{th} eigenvector and $\lambda_{\Lambda,\sigma}$ is the corresponding eigenvalue. Since two successive RG transformations are equivalent to one RG transformation with a larger rescaling factor, $R_{\Lambda_1} R_{\Lambda_2} = R_{\Lambda_1 \Lambda_2}$, the matrices $\hat{M}$ satisfy $\hat{M}^{\Lambda_1} \hat{M}^{\Lambda_2} = \hat{M}^{\Lambda_1 \Lambda_2}$ and therefore

$$\lambda_{\Lambda_1,\sigma} \lambda_{\Lambda_2,\sigma} = \lambda_{\Lambda_1 \Lambda_2,\sigma}. \tag{3.157}$$

This implies $\lambda_{\Lambda,\sigma} = \Lambda^{y_\sigma}$ where y_σ is a number that depends on σ but not on Λ. The usefulness of the eigenvalue analysis becomes apparent if we write in (3.155) $\delta \vec{K} = \sum_\sigma a_\sigma \vec{e}_\sigma$ so that $\delta \vec{K}' = \sum_\sigma a_\sigma \lambda_{\Lambda,\sigma} \vec{e}_\sigma$. We see now that if $|\lambda_{\Lambda,\sigma}| < 1$ and $a_\sigma \neq 0$ the renormalized parameters K' are closer to the fixed point that the original parameters. Therefore the critical manifold is spanned by those eigenvectors whose eigenvalue has a modulus less than one.

We can now be more specific about the definition of relevant, irrelevant, and marginal directions in the parameter space. The direction $\vec{e}_\sigma$ is called

- relevant, if $|\lambda_{\Lambda,\sigma}| > 1$ or $y_\sigma > 0$
- irrelevant, if $|\lambda_{\Lambda,\sigma}| < 1$ or $y_\sigma < 0$
- marginal, if $|\lambda_{\Lambda,\sigma}| = 1$ or $y_\sigma = 0$

Thus, the critical manifold is spanned by the directions that are irrelevant near the critical fixed point (hence the terminology: it does not matter if the system is displaced from a critical point to an irrelevant direction, upon RG transformations it will flow towards the critical point and reach it in the infinitely large scale limit). Relevant directions, in contrast, are directions that take the system away from the critical fixed point. Marginal directions require more careful analysis and often lead to logarithmic corrections in various quantities.

In Figure 29 I have depicted the RG flows for a model with three fixed points: one corresponding to the high-temperature phase, one corresponding to the zero-temperature phase, and one critical fixed point. The zero-temperature point corresponds to $K_2 = \beta J_2 = \infty$ and is not shown in the figure. The high-temperature and zero-temperature fixed points have no relevant directions and are therefore called stable fixed points or sinks: all systems that have their parameters $\vec{K}$ in the vicinity of either of those two flow to the fixed points under RG transformations. Stable fixed points correspond to stable bulk phases — in the example in the Fig. 29 there are only two stable phases, the (presumably ordered) low-temperature phase, and the (presumably disordered) high-temperature phase. The critical fixed point has one relevant and one irrelevant direction. There must be at least one relevant direction since otherwise all systems in the vicinity of a critical fixed point would flow to it, and we know that we can always take the system away from criticality by changing temperature slightly from the critical temperature. Let us now assume, for simplicity, that there is one relevant and one irrelevant direction at the critical point, and let the relevant

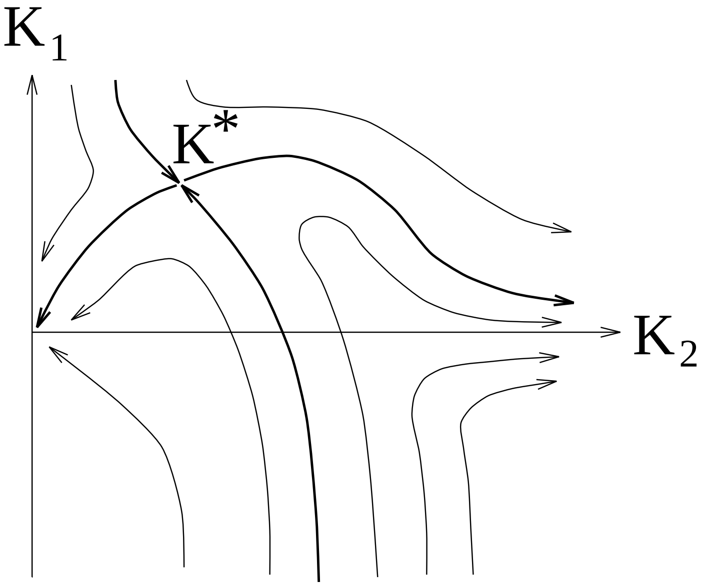


FIG. 29. Renormalization group flow for a model with three fixed points. Note that the critical fixed point K^* has one relevant and one irrelevant direction. The high-temperature fixed point in the origin has no relevant directions. The low-temperature fixed point is at $K_2 = \infty$ and does not appear in the picture.

direction be $\vec{e}_t$. Since the projection of $\vec{K}$ in the direction $\vec{e}_t$ tells how far the system is from criticality, it is natural to identify this projection as $t = (T - T_c)/T_c$ — since the length of the eigenvector $\vec{e}_t$ is arbitrary, we can adjust it so that the prefactor multiplying t is unity. From the renormalization group equation we know that rescaling by Λ changes t to $t' = t\Lambda^{y_t}$. Thus, we can conclude that the RG transformed Hamiltonian $\mathcal{H}'$ describes a system at a different temperature t'.

This is a major conceptual step: until now we have maintained that the RG transformation is a symmetry operation (rescaling of the lattice constant by Λ), not unlike rotation in quantum mechanics, and the measurable properties of the system are unaffected by the symmetry operation. Now, after the identification of one relevant direction as temperature, we can adopt another point of view: the RG transformation maps a system at temperature t to another system at temperature t' but with the original lattice constant a. This can be rephrased so that coarse graining the system is similar to taking it further from criticality. Viewing the RG transformation as a symmetry operation and writing $\xi = af(\vec{K})$, we concluded that $f(R_\Lambda \vec{K}) = \Lambda^{-1} f(\vec{K})$. Applying this information to the new interpretation we conclude that

$$\xi(t\Lambda^{y_t}) = \Lambda^{-1}\xi(t). \tag{3.158}$$

We have previously defined the critical exponent ν through $\xi \sim |t|^{-\nu}$, and inserting this to the previous expression yields

$$\nu = \frac{1}{y_t}. \tag{3.159}$$

This is the first connection between our RG analysis and the long sought-after critical exponents. Using scaling laws we obtain $2 - \alpha = d/y_t$. The other critical exponents are associated with the magnetic field which is a relevant operator as we saw in the analysis of the one-dimensional Ising model. To calculate them we would have to consider a system with two relevant directions near the fixed point.

Perhaps the most important consequence of the connection between critical exponents and the eigenvalues of the RG operations is that such a connection implies that the critical exponents are the same for all Hamiltonians that flow to the same fixed point. This is at the heart of universality: each universality class corresponds to a particular fixed point.

3. Position-Space Renormalization

In the previous discussion of the renormalization group there was one crucial step that we did not address in any detail: once we have $\mathcal{H} = -\beta H$, how do we get $\mathcal{H}' = -\beta H'$? If we can obtain $\mathcal{H}'$, it is (at least in principle) straightforward to find the fixed points, identify stable phases, linearize the transformations near critical fixed points, and obtain critical exponents. In this section we will discuss one approximate scheme to obtain the renormalized Hamiltonian for a set of lattice models. The method is called position-space renormalization. It is best illustrated with an example, and the easiest example to deal with is once again the two-dimensional Ising model. This time we will study the model on a triangular lattice.

The triangular lattice is shown in Figure 30. We start by constructing block spins $s_j^{(b)}$ from three spins s_{j1}, s_{j2} and s_{j3} that lie in the corners of a triangle j. The natural way to assign a value to the block spin is to use *majority rule*: if the majority of the spins s_{j1}, s_{j2} and s_{j3} are $+1$, the block spin $s_j^{(b)} = +1$, otherwise $s_j^{(b)} = -1$. The majority rule can be generalized to all position-space renormalization problems, although if the number of spins in the block is even we must come up with a special rule for the case when equally many spins in the block are $+1$ and -1. Note that the block spins also form a triangular lattice but the lattice constant has increased from a to $\sqrt{3}a$, hence, the RG parameter is $\Lambda = \sqrt{3}$ in this case.

The second step is to rewrite the original Hamiltonian

$$\mathcal{H} = -\beta H = K \sum_{\langle ij \rangle} s_i s_j + \beta h \sum_i s_i \tag{3.160}$$

as a sum of two terms $\mathcal{H}_0 + \mathcal{H}_1$ so that the term $\mathcal{H}_0$ does not couple different blocks. This gives

$$\mathcal{H}_0 = \sum_j [K(s_{j1}s_{j2} + s_{j2}s_{j3} + s_{j3}s_{j1}) + \beta h(s_{j1} + s_{j2} + s_{j3})] \tag{3.161}$$

$$\mathcal{H}_1 = \sum_j K(s_{j1}s_{j'2} + s_{j1}s_{j'3} + s_{j1}s_{j''2} + s_{j3}s_{j''2} + s_{j3}s_{j'''1} + s_{j3}s_{j'''2}) \tag{3.162}$$

where j' is the block right above block j and j'' is the block that is slightly above and to the right of block j, see Fig. 30. Since j runs over all blocks, including the interactions between

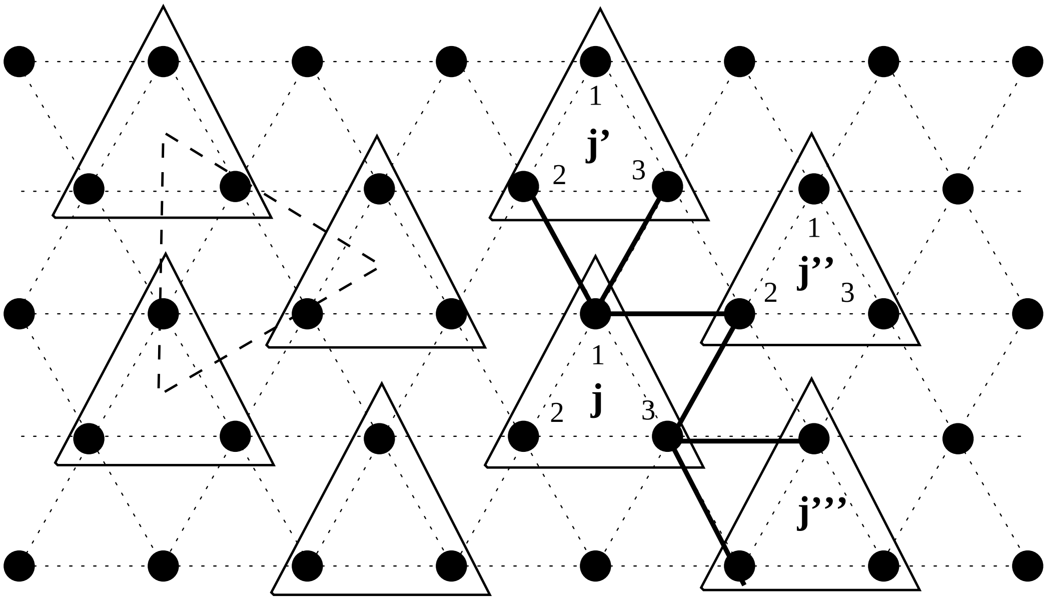


FIG. 30. Triangular lattice and the block spin construction. The blocks j, j', j'', and j''' have been named and interactions between them are indicated by solid lines.

the block j and its neighbors to the left or below would result in double counting in the expression for $\mathcal{H}_1$. The probability for a particular block spin configuration is given by the sum of the probabilities of those microscopic spin configurations that give rise to the right block spins. The coarse-grained Hamiltonian is thus given by

$$e^{\mathcal{H}'(\{s_j^{(b)}\})} = \sum_{\{s_i\}}{}' e^{\mathcal{H}_0(\{s_i\})+\mathcal{H}_1(\{s_i\})} \tag{3.163}$$

where the sum runs over configurations $\{s_i\}$ such that $s_j^{(b)} = \text{sign}(s_{j1} + s_{j2} + s_{j3})$ for all blocks j.

The difficulty in evaluating this expression comes from the term $e^{\mathcal{H}_1(\{s_i\})}$ which couples neighboring blocks. We will treat the difficult term perturbatively. To do that, we first define the average of the observable A with respect to $\mathcal{H}_0$ as

$$\langle A(\{s_j^{(b)}\})\rangle_0 = \frac{\sum'_{\{s_i\}} e^{\mathcal{H}_0(\{s_i\})} A(\{s_i\})}{\sum'_{\{s_i\}} e^{\mathcal{H}_0(\{s_i\})}}. \tag{3.164}$$

The average is a function of the block spins $\{s_j^{(b)}\}$ because the sums are restricted to correspond to a particular block spin configuration. With this definition the equation for $\mathcal{H}'$ can be written as

$$e^{\mathcal{H}'(\{s_j^{(b)}\})} = \sum_{\{s_i\}}{}' e^{\mathcal{H}_0(\{s_i\})} \langle e^{\mathcal{H}_1(\{s_j^{(b)}\})}\rangle_0. \tag{3.165}$$

The first factor in this expression is easy to evaluate since it contains terms only within one block. Therefore,

s_{j1}	s_{j2}	s_{j3}	$s_j^{(b)}$	$\exp[\mathcal{H}_0]$
+1	+1	+1	+1	$\exp(3K + 3\beta h)$
+1	+1	-1	+1	$\exp(-K + \beta h)$
+1	-1	+1	+1	$\exp(-K + \beta h)$
-1	+1	+1	+1	$\exp(-K + \beta h)$
+1	-1	-1	-1	$\exp(-K - \beta h)$
-1	+1	-1	-1	$\exp(-K - \beta h)$
-1	-1	+1	-1	$\exp(-K - \beta h)$
-1	-1	-1	-1	$\exp(3K - 3\beta h)$

TABLE III. Evaluation of $z_0(s_j^{(b)})$.

$$e^{\mathcal{H}_0(\{s_i\})} = \prod_j e^{[K(s_{j1}s_{j2}+s_{j2}s_{j3}+s_{j3}s_{j1})+\beta h(s_{j1}+s_{j2}+s_{j3})]} \tag{3.166}$$

and

$$Z_0(\{s_j^{(b)}\}) = \sum_{\{s_i\}}{}' e^{\mathcal{H}_0(\{s_i\})} =$$

$$\prod_j \sum_{\substack{s_{j1}=\pm1 \\ s_{j2}=\pm1 \\ s_{j3}=\pm1}} \delta[\mathrm{sign}(s_{j1}+s_{j2}+s_{j3}), s_j^{(b)}] e^{[K(s_{j1}s_{j2}+s_{j2}s_{j3}+s_{j3}s_{j1})+\beta h(s_{j1}+s_{j2}+s_{j3})]} \tag{3.167}$$

where $\delta(i,j)$ is the Kronecker delta. Here $Z_0(\{s_j^{(b)}\}) = \prod_j z_0(s_j^{(b)})$ is defined as an analog to the partition function. The sum is easiest to evaluate if we tabulate all the possible combinations of s_{j1}, s_{j2} and s_{j3} and the corresponding $s_j^{(b)}$. This is done in Table III. From the table we can read

$$z_0(s_j^{(b)}) = \exp[3K + 3s_j^{(b)}\beta h] + 3\exp[-K + s_j^{(b)}\beta h]. \tag{3.168}$$

For simplicity we will now concentrate on the case of no external magnetic field so that $z_0(s_j^{(b)}) = \exp[3K] + 3\exp[-K]$ is independent of $s_j^{(b)}$ and $Z_0 = [\exp(3K) + 3\exp(-K)]^{N/3}$ where $N/3$ is the number of blocks (N is the number of spins).

It remains to evaluate the average $\langle e^{\mathcal{H}_1}\rangle_0$. We do that using the *cumulant expansion*: we have

$$\begin{aligned} \langle e^{\mathcal{H}_1}\rangle_0 &= 1 + \langle\mathcal{H}_1\rangle_0 + \tfrac{1}{2}\langle\mathcal{H}_1^2\rangle_0 + \ldots \\ &= \exp\left[\langle\mathcal{H}_1\rangle_0 + \tfrac{1}{2}\left(\langle\mathcal{H}_1^2\rangle_0 - \langle\mathcal{H}_1\rangle_0^2\right) + \ldots\right]. \end{aligned} \tag{3.169}$$

The last line is called a cumulant expansion, and the first term in the exponent is called the first cumulant, the second one is the second cumulant *etc.*. The cumulant expansion is frequently much more accurate than the simple Taylor expansion. We will be satisfied with the first order cumulant expansion so we only need to evaluate $\langle\mathcal{H}_1\rangle_0$. The only operators that appear in $\mathcal{H}_1$ are products of two spins, and therefore we must evaluate terms like $\langle s_{j1}s_{j'2}\rangle_0$. Since the average is performed relative to $\mathcal{H}_0$ which does not couple different blocks, we have simply $\langle s_{j1}s_{j'2}\rangle_0 = \langle s_{j1}\rangle_0\langle s_{j'2}\rangle_0$. The averages $\langle s_{j1}\rangle_0$ can be read from Table III, which gives (for $h = 0$)

$$\langle s_{j1}\rangle_0 = \frac{1}{z_0}s_j^{(b)}\left[1\times\exp(3K) + (1+1-1)\times\exp(-K)\right] = s_j^{(b)}\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}. \tag{3.170}$$

The averages of s_{j2} and s_{j3} are obtained similarly. Inserting this into the expression for $\mathcal{H}_1$ gives

$$\langle\mathcal{H}_1(\{s_j^{(b)}\})\rangle_0 = 2K\sum_{\langle j,j'\rangle} s_j^{(b)}s_{j'}^{(b)}\left(\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}\right)^2. \tag{3.171}$$

Here the factor two arises since in our expression for $\mathcal{H}_1$ there are six products of two spins which describe the interactions between the block j and its nearest neighbors. Thus, each pair of nearest neighbors is coupled by two interaction lines. Hence, we have

$$e^{\mathcal{H}'} = [e^{3K}+3e^{-K}]^{N/3}e^{2K\sum_{\langle j,j'\rangle}s_j^{(b)}s_{j'}^{(b)}\left(\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}\right)^2}. \tag{3.172}$$

Taking the logarithms gives

$$\mathcal{H}' = \frac{N}{3}\log[e^{3K}+3e^{-K}] + 2K\left(\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}\right)^2\sum_{\langle j,j'\rangle}s_j^{(b)}s_{j'}^{(b)}. \tag{3.173}$$

The first term is an additive constant which we are not interested in, and the second term looks like an Ising coupling between the block spins. This is exactly what we hoped to get, and now we just read off the renormalized coupling constant

$$K' = 2K\left(\frac{e^{3K}+e^{-K}}{e^{3K}+3e^{-K}}\right)^2 \tag{3.174}$$

which is our RG equation.

The fixed point K^* is given by

$$\frac{1}{\sqrt{2}} = \frac{e^{3K^*}+e^{-K^*}}{e^{3K^*}+3e^{-K^*}} \tag{3.175}$$

which gives $e^{4K^*} = 2\sqrt{2}+1$ or $K^* = \frac{1}{4}\log[2\sqrt{2}+1] \approx 0.336$ or $T_c \approx 2.98J/k_B$. The exact result is $K_c = \frac{1}{4}\log[3] \approx 0.275$ so we are off by about 20%. To calculate the critical exponents we must differentiate the expression (3.174) with respect to K. This is most conveniently done with Mathematica or some other symbolic manipulation software, which gives

$$\frac{dK'}{dK} = 2\frac{(1+e^{4K})(3+4e^{4K}+e^{8K}+16e^{4K}K)}{(3+e^{4K})^3} \tag{3.176}$$

which yields at $K = K^*$

$$\lambda_t = \frac{dK'}{dK} = 2(1+\sqrt{2})\frac{4 + 3\sqrt{2} + \log(1+2\sqrt{2})}{(2+\sqrt{2})^3} \approx 1.62352. \qquad (3.177)$$

Recalling $\Lambda = \sqrt{3}$ and $\lambda_t = \Lambda^{y_t}$ gives $y_t \approx 0.882203$ or $\nu = 1/y_t \approx 1.13353$ and $\alpha = 2 - \nu d \approx -0.27$. This is not in very good agreement with the exact result $\alpha = 0$.

An obvious way to improve the analysis is to include the next term in the cumulant expansion. That, however, gets rather complicated since the second order cumulant introduces interactions between next-nearest neighbor blocks and third-nearest neighbor blocks. The analysis can nevertheless be carried through and the result is $K_c \approx 0.2575$ and $\alpha \approx 0.081$, which is a significant improvement over the first order result. However, conceptually the possibility of improving the calculation systematically is a major improvement over the previous theories: we now have a tool, position space renormalization, that allows us to analyze most critical phenomena. On the level that we have discussed the tool is still rather rough and actual calculations can get rather involved. Most practical applications of the renormalization group method are carried out in wavevector space rather than in position space, but the general idea is the same: careful elimination of those degrees of freedom that correspond to small scale structure (*i.e.* short length scales or large wavevectors).

APPENDIX A: ANALOGIES BETWEEN THE LIQUID-GAS PHASE TRANSITION AND THE ISING PHASE TRANSITION

The most straightforward way to see the analogy between the liquid-gas phase transition and the Ising phase transition is to consider the phase diagrams of the two systems. In the liquid-gas system the relevant variables are temperature, pressure, and density. The phase diagram of the system is sketched in Fig. 31(a). The solid lines are isotherms describing the systems behavior if the temperature is kept fixed and either the pressure or density (or equivalently, volume) is varied. At a large pressure for $T < T_c$ the system is in a large density liquid state. Once the pressure is reduced, the density of the liquid becomes smaller (in practice this change is quite small since liquids are almost incompressible). At some point, which is known as the equilibrium vapor pressure and is indicated by the dashed line in the Figure, there is a discontinuity in the density: for pressures slightly higher than this value all material in the system is liquid, for slightly lower pressures all material is gas, and precisely at the equilibrium vapor pressure the two phases coexist. The equilibrium vapor pressure is a function of temperature. Above the critical temperature T_c there is no phase transition — the density of the material decreases smoothly as the pressure is lowered. This implies that there is no fundamental difference between the states that we traditionally call liquid and gas — a more pedantic term for both of them is fluid. Indeed, the only way to unambiguously characterize a fluid as either "gas" or "liquid" is to prepare the system in the region of the parameter space where the two states coexist (i.e. set pressure equal to equilibrium vapor pressure), in which case the phase with the lower density is called gas and the other phase is called liquid. The property that distinguishes the two states is the order parameter. We could choose the order parameter to be any (monotonic) function of the density; the most simple choice is just a linear function. It is conventional to choose the

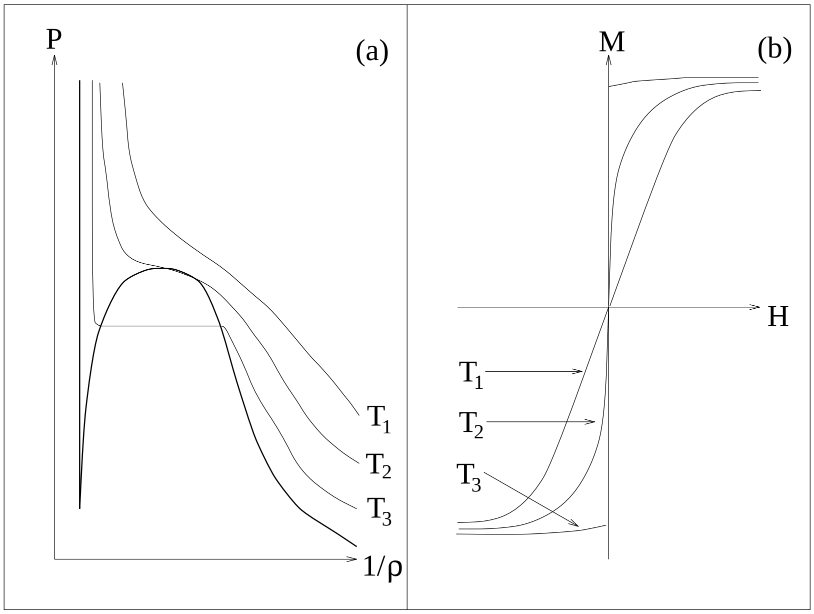


FIG. 31. Phase diagrams for the liquid-gas system (a), and the Ising ferromagnet (b). The thick line in Figure (a) indicates the boundary of the two-phase region. Here $T_1 > T_c$, $T_2 = T_c$, and $T_3 < T_c$.

order parameter so that it vanishes at the critical point, and therefore we define the order parameter to be $\rho - \rho_c$ where ρ_c is the density of the fluid at the critical point. For $T < T_c$ the order parameter can assume two values ($\rho_{\text{liquid}} - \rho_c$ and $\rho_{\text{gas}} - \rho_c$), whereas above the critical temperature it is single valued. The fact that $\rho - \rho_c$ is not zero for $T > T_c$ leads to some minor complications in the derivation of a Landau-type expansion for the liquid-gas transition [13].

Sometimes it is more convenient to consider the phase diagram in the control parameter space. For the liquid-gas system there are two control parameters, temperature and pressure. The phase diagram in the control parameter space is depicted in Fig. 32(a) where the line indicates the coexistence curve, i.e. the equilibrium vapor pressure as a function of temperature. The line terminates at the critical point where the difference between gas and liquid ceases to exist.

Now that we have considered the liquid-gas transition in some detail, let us turn to the Ising magnet. For a ferromagnet the relevant variables are magnetization per volume, m, external magnetic field H and temperature T. The isotherms connecting m and H are shown in Fig. 31(b). At first this diagram looks quite different from the corresponding liquid-gas diagram n Fig. 31(a), in particular, it appears that there is no two-phase region. This is, however, only an illusion: for $H > 0$ the magnetization at low temperatures is positive, $m > 0$, and for $H < 0$ it is negative, and for precisely $H = 0$ the system separates into domains which have nonzero magnetization $|m|$ but the direction of the magnetization of each domain is arbitrary (up or down). Thus, there is a two-phase region even in the magnetic problem, but since "coexistence field" is temperature independent, the region is not visible in the diagram — it is just the segment of the $H = 0$ line that is a part of an isotherm.

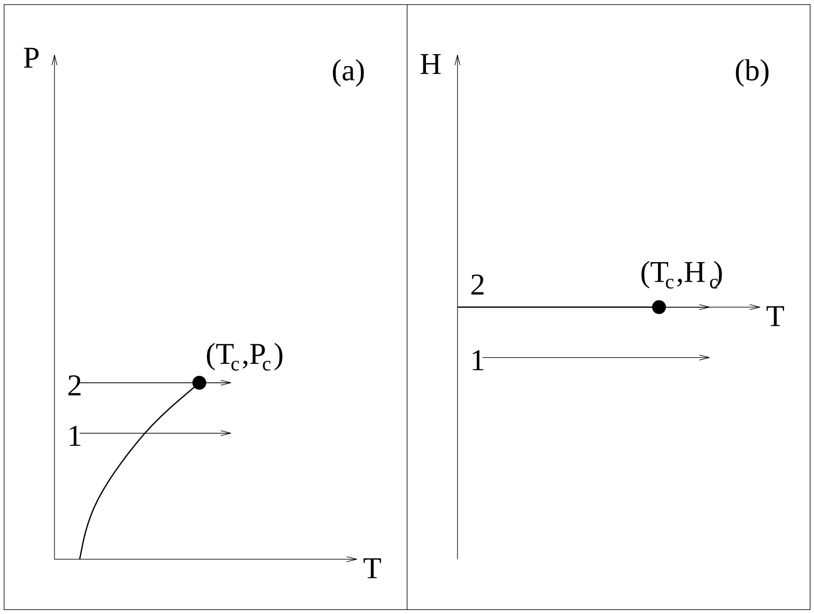


FIG. 32. Phase diagrams in the control parameter space for the liquid-gas system (a), and the Ising ferromagnet (b). The thick lines indicate coexistence curves, and the thin lines indicate typical experiments. Note that in Figure (b) path 2 is not visible as it lies on top of the $H = 0$ axis.

Now we can identify the order parameter as the difference between the magnetization and the magnetization at critical point. Since the magnetization at the critical point vanishes, the order parameter is simply given by m. The liquid and gas phases correspond to the two states at $T < T_c$, thus, we can identify liquid as the $m > 0$ state and gas with the $m < 0$ state or vice versa.

The phase diagram of the Ising magnet in the control parameter space is shown in Fig. 32(b). Since the coexistence field does not depend on temperature, the coexistence line is parallel to the temperature axis. In Fig. 32 I have also indicated paths in the parameter space that correspond to common experimental arrangements. Path number 1 corresponds to an experiment at a constant pressure or magnetic field. We see that since the coexistence curve in the liquid-gas system is not parallel to the temperature axis, an experiment of this type usually results in crossing the coexistence line, and therefore a discontinuous (first order) phase transition. The magnetic system, in contrast, has no phase transitions on the corresponding path. If, however, we fix the pressure (magnetic field) to the critical pressure P_c (critical magnetic field H_c), the path passes through the critical point and there is a continuous (second order) phase transition in both systems (path number 2).

The control parameter that couples directly to the order parameter is called the conjugate field. In the liquid-gas system the conjugate field is pressure, and in the magnetic system it is the external magnetic field. The derivative of the order parameter with respect to the conjugate field is a generalized susceptibility; in the magnetic system it is the usual isothermal magnetic susceptibility $\chi_T = \left(\frac{\partial m}{\partial H}\right)_T$, and in the liquid-gas system it is related to the isothermal compressibility $\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T$. Because of the analogy between the two systems, we know that for the continuous phase transition at the critical point the isothermal compressibility behaves similarly to the magnetic susceptibility, and in particular their critical exponents are equal. The divergence of κ_T at the critical point implies that there are large density fluctuations in the fluid (remember that the fluctuation-dissipation theorem connects fluctuations to a response coefficient; for density fluctuations we have $\langle(\rho - \langle\rho\rangle)^2\rangle = k_B T \kappa_T$). Using the definition of the static structure factor we can show that $\lim_{q\to 0} S(\vec{q}) = \rho k_B T \kappa_T$ and therefore the small-wavelength structure factor diverges at the critical point. This has a very physical consequence: in a light scattering experiment, the intensity of scattered light is proportional to $S(\vec{q})$, and we have the result that near the critical point the liquid-gas system scatters light very strongly. The phenomenon is called critical opalescence.

Although the continuous phase transition in the liquid-gas system is similar to the continuous phase transition in the Ising magnet, the discontinuous phase transitions in the two systems differ. The reason is that the entropy is much lower in the liquid phase than in the gas phase, and therefore there is a latent heat associated with the first order transition. The entropies of the $m > 0$ and $m < 0$ phases are, in contrast, identical due to symmetry, and there is no latent heat associated with crossing the coexistence line in the magnetic system.

REFERENCES

[1] K. H. Fischer and J. A. Hertz, *Spin Glasses*, (Cambridge University Press, Cambridge, 1991).

[2] C. W. J. Beenakker and H. van Houten, Solid State Physics **44**, Eds. H. Ehrenreich and D. Turnbull (Academic, Boston, 1991).

[3] Y. Imry, in *Directions in Condensed Matter Physics*, Vol. 1, Eds. G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986).

[4] M. Kardar, Lectures on Statistical Mechanics (8.334), (Massachusetts Institute of Technology, unpublished, 1988).

[5] K. Huang, *Statistical Physics*, (Wiley, New York, 1987).

[6] K. Shimoda, *Introduction to Laser Physics*, (Springer, Berlin, 1984).

[7] G. D. Mahan, *Many-Particle Physics*, (Plenum, New York, 1990).

[8] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems*, (McGraw-Hill, New York, 1971).

[9] J. R. Waldram, *The theory of thermodynamics* (Cambridge University Press, Cambridge, 1985).

[10] M. N. Barber, in *Phase Transitions and Critical Phenomena*, Vol. 8, eds. C. Domb and J. L. Lebowitz (Academic, London, 1983).

[11] N. D. Mermin, Reviews of Modern Physics **51**, 591 (1979).

[12] S. Kobe, in *Actas: Noveno Taller Sur de Fisica del Solido, Misión Boroa, Chile, 26–29 April 1995* (Universidad de la Frontera, Temuco, Chile, 1995) (`cond-mat/9605174`).

[13] P. M. Chaikin and T. C. Lubensky, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, 1995).